摘要

摘要

人脸识别研究的核心问题是使计算机具有身份辨别的能力。该研究具 有重大的理论意义和实际应用价值。人脸识别技术的系统研究能够推动人工 智能和计算机视觉等相关领域的发展。作为生物特征识别关键技术之一的人 脸识别技术则在身份认证、视觉监控、人机交互、公共安全、信息安全、金 融等领域具有潜在的应用前景。

经过几十年的发展,人脸识别技术取得了长足的进步,在理想情况下 好的人脸识别系统已经能够取得可以接受的识别性能。但非理想条件下的人 脸识别技术还远远未能满足人们的需要。本文重点探讨实现鲁棒人脸识别系 统的一些核心技术和关键问题,对面部特征精确配准、特征提取和表示、非 线性判别分析等方面进行专门研究。本文的主要研究内容总结如下:

1、研究面部特征精确配准问题,提出基于约束形状模型的面部特征 提取方法

面部特征精确配准是实现鲁棒的人脸识别系统的基本前提。本文首先 对面部特征点进行分类,对不同类型的特征点采取不同的建模和搜索方法, 从而提出一种约束形状模型(Constrait Shape Model, CSM)。该模型的要点包 括:(1)面部特征点分成边缘点,控制点两部分。(2)对于边缘点采取改进的 主动形状模型进行建模的策略,即考虑边缘强度信息对边缘点建模的贡献。 (3)对于控制点采用Gabor特征进行建模,以及基于启发式的搜索策略。实验 表明,CSM能够更精确定位面部关键特征。

2、提出基于 Gabor 象限特征的人脸识别方法

目前,Gabor特征在人脸识别领域获得了广泛应用,然而大多数人脸识别系统是基于幅值特征的。本文提出一种基于Gabor象限特征的鲁棒人脸识别方法(Histogram of Gabor Phase Pattern, HGPP)。该方法具有如下特点:(1)大多数传统人脸识别方法都是基于Gabor幅值特征,而HGPP方法完全基于Gabor相位特征。(2)提出一种新的局部异或算子(Local XOR Pattern, LXP)用来提取局部象限变化。(3)HGPP方法无需训练,直接提取直方图特征,因而避免了统计学习领域的泛化分析问题。

3. 对局部异或算子进行一般性扩展

本文提出的局部异或算子(LXP)成功的应用到 Gabor 象限特征上,并且 取得了非常好的性能。但是我们发现该算子仅仅适用于二值运算,针对这一 问题,本文提出一种对其进行一般性扩展的方法,即该算子可以适用于灰度 图像。基于此,进一步融合 Gabor 的幅值和相位信息,本文又提出基于 Gabor 特征的局部异或直方图(Histogram of Gabor-based Local XOR Pattern, HGLXP)方法。该方法具有如下特点: (1)根据不同的梯度方向对人脸图像进 行二值化。(2)LXP 算子被用来反映相邻点之间梯度方向的一致性。(3)LXP 算子成功和 Gabor 幅值和相位特征进行结合,最后提取空间直方图特征作为 判别特征。(4)HGLXP 方法无需训练,直接提取直方图特征,因而避免统计 学习领域的泛化分析问题。

4. 证明核费舍尔和支持向量机之间的内在联系,提出了基于支持向量的非线性判别分析方法

核费舍尔和支持向量机是核技巧的两种成功的应用,它们之间具有内在 联系。本文证明了二者的内在联系,即支持向量机的法向量具有零空间特 性。基于该特性,本文提出了基于支持向量的判别分析方法,实验表明该方 法是非常有效的。

5. 提出基于直方图特征的集成核费舍尔判别方法

直方图特征被成功的应用到人脸识别领域,并且无需训练,从而避免 了泛化分析问题,然而训练数据集合中包含大量先验知识没有得到应用,所 以本文提出一种应用先验知识的基于核费舍尔的统计学习方法,来进一步提 高人脸识别系统的性能。

本论文的上述内容探讨了实现鲁棒的人脸识别系统的一些关键问题, 尤其在面部特征精确配准、特征提取和表示、非线性判别分析方面进行了深 入研究,具有一定参考价值和借鉴意义。

关键词: 人脸识别, 主动形状模型, 核分析, 支持向量机, 直方图, 局部 异或模式。

Π

Abstract

The key issue of face recognition is how to make computer identify the specific person, and it is very significant in theoritic and practical aspects. It can speed up the development of the Artificial Intelligence and Computer Vision. At the same time, face recognition as one of important biometric recognition technologies is of wide potential applications, such as identity authentication, surveillance, human and computer interaction, public security, financial security, and so on.

Face recognition has made a big progress in the past thirty years, and the performance can be accepted in some good conditions. However, it can not satisfy the need of human being in the uncontrolled environment. This thesis focuses on the key technologies and problems to implement a robust face recognition system. The contributions of the thesis are:

1. Studied the facial feature extraction problem, and proposed a Constrained Shape Model(CSM)

Facial feature extraction is one of basic problems for robust face recognition system. The thesis first re-grouped facial feature points, for which different modeling and searching methods are used, and then proposed a Constrained Shape Model(CSM) to locate the facial feature points. The features of CSM include: (1) The facial feature points are divided into the edge and control parts. (2) A modified ASM is used to model the edge part, which considers the contribution of the edge information. (3)Gabor feature is used to model the control points, and a heuristic strategy is used to find the candidate points. Experimental results show that CSM can robustly extract the facial feature.

2. Initially proposed Gabor Quadrant based face recognition method.

Gabor feature has been widely used in the field of object recognition, however, most face recognition systems are based on the Gabor magnitude part. This thesis proposed a Gabor phase based method for face recognition, named Histogram of Gabor Phase Pattern(HGPP). The features of HGPP are (1) Different from traditional face recognition methods exploiting the Gabor magnitude information, HGPP is completely based on the Gabor phase information. (2)Local XOR Pattern(LXP) is proposed to capture local phase variations. (3)HGPP does not need the train procedure, and directly extracts the histogram feature, therefore, avoids the generalization problem existing in the statistic-based methods.

3. Generalization of the Local XOR Pattern(LXP) operator

LXP proposed in the thesis has been successfully applied to the Gabor quadrant information, and it has achieved a good performance. However, it seems that LXP can be only a binary operator, and this thesis proposed a generalization method for LXP, which can be used on the grey-level images. The Gabor magnitude and phase are then combined with LXP, and the Gabor-based Local XOR Pattern histograms(HGLXP) is finally proposed for face recognition. The features of HGLXP are (1) Binary images for face recognition are calculated based on different gradient directions. (2) LXP is used to reveal whether the neighbors has the same directions.(3)LXP has been successfully combined with Gabor magnitude and phase to extract spatial histograms as the discriminant feature.(4)HGLXP does not need the training procedure, and the histogram feature is directly extracted to avoid the generalization problem.

3. The relationship between kernel fisher and support vector machine was proven, and then this thesis proposed the margin-based nonlinear discriminant analysis method.

Kernel fisher analysis and Support Vector Machine(SVM) are two successful applications of the kernel trick. In this thesis, the relationship between them has been proven, that is, the normal vector of SVM is of nullspace property in terms of the support vectors based within-class scatter matrix. The margin-based discriminant analysis method is then proposed for face recognition, and experimental results show that it is very efficient.

4. Proposed a histogram and ensemble-based kernel fisher analysis method for face recognition

The histogram feature has been successfully used in the field of face recognition without the training procedure, therefore, avoiding the generalization problem. However, the background information contained in the training database offered by the experimental protocal is not exploited, therefore, this thesis proposed a kernel fisher method to use the background information contained in the training database, and enhanced the performance of the face recognition system.

This thesis makes research about some key problems to implement a robust face recognition system, especially in the aspects of facial feature extraction, feature representation and nonlinear discriminant analysis, which are helpful for other researchers.

Key words: Face Recognition, Active Shape Model, Kernel analysis, Support Vector Machine, Histogram, Local XOR Pattern.

摘要	I
Abstract	III
第1章 绪论	1
1.1 人脸识别的研究背景和意义	1
1.2 人脸识别问题描述	1
1.3 人脸识别作为生物特征识别的优势和不足	2
1.4 人脸识别研究综述	5
1.4.1 人脸识别主要技术方法概述	5
1.4.2 国内外主要公共人脸图像数据库及 FERET 测试	12
1.4.3 人脸识别领域的技术发展趋势	16
1.5 本文工作的主要路线	16
1.6 本文的主要贡献	17
1.7 本文的组织结构及各章节间的相互关系	
第2章 基于约束形状模型的面部特征提取	20
2.1 前人工作简介	20
2.2 统计形状模型	22
2.3 Gabor 变换原理	23
2.4 基于边缘约束和 Gabor 相位估计的约束形状模型	25
2.4.1 边缘点搜索方法	25
2.4.2 控制点搜索方法	
2.4.3 面部特征点的提取算法	
2.5 实验结果	27
2.6 本章小结	29
第 3 章 基于 Gabor 相位信息的特征提取和识别方法	
3.1 人脸识别中判别特征提取的回顾	
3.1.1 基于统计学习方法的判别特征提取	
3.1.2 非统计学习方法判别特征提取方法	
3.1.3 小结	
3.2 人脸识别中 Gabor 特征表示的简述	
3.3 基于空间直方图的局部 Gabor 相位编码模式	
3.3.1 Gabor 相位角的象限量化	
3.3.2 局部 Gabor 相位编码模式	

3.3.3 全局 Gabor 相位编码模式	40
3.3.4 基于空间直方图的 Gabor 相位模式	41
3.3.5 进一步分析	42
3.4 基于 HGPP 的人脸识别	43
3.4.1 基于直方图的交的匹配方法	43
3.4.2 基于 Fisher 原则的加权方法	44
3.5 实验分析	45
3.5.1 基本 FERET 测试	45
3.5.2 CAS-PEAL-R1 数据库	53
3.6 本章小结	54
第4章 局部异或算子的一般性扩展	55
4.1 局部二值模式(LBP)	55
4.2 局部异或模式(LXP)	56
4.3 基于梯度方向二值化的 LXP 模式	56
4.4 基于 LXP 的空间直方图	62
4.5 进一步分析	62
4.6 基于 LXP 的人脸识别	62
4.7 实验分析	63
4.7.1 基于灰度图像的 LBP 和 LXP 的实验对比	63
4.7.2 基于 Gabor 特征的 LBP 和 LXP 的实验对比	64
4.7.3 空间直方图特征的降维	65
4.7.4 基于基本 FERET 测试的实验对比	66
4.7.5 基于 CAS-PEAL-R1 实验对比	67
4.8 本章小结	67
第5章 基于核学习的人脸识别方法	69
5.1 基于子空间人脸识别方法	69
5.1.1 主元分析方法	70
5.1.2 线性判别分析方法	71
5.1.3 基于核技术的子空间分析方法	75
5.1.4 独立元分析方法	78
5.1.5 总结	79
5.2 支持向量机的零空间特性	79
5.2.1 支持向量机	

5.2.2 支持向量机的零空间特性	81
5.3 基于 SVM 的零空间核判别分析方法	83
5.3.1 决策边界特征矩阵	84
5.3.2 基于 SVM 的零空间判别分析	
5.4 基于最近邻样本空间的 Fisher 判别分析方法	
55 基于核费舍尔判别分析和直方图特征的人脸识别方法	87
56 实验分析	89
561 基于灰度图像性能对比	
5.6.2 基于 HGPP 的核学习方法性能测试	
5.6.3 基于 HGLXP 的核学习方法性能测试	92
5.7 本章小结	94
第6章 面部特征定位和基于人脸识别的原型系统	96
6.1 面部特征定位系统	96
6.2 照片的半自动标注系统	97
6.3 基于人脸识别的考勤系统	
6.4 本章小结	
结论	104
参考文献	106
个人简历	
哈尔滨工业大学博士学位论文原创性声明	
哈尔滨工业大学博士学位论文使用授权书	
致谢	
· · · · · · · · · · · · · · · · · · ·	

Contents

Chinese Abstract	I
Abstract	III
Chapter 1 Introduction	1
1.1 Backgroud of face recognition	1
1.2 Introduction of face recognition	1
1.3 Metrits and demetrits of face recognition	2
1.4 Introduction of face recognition	5
1.4.1 Main face recognition methods	5
1.4.2 Main face databases and FERET protocal	12
1.4.3 Future of face recognition	16
1.5 Main strategies	16
1.6 Main contributions	17
1.7 Organization of the Thesis	
Chapter 2 Facial feature localization based Constrainted Shape Mod	del20
2.1 Previous work	20
2.2 Statistic shape model	22
2.3 Theory of Gabor transformation	23
2.4 Constrainted Shape Model based on edge constraint and Gal	oor phase
information	
2.4.1 Searching strategy for edge points	25
2.4.2 Searching strategy for control points	26
2.4.3 Algorithm of facial feature localization	
2.5 Experimental results	
2.6 Conclusions and future work	
Chapter 3 Faical Feature Extraction and Face Recogniton based of	on Gabor
Phase	
3.1 Review work about the discriminant feature extraction for face r	ecogntion

3.1.1 Discriminant feature extraction based on statistic-based meth	ods30
3.1.2 Discriminant feature extraction based on non-statistic based	d methods
3.1.3 Conclusions	
3.2 Reviews about the Gabor representation	34
3.3 Gabor Phase Pattern for the spatial histogram	
3.3.1 Qudrant Bit Coding of the Gabor phase	
3.3.2 Local Gabor Phase Pattern	
3.3.3 Global Gabor Phase Pattern	40
3.3.4 Spatial Histogram for Gabor Phase Pattern	41
3.3.5 Further analysis	42
3.4 Face recognition based on HGPP	43
3.4.1 Face recognition based on histogram intersection	43
4.4.2 Weighting strategy based on the Fisher principal	44
3.5 Experimental analysis	45
3.5.1 Standard FERET protocal	45
3.5.2 CAS-PEAL-R1 protocal	53
3.6 Conclusions and future work	54
Chapter 4 Generalization of Local XOR Pattern operator	55
4.1 Local Binary Pattern(LBP)	55
4.2 Local XOR Pattern(LXP)	56
4.3 LXP based on gradient direction binary results	56
4.4 Spatial histogram based on LXP	62
4.5 Further analysis	62
4.6 Face recognition based on LXP	62
4.7 Experimental Analysis	63
4.7.1 Experimental comparisons of LBP and LXP based on	grey-level
images	63
4.7.2 Experimental comparisons of LBP and LXP based on t	he Gabor
feature	64
4.7.3 Feature reduction of spatial histogram	65
4.7.4 Experimental results based on the standard FERET database.	66
5.7.5 Experimental results based on CAS-PEAL-R1 database	67

4.8 Conclusions and future work	67
Chapter 5 Face recognition based Kernel Fisher analysis	69
5.1 Face recognition based subspace method	69
5.1.1 Principal component analysis	70
5.1.2 Linear discriminant analysis	71
5.1.3 Kernel Fisher analysis	75
5.1.4 Independent Component Analysis	78
5.1.5 Conclusions	79
5.2 Nullspace property of Support Vector Machine(SVM)	79
5.2.1 Support Vector Machine	80
5.2.2 Nullspace of SVM	81
5.3 SVM-based nullspace kernel discriminant analysis for face recognition	.83
5.3.1 Decision boundary feature matrix	84
5.3.2 Nullspace discriminant analysis based on SVM	84
5.4 Nearest Neighbor-based Fisher discriminant analysis	86
5.5 Face recognition based on the combination of kernel Fisher analysis	and
spatial histogram	87
5.6 Experimental analysis	89
5.6.1 Experimental comparison based on the grey-level images	89
5.6.2 Performance of HGPP-based kernel discriminant analysis for	face
recognition	90
5.6.3 Performance of HGLXP-based kernel discriminant analysis for	face
recognition	92
5.7 Conclusions and future work	94
Chapter 6 Demo System based on Facial Feature Localization and l	Face
Recongnition	95
6.1 Facial Feature Localization system	95
6.2 Semi-automatic annotation of pictures	6
6.3 Time Card system based on face recognition	99
6.4 Conclusions	. 101
Conclusion	.103
References	.105
Papers publised in the period of PH.D. Education	.115

哈尔滨工业大学博士学位论文

Resume	
Letter of Authentification	
Statement of Copyright	
Acknowledgement	

图表目录

图 1-1 节点
图 1-2 适应对象的网格标记图
图 1-3 人脸的束图表示
图 1-4 统计识别方法模型
图 2-1 面部控制点
图 2-2 面部特征提取结果, a,b)具有一定光照的面部特征的提取结果 c)表
情变化情况下的面部特征的提取结果 d)正常人脸的面部特征的提取结果
图 3-1 Gabor 幅值特征可视化结果
图 3-2 Gabor 相部分可视化结果
图 3-3 Daugman 的相位角的象限量化示意图
图 3-4 局部异或运算算子示意图
图 3-5 a) 实部的局部 Gabor 相位编码模式(Real-LGPP) b)虚部的局部
Gabor 相位编码模式(Real-LGPP)
图 3-6 不同的人脸图像提取的局部 Gabor 相位编码模式40
图 3-7 模式分块方法示意图
图 3-8 a) 实部的全局 Gabor 相位编码模式(Real-GGPP) b)虚部的全局
Gabor 相位编码模式(Real-GGPP)
图 3-9 HGPP 万法的流程图
图 3-10 FERET 数据库的样本45
图 3-11 Gabor 核的幅值部分 f_{max} . a) $f_{\text{max}} = \pi$; b) $f_{\text{max}} = \sqrt{2} / 2\pi$, and
c) $f_{\text{max}} = \pi$
图 3-12 正面图像
图 3-13 第一行人脸图像向左旋转 2,6,10 度,第二行表示人脸图像向右旋
转 2,6,10 度
图 3-14 前两行表示人脸图像在 Y 左平移 2,3,4, 第三行和第四行表示人
脸图像在 X 轴上进行平移 2,3,4 像素

哈尔滨工业大学博士学位论文

图 3-15 在 FB 测试集合上验证 HGPP 对旋转的鲁棒性 (2 表示向左	旋转
2度,-2表示向右旋转2度)	
图 3-16 在 FB 测试集合上验证 HGPP 对 X 轴上的平移的鲁棒性(例	」如 3
表示向左平移3个像素,-3表示向右平移3个像素)	49
图 3-17 在 FB 测试集合上验证 HGPP 对 Y 轴上的平移的鲁棒性(例	」如 3
表示向左平移3个像素,-3表示向右平移3个像素)	49
图 3-18 子块的个数对识别率的影响	50
图 3-19 不同的直方图 Bins 对识别率的影响	51
图 3-20 CAS-PEAL 数据库上样本	53
图 4-1 局部二值模式(LBP)算子的示意图	55
图 4-2 LXP 的含义示意图 a) '1'表示的含义(LXP) 0 表示的含义	57
图 4-3 基于 Gabor(5 个尺度 8 个方向)幅值特征的 LXP 示意图 a)表示	梯度
方向是 0 度时 LXP 结果 b)表示梯度方向是 45 度时 LXP 结果 c)表示	梯度
方向是 90 度时 LXP 结果 d)表示梯度方向是 135 度时 LXP 结果	59
图 4-4 基于 Gabor(5 个尺度 8 个方向)相位特征的 LXP 示意图 a)表示	梯度
方向是 0 度时 LXP 结果 b)表示梯度方向是 45 度时 LXP 结果 c)表示	梯度
方向是 90 度时 LXP 结果 d)表示梯度方向是 135 度时 LXP 结果	61
图 5-1 x_1, x_2 是支持向量, $h'(x)$ 是法向量, Margin 是类问距	80
图 5-2 NN-FID 的示音图	87
图 5-3 CAS-PEAL-R1 数据库上的性能对比	91
图 5-4 子中数对识别家的影响(CAS-PEAL-R1)	91
图 5-5 FFRFT 数据库上的性能对比	92
图 5-6 百方图 Bins 数对识别率的影响	93
图 5-7 CAS-PEAL-R1 数据库上的性能对比	93
图 5-8 FERET 数据库上的性能对比	94
图 6-1 面部特征定位系统	95
图 6-2 基于视频的面部特征定位系统	
图 6-3 单张标注图片	97
图 6-4 标注第一张图片	98
图 6-5 自动标注后续图片	98
图 6-6 自动标注(聚类)	99
网人,其工人队门则的北热乏众	100

图表目录

表	1-1	生物特征识别方法比较	3
表	2-1	实验结果对比	29
表	2-2	不同频段对误差精度的影响	29
表	3-1	不同图像大小对识别率的影响	47
表	3-2	各种局部 Gabor 相位编码模式的首选识别率对比	51
表	3-3	和各种著名的人脸识别系统的对比结果(FERET)	52
表	3-4	CAS-PEAL-R1 数据库首选识别率的对比	53
表	4-1	基于灰度图像的 LBP 和 LXP 结果对比(16 Bins)	63
表	4-2	基于 Sobel 算子的 LBP 和 LXP 结果对比	64
表	4-3	基于 Gabor 特征的 LBP 和 LXP 结果对比	64
表	4-4	基于 Gabor 幅值部分的 LXP 特征(单一方向)和 LGBPHS 的对比	65
表	4-5	Image_LXP 的直方图 Bins 数和识别率之间的关系	66
表	4-6	GHLXP 的直方图 Bins 数和识别率的关系	66
表	4-7	和各种著名的人脸识别系统的对比结果(FERET)	67
表	4-8	和各种著名的人脸识别系统的对比结果(CAS-PEAL-R1)	67
表	5-1	CAS-PEAL-R1 数据库的识别结果(r=2)	89

List of Figures and Tables

Fig.1-1 Jet	7
Fig.1-2 Object Adapted Grids	7
Fig.1-3 Face bunch graph	8
Fig.1-4 Model of statistic-based recognition method	9
Fig.2-1 Control points	2
Fig.2-2 Samples of facial feature extraction results a,b)Samples with	h
illumination changes, c) Samples with expression changes d) Samples in th	e
normal condition	8
Fig.3-1 Visualization of Gaborfaces(Magnitude part)	5
Fig.3-2 Visualization of Gabor Phase	7
Fig.3-3 Illustration of the Quandrant Bit Coding method	7
Fig.3-4 Illustration of Local XOR Pattern Operator	8
Fig.3-5 a) Real Part Local Gabor Phase Pattern b) Imaginery Part Local Gabo	r
Phase Pattern	9
Fig.3-6 LGPPs for Different faces	0
Fig.3-7 Illustration of the division method for Gabor Phase Pattern4	1
Fig.3-8 a) Real Part Global Gabor Phase Pattern b) Imaginery Part Globa	ıl
Gabor Phase Pattern	1
Fig.3-9 Flow chart of HGPP	2
Fig. 3-10 Samples of the FERET database	5
Fig.3-11 Magnitude Part of Gabor kernels. a) $f_{\text{max}} = \pi/2$; b) $f_{\text{max}} = \sqrt{2}\pi/2$, an	d
c) $f_{\text{max}} = \pi$	6
Fig.3-12 Frontal faces	7
Fig.3-13 The first row with 2,6,10 degrees rotation, the second row with	h
2,6,10 degrees rotation	7
Fig.3-14 The top two rows with X-translation 2,3,4 pixels, the two bottom	n
rows with Y-translation 2,3,4 pixels	8

Fig.3-15 Robustness of HGPP against rotation on FB probe(2 means 2 degress
Left rotation, -2 means 2 degress Right rotation)
Fig.3-16 Robustness of HGPP against X-translation on FB probe(3 means 2
pixels Left translation, -3 means 3 pixels Right translation in the X-direction)
Fig.3-17 Robustness of HGPP against Y-translation on FB probe(3 means 3
pixels Left translation, -3 means 3 pixels Right translation in the Y-direction)
Fig.3-18 Recognition rates for different number of sub-regions
Fig.3-19 Recognition rates for different number of histogram bins
Fig.3-20. Samples of the CAS-PEAL database
Fig.4-1 Illustration of Local Binary Pattern(LBP) operator
Fig.4-2 Illustration of LXP, a) the meaning for '1' b) the meaning for '0'57
Fig.4-3 Illustration of Gabor Magnitude based LXP results(5 scales, 8
orientations), a) LXP results for 0-degree gradient direction b) LXP results for
45-degree gradient direction c) LXP results for 90-degree gradient direction d)
LXP results for 135-degree gradient direction
Fig.4-4 Illustration of Gabor Phase based LXP rsults(5 scales, 8 orientations),
a) LXP results for 0-degree gradient direction b) LXP results for 45-degree
gradient direction c) LXP results for 90-degree gradient direction d) LXP
results for 135-degree gradient direction
Fig.5-1 x_1, x_2 are support vectors, $h'(x)$ is the normal vector, Margin is the
distance between different classes
Fig.5-2 Illustration of NN-FLD
Fig.5-3 Comparison of the performance on CAS-PEAL-R1 database91
Fig.5-4 Recognition rates for different number of sub-regions(CAS-PEAL-R1)
Fig.5-5 Comparison of the performance on the FERET database92
Fig.5-6 Recognition rates for different number histogram bins
Fig.5-7 Comparison of the performance on the CAS-PEAL-R1 database93
Fig.5-8 Comparison of the performance on the FERET database94
Fig.6-1 Facial feature localization system
Fig.6-2 Video-based facial feature localization system

哈尔滨工业大学博士学位论文

Fig.6-3Annotating the picture one by one	97				
Fig.6-4Annotation of the first picture					
Fig.6-5Automatic annotating other pictures	98				
Fig.6-6Automatic annotating the pictures(Clustering)	199				
Fig.6-7 Time Card system based on face recognition	100				
Table 1-1 Comparison of biometrics technologies	3				
Table 2-1 Experimental result comparisons	29				
Table 2-2 Error for different frequencies	29				
Table 3-1 Recognition rates for different sizes of basic HGPP	47				
Table 3-2 Rank-1 recognition rate comparisons for different kinds of C	Jabor				
Phase Patterns	51				
Table 3-3 Recognition rate comparisons with other well-known	face				
recognition systems	52				
Table 3-4 Rank-1 recognition rate comparison on the CAS-PEAL-R1 data	abase				
	53				
Table 4-1 Comparison results for LBP and LXP based on grey-level image	es(16				
bins)	63				
Table 4-2 Comparison results of LBP and LXP based the Sobel operator	64				
Table 4-3. Comparison results of LBP and LXP based on the Gabor feature	e64				
Table 4-4 Comparison results of LXP(single direction) and LGBPHS base	ed on				
the Gabor magnitude	65				
Table 4-5 The relationship between the number of histogram bins	and				
recognition rates for Image LXP	66				
Table 4-6 The relationship between the number of histogram bins	and				
recognition rates for GHLXP	66				
Table 4-7 Comparison results with other well-known face recognition sys	stems				
	67				
Table 4-8 Comparison results with other well-known face recogn	nition				
systems(CAS-PEAL-R1)	67				
Table 5-1 Experimental results for the CAS-PEAL-R1 database(r=2)	89				

第1章 绪论

1.1 人脸识别的研究背景和意义

本项研究工作部分得到了国家自然科学基金重点项目"基于生物特征的 身份识别研究"(批准号: 60332010)、中科院百人计划、上海市科委项目(项 目号: 03DZ15013)、以及上海银晨智能识别科技有限公司的大力资助。人 脸识别技术主要通过计算机分析人脸图像,从人脸图像中提取有效信息,从 而进行身份识别。人脸识别技术因其在公安(罪犯识别等)、安全验证系统、 信用卡验证、医学、档案管理、视频会议、人机交互系统等方面的巨大应用 前景而成为当前模式识别和人工智能领域的一个研究热点^{[1][2]}。虽然人类能 毫不费力地识别出人脸及其表情,但人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大 的课题。人脸有复杂的表面结构,因而人脸的自动机器识别却是一个难度极大

人脸识别具有广泛的应用前景,尤其在"911"事件之后,安全和反恐问题日益得到各国政府的重视,而生物特征识别技术被认为是满足这一需要的最好方案之一。人脸识别技术作为生物特征识别领域的热门方向之一,可以 广泛的应用于安全、视频监控、人机交互、电子银行等诸多应用系统中。

1.2 人脸识别问题描述

人脸识别(Face Recognition),实际上是指计算机以人的输入图像或视频序列作为特征的身份识别,是生物特征识别的一个分支。为了准确辨别人的身份,计算机需要从中提取出有效的识别信息,并利用预先存储的人脸数据,识别其中一人或多人的身份。简单的说,人脸识别的任务就是使计算机具有人类所具有的身份辨别能力,为了完成这一任务,一般来说一个人脸识别系统应该包括如下四部分:

1. 人脸检测与跟踪

人脸检测的主要任务是判断输入图像或者视频序列中是否存在人脸。 如果存在,则需确定其所在的位置和大小。人脸检测是实现人脸识别完全自 动化最为关键的一步,而人脸跟踪则进一步确定人脸的位置、大小的变化信 息。

2. 面部特征定位

面部特征定位的主要任务是给出面部显著特征包括眼睛、鼻子、嘴巴 等器官的位置信息,同时可对各个特征的轮廓信息作出描述。

3. 特征提取/选择

特征提取/选择作为人脸识别中最为重要的一部分,直接决定了最终的 识别性能。该步骤的主要任务是提出或者选择出那些能够区分不同人的差异 性而对同一个人具有稳定性的特征或者度量。

4. 分类器设计

分类器的主要任务是根据提取的特征确定某人的身份,即解决输入图 像是人脸数据库中的哪一个人的问题,一般来说首先需要对输入人脸是否在 给定的数据库中进行判断;分类器的另外一个任务是解决身份确认问题,即 系统需要对输入图像的身份与其所宣称的身份是否符合作出判断。

实际应用系统还应该重视输入图像的预处理问题,它对识别算法的性 能影响不可低估。

1.3 人脸识别作为生物特征识别的优势和不足

身份识别与验证是人类社会日常生活中的基本活动之一。传统的身份 识别方法主要包括两个方面:利用身份标识物品,比如证件、钥匙等;或利 用身份标识知识,比如帐号、密码等。然而这些方式都具有不方便、不可 靠、不安全等特点,比如证件可以丢失,密码可以忘记等等。传统身份识别 的缺点以及人们对高可靠身份验证技术的需求促进了生物特征识别的快速发 展。同时,生物特征识别也被认为是解决身份认证问题的终极解决办法。

所谓生物特征识别或生物测定学(Biometrics)在这里特指生物特征识别,即利用计算机技术提取人类所固有的生理特征或行为特征,并与预先存储在数据库中的数据进行比较,以达到个人身份识别目的的自动化技术^[3]。 生理特征通常指与生俱来,先天性的特征,即要满足普遍性、独特性、稳定 性等;行为特征则是习惯使然,后天性的特征,两者统称为生物特征。生理 第1章 绪论

特征^[4]包括视网膜、虹膜、指纹、掌纹、语音、脸像等;行为特征则包括声纹、笔迹、步态等。基于生物特征识别的身份识别技术由于利用了人的某些生物特征的独特性和不变性,因此比传统的身份识别方法更加准确、可靠;同时由于生物特征始终与个人密切相关、不可分离,因此使用起来也更加安全、方便。当然,这些技术所需的软硬件系统也都更为复杂。另外,不同的生物特征有各自不同的特点和实用性,表 1-1 从七个方面对它们进行了对比^[4]。

生物特征	普遍性	唯一性	稳定性	可采集性	准确性	可接受性	防伪性
脸像	高	低	中	高	低	高	低
指纹	中	高	一同	中	高	中	高
虹膜	高	高	高	中	高	低	高
视网膜	一同	高	中	低	高	低	高
语音	中	低	低	中	低	高	低
红外温谱	高	高	低	高	中	高	高
步态	中	低	低	高	低	高	中
耳形	中	中	高	中	中	高	中
笔迹	中	中	中	高	中	中	中
足迹	中	中	中	低	中	中	低
掌纹	高	中	中	中	中	中	高

表 1-1 生物特征识别方法比较

Table 1-1 Comparison of biometrics technologies

如表 1-1 所示,人脸与虹膜、指纹等特征相比更容易受到各种因素的 干扰,但是人脸识别技术仍然有其巨大的发展空间。人脸识别技术在可用性 方面具有如下三个技术优势:

1. 可操作性强

人脸识别技术最大的优点就是自然、友好、无侵犯性,它不需要被识 别者的主动合作,甚至可以在被识别者不知情的情况下进行,这在某些场合 (例如安全监控)是非常重要的;

2. 性价比高

人脸识别技术所需的硬件设备较为简单,因此系统成本较低^[5];而且 在识别的基础上,通过对人脸的表情、姿态加以分析,还能获得其他识别方 法难以获得的一些信息。因此,人脸识别技术近年来越来越受到研究人员的 广泛关注。

3. 可接受性强

与其他的生物特征,如指纹、虹膜识别等等进行对比,人脸识别更容 易被人接受。

人脸识别技术作为一种重要的基于生物特征识别的身份识别技术,同 时又是计算机视觉、模式识别和图像处理等领域研究的热点问题之一,每年 都有大量科技论文出现在相关领域的重要国际会议(CVPR,ICCV,ICPR, ACM MM 等等)上,而且还出现了 AFGR 和 AVBPA 等专业性的国际会议和 一些专题著作^[6,7]。人脸识别技术的发展在某种程度上推动了人工智能、计 算机视觉等领域的发展。在实际应用中,计算机软硬件的快速发展在计算能 力和经济成本上保证了人脸识别技术的可行性,使其广泛应用成为可能。在 司法验证、安全监控、档案管理、人机交互等方面的巨大应用前景的推动 下,几十年来相关领域的理论研究成果使人脸识别技术不断完善,也越来越 多地引起了人们的广泛关注。同时也涌现出大量的方法,详细内容可参看综 述性文献[1,2]。人脸识别技术的相关研究虽然已经取得了很大的进展,但是 在人脸识别系统的实用化方面,其准确率主要还受到以下几方面因素的影 啊:

(1)面部特征的复杂性以及图像信息的不完整性:人脸具有极强的可塑性所以具有不确定性,而无论 2D 还是 3D 图像所获取的信息都具有不完整性。

(2) 姿态(Pose):由于采集图像的视角变化是一个非常困难的视觉问题,所以经常导致姿态上的变化,其中包括姿态平面内旋转,深度旋转等等。人脸识别必须解决该问题,才能使其性能保持稳定。

(3)表情(Expression):一个人的心情是不断变化的,而"貌由心生"导致人的面部表情不可能保持稳定。人脸识别相对于其他物体识别问题的一个困难就是人脸的非刚体性,主要体现为人脸的面部表情变化非常大。

(4) 光照(Illumination): 光照的方向及强度变化会严重影响人脸图像的纹理特征,因而影响大多数基于表观的识别方法的效果;对于基于几何特

第1章 绪论

征的识别方法,光照同样会严重影响几何特征的准确提取。

(5) 遮挡(Occlusion/Accessory): 人脸识别的另一个特殊困难之处是 人脸受到各种各样遮挡的影响,例如:头发,胡须,眼镜,帽子等。

(6)性别和年龄(Gender/Aging):不同性别的人对于人脸识别影响也 是不一样的,女性由于喜欢化妆等原因导致了识别上的困难。对年轻人的识 别要较老年人困难,由于年轻人会较多的改变发型,生理变化等等,而老年 人相对稳定。无论男性,还是女性,年龄的增长会使得人脸识别的难度增 大。

(7) 其他: 整容、病变、创伤等所造成的脸部细节的变化等。

1.4 人脸识别研究综述

正如前文所述,人脸识别技术走向实际应用还需要解决很多实际问 题。最近十年来,人脸识别技术取得了飞速发展,其中Gabor特征获得了非 常成功的应用,成为占主导地位的特征提取方法之一^[8-11]。另外还有一些通 过提取人脸形状特征来进行人脸识别的方法,其中主动形状模型(ASM)最有 影响力^[12]。统计学习的发展同样促进了人脸识别技术的发展,比如线性判 别分析方法^[10]、非线性判别分析方法^[13]。而核方法^[14-17]作为目前统计学习 领域最成功的方案之一是当前的研究热点。事实上,早在1965年Chan & Bledsoe 就在Panoramic Research Inc.发表了人脸识别技术报告^[18],到现在已 有三十多年的历史。人脸识别作为一个经典模式识别问题,得到了诸多研究 人员的青睐,涌现出了诸多技术方法。尤其是1990年以来,人脸识别更得到 了长足的发展,每年都有大量的学术论文发表,现在,几乎所有知名的理工 科大学和IT 产业的主要公司都有研究组在从事人脸识别的研究。每年都有 大量的研究论文发表,然而这在某种程度上给研究者带来麻烦。对人脸识别 的历史有一个清晰地认识可以更好地把握自己的研究方法,从而避免盲目的 尝试。为了更好的了解人脸识别研究的历史和现状,下面对人脸识别技术的 研究进展情况作一下简单介绍。

1.4.1 人脸识别主要技术方法概述

人脸识别的发展是一个渐变的过程,每一个阶段都有其代表性的方法。这些方法对于研究人员来说是非常有意义和启发性的。不但一些经典的 模式识别理论和方法被应用到人脸识别问题中,人们还对大量新的理论和方 法做了有益的尝试。其中较为成熟和广泛应用的方法,按照特征提取和选择 方法的不同,以及出现的时间顺序,大致可以分为四大类:

- 几何特征方法(Geometric Feature-based Approaches)
- 弹性束图匹配方法(Elastic Bunch Graph Matching Method)
- 统计学习方法(Statistic Approaches)
- 非统计学习方法(Non-statistic Approach)

这里只是根据作者的理解简单地进行了分类,这样的分类不是绝对 的,几大类方法之间也不是互斥的,其中统计学习方法还可以进一步细分。 而非统计学习的方法也是本文的研究重点之一,将在第3章,第4章进行系 统地分析。所以,本章将简单介绍这四大类方法,并提供相关文献以供进一 步研究。

1.4.1.1 几何特征方法

人脸识别最初只是从生物学上最直观的特征出发,即基于几何特征的 人脸识别方法,并且在很长时间内占据着主流地位,文献[18-20]对几何特 征方法做了详尽的综述。这种方法的基本思想是提取人脸上突出特征(例如 眉毛,眼睛,鼻子,嘴巴等)的相对位置和相对大小作为特征参数,再辅以 人脸轮廓的形状信息,形成特征向量,最后利用适当的距离度量和分类方法 对特征向量加以分类。

几何特征方法利用的是单纯的几何信息,其优点是所需的存储空间和 分类时间代价较小,而且在图像分辨率较低的情况下仍然可以使用。其缺点 是识别的准确率完全依赖于几何特征的准确提取,而几何特征提取问题直到 目前仍然没有完全解决,例如光照和遮挡的变化可能严重干扰几何特征的准 确提取;而且几何特征本身容易受姿态和表情变化的影响,稳定性不高,即 使几何特征提取准确,仍然可能因为有较大的变化而导致分类错误。较早从 事人脸识别研究的研究人员除了 Bledsoe 外还有 Goldstein,Harmon 以及 Kanade^[20]等。Kanade 博士于 1973 年在京都大学完成了第一篇人脸识别方 面的博士论文^[20]。文献[21]中比较了几何特征方法和模版匹配(Template Matching)方法,并且指出,至少在该文中所采用的算法实现上,模版匹配 方法优于几何特征方法。因此目前单纯利用几何特征的人脸识别方法已经不 常使用了,但几何特征仍然可以作为其他特征的有益补充。

1.4.1.2 弹性束图匹配方法

在1996年进行的FERET(FacE Recognition technology Test)^[22]人脸识别 测试中,南加州大学的弹性束图匹配方法^[9]取得了最好的成绩。该方法基于

第1章 绪论

动态链接结构(Dynamic Link Architecture)^[8]的弹性图束图匹配(Elastic Graph Bunch Matching)方法,也被一些研究人员归为神经网络方法^[23]。其基本思想是把人脸图像表示成标记图(Labeled Graph),用标记图之间的相似度来衡量人脸图像之间的相似度。标记图是由一组边连接起来的节点构成的,如图 1-1所示^[8]。其中,边被标记为节点间的距离,节点则被标记为一个特征向量,称为Jet,在[8]中,采用Gabor小波系数的相位信息标记Jet,使节点定位更加准确。具体而言,在[9]中的Jet是由Gabor小波变换提取的幅值特征构成的。采用适应对象的网格(Object Adapted Grids)标记图取代规则网格标记图,使节点对应于人脸上突出的特征点(例如眼睛、鼻子、嘴巴及某些轮廓点等)。如图1-2所示^[8],弹性束图更准确地包含了人脸的几何特征信息;采用束图(Bunch Graph)取代单一标记图来进行弹性图匹配,如图1-3所示^[8]。 束图在标记图的基础上,在每个节点上不止标记一个而是标记一束Jet,这些Jet来自于不同的图像,因此每个Jet表示了该节点的一种变化,通过在不同的节点上选择不同的Jet,可以得到不同的标记图,这样束图的表达能力就比标记图强得多。



图 1-1 节点 Fig.1-1 Jet



图 1-2 适应对象的网格标记图

Fig.1-2 Object adapted grids

标记图的边可以度量人脸的几何特征,而节点则表达了人脸图像Gabor

特征的变化。人脸图像之间的相似度由标记图之间的相似度来衡量,而标记 图之间的相似度则由各节点的特征向量(Jet)的相似度和标记图相对形变的程 度决定。所谓弹性束图匹配,就是将待识别图像的标记图进行平移、旋转、 放缩及弹性形变以增加生成的标记图与已知图像的标记图的相似度的过程。 标记图的平移、旋转、放缩及弹性形变就包含了人脸的各种变化。该方法的 优点是既保留了面部的全局结构特征,又对人脸的关键局部特征进行了建 模。另外,近年来还出现了一些对该方法的扩展^[24,25]。



图 1-3 人脸的束图表示 Fig.1-3 Face bunch graph

1.4.1.3 统计学习方法

统计学习方法最近几十年来取得了非常迅速的发展,其中支持向量机 从理论上通过 VC-维和结构风险最小化原则讨论了泛化问题。另外,它也越 来越多的应用到实际问题中,在模式识别领域,统计学习方法也为人脸识别 的发展起到极大的推动作用。

在最近二十年来,统计学习方法在人脸识别领域一直占据着统治地 位。基于统计学习方法的人脸识别具有如下优点:(1)不依赖于人脸的先验 知识和参数模型。(2)采用基于训练集合的方法获取参数,具有统计意义上 的可靠性。(3)可以通过扩充训练集合达到扩充识别模式的范围,从而提高 系统的鲁棒性。统计方法是人脸识别方法中一类重要的方法,虽然出现较 晚,但已经占据了重要的地位,本节介绍几种常见的统计学习方法。

统计学习方法模型





子空间分析方法

子空间分析(Subspace Analysis)方法的基本思想是将待分类样本看作空 间中的点,首先寻找高维原空间的一个低维子空间,将原样本投影到这个子 空间上再进行分类。麻省理工学院媒体实验室的 Turk 和 Pentland 提出的 "Eigenface"方法^[26,27]是早期人脸识别领域最具影响力的方法之一。该方法是 目前占主导地位的统计学习方法的核心部分之一,其后的很多人脸识别技术 都或多或少与 Eigenface 有关系^[28-32],现在 Eigenface 已经成为人脸识别技术 都或多或少与 Eigenface 有关系^[28-32],现在 Eigenface 已经成为人脸识别的 性能测试基准算法。由于统计学习在机器学习领域获得了快速发展,也直接 推动了其在人脸识别领域的应用。基于表观(Appearance-based)的线性子空 间建模和基于统计模式识别技术的人脸识别方法的发展——他们逐渐成为主 流的人脸识别技术。Belhumeur 等提出的 Fisherface^[26]人脸识别方法是这一 时期的另一重要成果。该方法首先采用主成分分析(Principal Component Analysis, PCA,亦即 Eigenface)对输入图像进行降维,在此基础上,采用 线性判别分析(Linear Discriminant Analysis,LDA)的方法变换降维后的主成分 以期获得"尽量大的类间散度和尽量小的类内散度",该方法目前仍然是主流 的人脸识别方法之一,产生了很多不同的变种,比如零空间法^[33],增强判 别分析模型^[31],直接的 LDA 判别方法^[34,35],以及近期的一些基于核学习的 改进策略^[36-38]。MIT 的 Moghaddam 则在 Eigenface 的基础上,提出了基于 双子空间进行贝叶斯概率估计的人脸识别方法^[29,30]。该方法通过"作差法", 将两幅人脸图像对的相似度计算问题转换为一个两类(类内差和类间差)分类 问题,类内差和类间差数据都要首先通过主成分分析(PCA)技术进行降维, 计算两个类别的类条件概率密度,最后通过贝叶斯决策(最大似然或者最大 后验概率)的方法来进行人脸识别。除此以外,还有独立元分析(Independent Component Analysis/ICA)方法^[39-41],以及基于核技术的子空间分析(Kernelbased Subspace Analysis)方法^[37-38]等。值得一提的是,核判别分析在这一阶 段取得了非常好的识别率,刘^[42]在 FRGC 的最难的测试集合上取得最好的 识别性能。他的方法主要是利用分数 KPCA 方法结合线性判别分析方法, 该方法在输入空间是一种非线性分析方法,但是在特征空间却是线性的。刘 [43]验证了非正定的核函数同样能够得到非常好的识别性能,所以扩展了核 方法的应用范围。

这部分工作也是本文的工作重点之一,将在第 5 章中详细介绍各种子 空间分析方法,并加以比较和评价。

隐马尔科夫模型方法

用于描述信号统计特性的隐马尔科夫模型(Hidden Markov Model/HMM) 也被用到了人脸识别领域。隐马尔科夫模型应用于人脸识别的基础在于人脸 的主要特征(例如头发、前额、眼睛、鼻子、嘴巴等)的位置顺序固定,因此 可以表示为隐马尔科夫模型中的一系列状态。这方面较早的工作是由 F. Samaria 等人完成的^[44],最近的推广及改进见于[45,46]。特别值得一提的 是,著名的 OpenCV 系统中提供的人脸识别程序采用的就是嵌入式隐马尔 科夫模型。

支持向量机方法

支持向量机(Support Vector Machine)是近年来兴起的源于统计学习理论,针对一般性模式识别问题的有效方法。其基本思想是通过一个隐函数实现从输入空间到高维空间的映射,而隐函数可以通过核函数来实现。统计学

第1章 绪论

习理论是针对小样本情况研究统计学习规律的理论,是传统统计学习的重要 发展和补充,为研究有限样本情况下机器学习的理论和方法提供了理论框 架,其核心思想是通过控制学习机器的容量实现对推广能力的控制。由 Vapnic 等人提出的支持向量机是一种新的通用基于统计学习机制的分类 器,较以往的方法表现出很多理论和实践上的优势^[14]。由于支持向量机主 要用于解决两类模式分类问题,因此支持向量机应用到人脸识别上,其关键 在于如何构造出一个两类问题,人们已经提出了一些策略来解决这个问题, 比如类内差/类间差,一对多以及一对一等等。对于人脸识别问题,一般也 采用这些策略,具体可参考[47-49]。

其他统计方法

除了上述统计方法之外,LFA^[50]在本质上是一种基于统计学习方法的 低维对象描述方法,与只能提取全局特征而不能保留局部拓扑结构的 PCA 相比,LFA 在全局 PCA 描述的基础上提取的特征是局部的,并能够同时 保留全局拓扑信息,从而具有更佳的描述和判别能力。另外,主动形状模型 (Active ShapeModels, ASMs)^[12]和主动表观模型(Active Appearance Models, AAMs)^[51]是这一时期内在人脸建模方面的一个重要贡献;Adaboost 方法在 人脸识别领域也获得广泛的应用^[11,52,53],其基本思想是通过一系列的弱分类 器构建强分类器。另外,还有最新的流形学习的方法^[54,55]用来模拟人脸图像 的分布,但是在实际应用中,该方法还在不断的发展中。

1.4.1.4 非统计学习方法

此外,最近几年来一些非统计学习的人脸识别方法也受到了越来越多 研究者的关注。其中,直方图特征^[56,57]被成功地应用到人脸识别领域,取得 了非常好的实验结果。如 Ahonen^[57]提出的基于局部二值模式(Local Binary Pattern, LBP)直方图特征的人脸识别方法,他认为人脸图像是由 LBP 为模式 来组成的,从而提出一种具有强纹理描述能力的算子用于人脸识别。原始的 LBP 算子通过对图像的每个像素的 8 邻域点的亮度比较进行编码。如果周 围像素点大于中心点像素值,则中心点在这个邻域位置上编码为 1,否则为 0。8 个邻域正好构成一个字节作为微模式。最后空间直方图用来统计这些 微模式出现的频次作为判别特征,取得非常好的识别性能。张 ^[56]把 LBP 和 Gabor 幅值信息进行了结合(Local Gabor Binary Pattern Histogram Sequence, LGBPHS),在 FERET 数据库上取得了最好的识别性能。这一类方法最大的特点是无需训练,所以可以避免泛化分析问题。LBP 特征由于其先进的思想和良好的效果从一开始就吸引了很多人的关注,现在已经有人把它和Adaboost^[54]、Fisher 原则^[56]进行结合用于人脸识别,并取得了一定的效果。

1.4.2 国内外主要公共人脸图像数据库及 FERET 测试

人脸数据库对算法的研究与开发、模型训练、算法性能测试和比较是 不可或缺的。而性能评测则是在一个共同的测试协议下对各种不同算法性能 进行比较、特定算法改进和模型参数的调整、从而发现问题进而提出新算法 的重要环节。本节将介绍国内外主要公共人脸图像数据库的情况,以及 FERET、FRVT、FRGC 等人脸识别性能评测的情况^[21,58,59],借此分析人脸 识别技术当前的最高学术水平以及面临的主要难题。

1.4.2.1 国内外主要公共人脸图像数据库

人脸数据库对算法的研究与开发、模型训练、算法测试是最为重要的 一部分之一。大部分人脸识别算法都涉及到有训练过程参与的参数选择问 题,因而训练所采用的人脸库的人数规模、数据采集的条件变化在很大程度 上影响着算法的精度和鲁棒性能;而算法性能测试所用到的人脸库的规模和 特点同样决定了评测的合理性和测试结果的有效性。下面简单介绍人脸识别 领域经常被采用的人脸图像数据库的情况。

1. FERET人脸数据库^[21]

FERET项目创建,包含14,051幅多姿态、不同光照条件的灰度人脸图像。其公布了一个最有影响力的测试协议,本文称之为基本FERET测试协议。该协议严格划分了训练集合、Gallery、不同的测试集合等,是人脸识别领域应用最广泛的人脸数据库之一,参见:

http://www.itl.nist.gov/iad/humanid/feret/feret_master.htmlo

2. PIE 人脸数据库^[60]

由美国卡耐基梅隆大学创建,包含68 位志愿者的41,368 张多姿态、 光照和表情的面部图像。其中的姿态和光照变化图像也是在严格控制的条件 下采集的。详情参见: http://www.ri.cmu.edu/projects/project 418.html。

3. BANCA 人脸库^[61]

人脸库是欧洲BANCA 项目资助创建的,包含208 人的多模态生物特征,覆盖了不同图像质量、不同时间段等变化条件。该人脸库划分了不同的

训练、测试条件。详情参见: <u>http://www.ee.surrey.uk/research/vssp/banca/</u>。

4. CAS-PEAL 人脸数据库^[62]

人脸库包含了1,040 名中国人共99,450 幅头肩部图像。所有图像在专门的采集环境中采集,涵盖了姿态、表情、饰物和光照四种主要变化条件,部分人脸图像具有背景、距离和时间跨度的变化。类似于FERET数据库,他们也公布了一个测试协议,即CAS-PEAL-R1协议。该测试数据库逐渐地被人脸识别研究人员所接受。介绍该人脸库的网址为: http://www.jdl.ac.cn/peal/index.html。

5. ORL 人脸数据库^[63]

剑桥大学AT&T 实验室创建,包含40人共400 张面部图像,部分志愿 者的图像包括了姿态、表情和面部饰物的变化。该人脸库在人脸识别研究的 早期经常被人们采用,但由于变化模式较少,多数系统的识别率均可以达到 90%以上。

6. XM2VTS 多模态生物特征数据库^[64]

数据库包括295 人在4 个月时间内4 次录制的人脸和语音数据。每次 采集都包括2 个头部旋转视频片断和6种不同语音视频片断。另外,其中293 人的3D 模型也是可以得到的。

7. MIT 人脸数据库^[65]

麻省理工大学媒体实验室创建,包含16 位志愿者的2,592 张不同姿态、光照和大小的面部图像。

需要说明的是,限于篇幅所限,还有很多数据库没有列出来。

1.4.2.2 FERET 测试

由美国国防部反毒品技术发展计划办公室资助的FERET项目是人脸识别领域具有里程碑意义的事件^[22]。FERET项目的目标是要开发能够为安全、情报和执法部门使用的人脸识别技术。该项目包括三部分内容:资助若干项人脸识别研究、创建FERET人脸图像数据库、组织FERET人脸识别性能评测。该项目分别于1994 年,1995 年和1996 年组织了3 次人脸识别评测,几种最知名的人脸识别算法都参加了测试,极大的促进了这些算法的改进和实用化,包括弹性束图匹配,线性判别分析,以及贝叶斯方法等等。该测试的另一个重要贡献是给出了人脸识别的进一步发展方向:光照、姿态等非理想采集条件下的人脸识别问题逐渐成为热点的研究方向,其测试结果被公认反映了当时人脸识别技术的最高学术水平。事实上该测试结果直到2005 年还是非常好的结果,后来的基于非统计学习^[56,57]的方法取得了比1996测试

结果好得多的性能。根据2000 发表在IEEE Trans. on PAMI 上的FERET'96 测试报告,南加州大学(USC)的弹性图匹配技术、马里兰大学(UMD)的子空 间判别分析方法、麻省理工大学(MIT)的贝叶斯方法等人脸识别技术具有最 好的识别性能。测试表明,在训练集和测试集的摄像条件相似的1196 人的 识别测试(测试集FB)中,性能最好的系统的首选识别率达到了95%; 然而, 对用不同的摄像机和不同的光照条件采集的测试图像(测试集fc),最高首选 识别率骤降为82%; 而对一年后采集的图像测试(测试集Dup II),最大的准 确率仅仅接近51%。而事实上,测试集Dup I和测试集Dup II上的数据姿态和 光照变化是非常大的。这表明,当时的人脸识别算法对于不同的光照条件、 不同姿态、不同的摄像机和老化的适应能力还是比较差,是人脸识别实用化 的主要瓶颈。

1.4.2.3 FRVT2000 、 FRVT2002

FERET 项目之后,涌现了若干人脸识别商业系统,为了比较商业系统 的实用性能,为应用部门提供详实的人脸识别系统性能资料,美国 DARPA, DoS, FBI 和NIST等多家机构分别于2000 年, 2002年联合资助了 2次专门对商业人脸识别产品进行评测的项目,即Face Recognition Vendor Test(FRVT),至今已经举办了2次:FRVT2000,FRVT2002^[58]。FRVT2000 和FRVT2002测试一方面对知名的人脸识别系统进行了性能比较,例如 FRVT2002 测试就表明Cognitec, Identix 和Eyematic 三个商业产品远远好于 其他系统,而它们之间的差别不大。另一方面则全面总结了人脸识别技术发 展的现状: 较理想条件下(正面签证照), 针对37437 人121,589 幅图像的人 脸识别(Identification)最高首选识别率为73%,人脸验证(Verification)的等错 误率大约为6%。FRVT 测试的另一个重要贡献是进一步指出了目前的人脸 识别算法亟待解决的若干问题。例如, FRVT2002 测试就表明: 目前的人 脸识别商业系统的性能仍然对于室内外光照变化、姿态、时间跨度等变化条 件非常敏感, 大规模人脸库上的有效识别问题也很严重, 这些问题都仍然需 要进一步的努力。总体而言,目前非理想成像条件下(尤其是光照和姿态)、 用户不配合、大规模人脸数据库上的人脸识别问题逐渐成为研究的热点问 题。而非线性建模方法、统计学习理论、基于Boosting 的学习技术、基于 3D 模型的人脸建模与识别方法等逐渐成为备受重视的技术发展趋势。

FRVT2000 在FERET 测试的基础上又增加了对压缩、距离、表情、光照、存储媒体、姿态、分辨率和老化等不同情况的测试,结果仍然表明这些系统的性能对于距离、光照、姿态和老化的适应性很差,这些变化越大,其

第1章 绪论

性能下降的就越大。而FRVT2002^[40]的测试则更为详尽,数据库的规模也达到了37,437人的121,589幅图像,大多数知名的人脸识别商业系统公司都参加了此次评测。此次测试分为高计算强度(HCInt)和中等计算强度(MCInt)两种。其中HCInt 测试使用的数据库规模达到了37,437人的121,589幅图像,其中所有图像全部由美国国务院(DOS)下属的领事事务局(Bureau of Consular Affairs)签证服务处提供。其中,每人至少三幅标准的签证照,图像质量很高,采集环境也非常一致。而MCInt 测试使用的数据库则包括两部分,一部分来自美国国家标准和技术研究所(National Institute of Standards and Technology,NIST)、美国海军武器研制中心(Naval Surface Warfare Center)和南佛罗里达大学(University of South Florida, USF)三个机构1999年到2002年间采集的静态图像,其中既包括室内也包括室外采集的图像,最长时间跨度为三年;另一部分则是由德州大学达拉斯分校(University of Texas at Dallas)于2001年采集的序列或者静态图像。FRVT2002测试了三类不同的人脸识别任务的性能,分别是身份认证、闭集识别、开集识别,有关FRVT2002的主要测试可参考[58]。

1.4.2.4 FGRC 测试

FRGC是人脸识别领域最近的一次也是最大规模的测试,该测试总结了 人脸识别领域的若干挑战^[59],其中包括高分辨率的人脸图像、基于3D图像 的人脸识别方法、不同预处理方法对识别性能的影响。现有的人脸识别系统 都利用分辨率较低的图像,在两个眼睛之间有40-60个像素点,而FRGC所用 人脸图像的两个眼睛距离为250像素点,所以FRGC关注能够利用高分辨率 信息的人脸识别方法。3D人脸识别方法对光照和姿态都是非常鲁棒的,所 以3D人脸识别方法成为解决这两个问题一个可行的方案。另外,人脸识别 算法更容易受到光照和姿态的影响,FRGC侧重于测试不同的预处理方法对 这两种变化的效果。同时FRGC测试对目前人脸识别所面临的重大问题进行 了探讨,其中包括高分辨率图像对人脸识别的作用、2D和3D哪个更有利于 人脸识别以及各种不同的预处理方法对人脸识别性能的影响。除此以外,该 测试组织者给出了一些非常有意义的推测:

Bowyer 认为一张 3D 图像比 2D 图像更有利于人脸识别。

Philips 认为一个高分辨率的 2D 图像比一个 3D 图像更有利于人脸识别。

Bowyer 认为 4-5 通过选择的 2D 图像比一个 3D 图像更有利于人脸识别。

Philips 认为 FGRC 能够使得人脸识别研究人员重新思考人脸识别该如何进行。

从 FRGC 组织者的观点可以发现人脸识别处于瓶颈期,但是 FRGC 的 贡献在于为人脸识别研究人员指明了一些方向。

1.4.3 人脸识别领域的技术发展趋势

目前,人脸识别实用化的主要障碍来自于非理想采集条件下识别性能的快速下降。所以一个可行的方案是只考虑某种约束或可控制条件下的人脸 识别,比如光照、姿态相对稳定等等。非可控条件下人脸识别的困难来源于 很多方面,其所带来的问题既有信号层面上数据获取本身的稳定性问题,也 有识别所采用的人脸描述特征对采集条件变化的鲁棒性问题,同时还与核心 识别算法的泛化能力密切相关^[66]。因此,要解决这些问题就需要从多方面 探讨人脸识别的未来。其中,多摄像机融合识别技术、序列图像分析、基于 非线性学习的人脸建模方法、3D 变形模型和 3D 人脸识别技术、面部特征 精确配准、鲁棒的人脸表示特征、泛化能力强以及对各种变换鲁棒的分类算 法等研究课题尤其值得关注。另外,在 FRGC 测试中提到的高分辨率图像 以及预处理办法等也将成为人脸识别领域至关重要的研究方向^[59]。

1.5 本文工作的主要路线

作为视觉研究的一个重要子领域,人脸识别技术具有重要的理论意义 和实际应用价值。正如前文所述,人脸识别技术达到实际应用还面临许多困 难。一般来说人脸识别系统应该包括面部特征提取、特征表示、判别特征提 取、分类器设计等等。本文针对人脸识别中的几个关键问题进行了研究,包 括面部特征提取、特征表示,非线性判别分析等等。

近几年来,Gabor 特征在人脸识别领域获得了成功的应用。由于面部 关键特征精确定位是实现鲁棒的、实用的人脸识别系统的基本前提,本文利 用 Gabor 特征对面部部分特征点(控制点)进行建模,结合主动形状模型 (Active Shape Model,ASM)^[12]达到鲁棒定位面部特征的效果。尽管 Gabor 特 征得到非常好的识别性能,然而我们很容易发现,现有的人脸识别系统大多 数是基于 Gabor 的幅值特征,事实上,相位特征对识别是非常有利的。本文 首次提出了基于 Gabor 象限特征的人脸识别方法,并和其它方法进行了大量 实验对比。

第1章 绪论

局部特征具有对光照、遮挡、姿态等变化的鲁棒性,因而在人脸识别 领域获得了非常广泛的应用。本文提出一种新的局部异或算子,并在此基础 上提出一种新的基于直方图的特征表示方法。另外,非线性判别分析,主要 是核判别分析在人脸识别领域获得了非常广泛的应用。本文深入探讨了核 Fisher^[13]和支持向量机^[14]的内在联系,提出了基于支持向量的非线性判别分 析方法。由于本文的特征表示是基于直方图特征的,而基于直方图特征的相 似度函数已经被证明是正定的,或者在实践中被验证具有良好的应用效果 ^[15-17],所以本文进一步利用核方法来提高基于空间直方图特征的人脸识别 系统的性能。

1.6 本文的主要贡献

经过三十多年的发展,尤其是近十年来的研究,人脸识别技术已经取 得了长足的进步。目前最好的人脸识别系统在注册和认证环境条件比较一 致、用户比较配合的情况下已经能够达到令人满意的效果,对1000 人左右 的识别系统,其正确识别率可以在95%左右;验证系统的等错误率(EER)性 能也在2%以下。然而,这并不意味着人脸识别技术已经非常成熟了,恰恰 相反,因为更大量的人脸识别应用系统需要在更大规模人脸库、摄像环境不 可控、用户不配合的情况下使用,目前即使最好的识别系统在这样的情况下 识别性能也会下降得非常快,很多情况下识别系统正确识别率会有大幅度的 下降,从而导致应用系统用户无法接受。因此,现有的人脸识别系统尚未成 熟,开发出真正鲁棒、实用的人脸识别应用系统还需要解决大量的关键问 题。基于此,本文重点探讨了实现鲁棒人脸识别系统的一些核心技术和关键 问题,对面部特征精确配准、基于Gabor 相位特征的人脸表示和识别方法、 非线性判别分析方法进行了专门研究。本文的主要贡献总结如下:

1、研究面部特征定位或配准问题,提出了基于约束形状模型的面部特征提取方法。面部特征精确配准是实现实用的人脸识别系统的基本前提。本 文首先对面部特征进行合理分类,然后对不同类型的特征点采取不同的建模 和搜索方法,从而提出一种基于统计学习理论的约束形状模型(CSM)。该模 型的要点包括,面部特征分成边缘点和控制点两部分,对于边缘点采取改进 的主动形状模型进行建模的策略,即考虑了边缘强度信息对边缘点建模的贡 献。而对于控制点采用 Gabor 特征进行建模,以及基于启发式的搜索策略。 实验表明,CSM 能够更精确的定位面部关键特征。 2、研究了特征提取问题,首次提出基于 Gabor 象限特征人脸识别方法。目前,大多数的人脸识别系统中都用的是 Gabor 幅值特征,本文提出了一种完全基于 Gabor 象限特征的鲁棒人脸识别方法(Histogram of Gabor Phase Pattern, HGPP)。该方法基于一种新的局部异或算子用来提取局部象限变化,本质上反映了二值图像的纹理特性。不同于传统的统计学习方法,HGPP 方法无需训练,直接提取直方图特征,因而避免了泛化分析问题。

3. 提出基于梯度方向二值化的方法对局部异或算子进行一般性扩展。 局部异或算子(LXP)成功的应用到了 Gabor 象限特征上,取得了非常好的性能,但是我们发现该算子仅仅适用于二值运算,针对这一问题,本文提出梯度方向二值化的方法对其一般性进行扩展的方法。基于此,进一步融合 Gabor 的幅值和相位信息,本文又提出了基于 Gabor 的局部异或模式直方图 (Histogram of Gabor-based Local XOR Pattern, HGLXP)方法。该方法沿着不同的梯度方向对 Gabor 结果图像进行二值化,而 LXP 算子被用来反映邻域 点之间的梯度方向的一致性。LXP 算子成功的和 Gabor 幅值和相位特征进 行结合,最后提取空间直方图特征作为判别特征。HGLXP 方法无需训练, 直接提取直方图特征,因而避免了统计学习领域的泛化分析问题。

4. 核费舍尔和支持向量机是核技巧的两种成功的应用,本文研究了它 们之间的内在联系。证明了核费舍尔和支持向量机之间的内在联系,即支持 向量机的法向量具有零空间特性。在此基础上,提出了基于支持向量的非线 性判别分析方法,本质上该方法是在最近邻样本空间进行判别分析,实验表 明该方法是非常有效的。

5. 提出了基于直方图特征的集成核费舍尔判别方法。直方图特征在本 文中被成功的应用到人脸识别领域,并且无需训练,从而避免泛化分析问 题,然而大量先验信息没有得到应用,所以本文提出一种基于核费舍尔的统 计学习方法来利用训练集合中提供的先验信息,来进一步提高人脸识别系统 的性能。

1.7 本文的组织结构及各章节间的相互关系

第 1 章为绪论,介绍了研究背景,研究意义,前人工作概述,面临的 主要问题,主要技术路线和主要贡献。第 2 章提出基于约束形状模型进行面 部特征提取的方法。第 3 章提出基于 Gabor 相位模式的判别特征提取方法。 第 4 章提出对局部异或算子进行一般性扩展的方法。第 5 章从理论上证明了
第1章 绪论

支持向量机和核费舍尔的内在联系,并在此基础上提出基于支持向量的判别 分析方法。最后,把直方图特征和核判别分析方法进行有机的结合,从而有 效的提高了人脸识别系统的性能。第6章设计了三个原型系统,即面部特征 定位、照片的半自动标注系统、基于人脸识别技术的考勤系统。

下面对论文后续章节之间的关系进行总结。第 2 章提出一种新的方法 解决面部特征定位问题,面部特征定位是实现自动人脸识别系统的基本前 提,所以作为独立的一章。第 3 章提出一种新的特征提取方法,并且基于最 近邻分类器进行验证,从而提出一种新的人脸识别方法。第 4 章是对第 3 章 内容的扩展。第 5 章基于第 3 章和第 4 章的特征提取方法,进一步提出非线 性人脸识别方法。第 6 章基于前几章提出的方法设计了三个原型系统。

第2章 基于约束形状模型的面部特征提取

面部关键特征的自动定位是人脸识别系统的关键环节之一,在很大程 度上影响着人脸建模的精度和最终识别系统的识别性能。人脸检测领域的研 究人员往往并不关心特征定位的准确度,而大多数识别算法研究人员在做识 别实验时都假定面部关键特征点(通常是双眼中心位置)是手工标定的,而且 非常精确。反过来,面部特征定位的精确度在很大程度上影响着最终识别系 统的性能,而现有的多数特征定位算法的定位精度都随着光照、姿态等外界 条件变化而快速下降,因此,特别有必要大力研究鲁棒、准确的特征定位问 题。

广义上来说,面部关键特征定位也是特征提取的一部分,由于面部特 征定位是实现自动人脸识别系统的第一步,起到非常关键的作用,所以把这 一个课题作为独立章节来阐述。

2.1 前人工作简介

在过去的 20 年中,国内外学者们做了大量关于人脸对齐的研究。最 开始人们是基于图像处理的一些简单方法,例如滤波、边缘提取、形态学处 理。但这些算法都很不鲁棒。后来,人们意识到仅仅用一些低级特征 (LowFeature)和处理方法是不够的,应该对研究对象建立相应的先验模型, 以便鲁棒地进行人脸对齐。基于这样一种思想, Kass 等人在 1987 年提出 了称为 Snake^[67] 的主动轮廓线模型(Active Contour Model, 简称 ACM)。活 动轮廓线模型或称作蛇模型(Snake),被定义为能量达到最小的曲线,Snake 的能量根据它的形状和在图像中的位置确定,内力约束它的形状,外力引导 它的行为。不像大多数其他的模型那样, Snake 是主动(active)的, 总是最 小化它的能量函数,因此表现出一种动态特性。变形模板是一种已知先验形 状的主动轮廓线方法,对分割整体形状、克服噪声很有利。由于人脸的多样 性和复杂性以及图像的噪声等复杂因素的影响,在使用刚性模型分割人脸轮 廓时遇到了极大的困难。而 Snake 模型由于没有利用关于对象的知识,所 以过于灵活,很难做到精确的分割。我们知道,特定人脸的形状一般变化不 大,如果能够利用关于人脸形状的先验知识,那么就可能实现更准确的分 割。基于 Snake 模型的变形模板方法,为解决这一问题展现了令人鼓舞的

前景。相对于刚性模板,变形模板使自己发生变形以匹配显著的图像特征。 相对于 Snake 模型,变形模板又能够使自己的变形限制在关于人脸形状的 合理范围内。人脸对齐中的主要问题就是如何有效地利用先验知识。从上面 的分析可以看到,主动轮廓线模型只是利用关于一般对象的先验知识,没有 利用关于特定对象的先验知识,因此,在对齐精度要求高的地方不太适用。 在利用先验知识方面,变形模板对主动轮廓线模型做了改进,是一种已知先 验形状的主动轮廓线方法。因此,变形模板被用在分割眼睛和嘴巴上。然 而,使用变形模板需要人为设计参数化的几何模型,这对于形状简单的形状 规则的对象(如眼睛和嘴巴)还比较可行,但是对于形状复杂的对象,如鼻 子、眉毛和人脸外轮廓等就很麻烦了,不仅参数很多,而且能量函数也很复 杂。为了克服主动轮廓线模型和变形模板的不足,Cootes 等人提出了主动 形状模型(Active Shape Model, ASM)和主动表观模型(Active Appearance Model, AAM)。这两种模型是基于统计学习的方法来学习模型的参数。

ASM(Active Shape Models)和 AAM(Active Appearance Models)是目前解 决稀疏特征配准问题的主流方法,它们都是采用形状的点分布模型(Point Distribution Model, PDM)^[12]。ASM 将局部纹理匹配和全局形状子空间约束 融合起来,通过局部搜索和全局形状约束的交替迭代,以期收敛到一个最优 的结果。AAM 与 ASM 相似, 建立了一个融合形状和纹理于一体的外观 (Appearance)模型。通过优化外观模型的参数实现特征配准,最终目标是期 望合成的图像纹理能够最佳的匹配输入图像纹理。但是由于人脸外轮廓点邻 域的纹理信息相对较少,因此 AAM 对外轮廓的定位不十分准确,并且它依 赖于训练集合中的图像外观的变化情况,推广能力较差。而且,由于 ASM 和 AAM 本身具有的优化迭代过程同时也不可避免的导致了算法具有较高的 复杂度,并且易于陷入局部极小。基于这些算法上的缺陷,在传统的 ASM, AAM 基础上, 很多研究人员对此进行了扩展, 提出了更加实用, 鲁 棒的统计模型或改进策略。文献[68]中提出了基于不同视角视图(View)的主 动外观模型方法,通过若干不同视角外观统计模型来覆盖较大范围的视角的 变化,这些模型还可用于在大的视角变化下估计头部方向,跟踪人脸。同 时,这些不同视角的外观模型参数之间的约束关系还可以用于由一幅某个视 角的视图预测生成该人脸新视点下的图像,或者在对同一人脸的多幅图像同 时进行特征定位时作为优化过程的一个约束条件。还有一些改进的模型,包 括直接表观模型(Direct Appearance Model, DAM)^[69]、纹理约束的主动形状 模型(Texture Constraint ASM, TC-ASM)^[70]、贝叶斯切空间形状模型

(Bayesian Tangent Shape Model, BTSM)^[71]。Wiskott 等人利用 Gabor 小波 设计出一种称为弹性束图(Elastic Bunch Graph, EBG)的数据结构来定位人 脸关键点。用 Gabor Jet 来对关键点进行建模,通过 Gabor 系数对关键点进 行描述。而弹性图是用来描述人脸的先验信息的。表明人脸各个器官之间的 相对位置是受到约束的。该方法可以适应一定程度的姿势和表情的变化,经 实践证明该方法是可行的。它首先在整幅图像中查找出脸部的关键点,利用 弹性图的变化来调整人脸关键点。遗憾的是,该方法的计算过程是非常耗时 间的。

本文提出的约束形状模型是利用边缘约束信息和 Gabor 相位特征对边 缘点和控制点分别进行建模和搜索的方法,从而获得面部特征的一种方法。 对于边缘点,充分利用了边缘强度信息来建立局部特征模型。另外,对于控 制点,利用 Gabor 小波的多个频段进行搜索的方法。本文利用定量的方法来 衡量特征定位的优劣,同时进行了大量的实验,实验结果表明这种方法是非 常有效的。面部特征点的定位有很多用途,比如身份验证,表情分析,表情 动画等。弹性图匹配的方法就是利用面部特征定位的方法来获得较高的识别 精度,并且对于均匀光照以及表情具有一定的适应性。现在比较好的方法, 比如主动形状模型和主动表观模型被成功用在面部特征提取上。主动形状模 型在全局上能够很好的获得对象的形状,也就是对于边缘点是有效的^[66], 但对于面部的关键点还不能够很好的定位,其中边缘点指的是面部,眉毛, 眼睛,鼻子和唇部的边缘部分的边缘信息较强的特征点。关键点指的是控制 形状或边缘信息不强的点,如图 2-1 所示的面部控制点。



图 2-1 面部控制点 Fig.2-1 Control points

2.2 统计形状模型

本节主要讨论如何利用统计形状模型来表示非刚体对像,统计形状模

型是通过一组已经标注的特征点进行训练得到的^[12]。在训练的过程中,每 一幅图像的像素点的坐标按着相同的方式进行对齐,构成向量x。通过对这 些向量进行主成分分析,得到本征矩阵P,则统计形状模型可以用一个参数 B来决定,表示为 $x = \overline{x} + PB$,其中 \overline{x} 是一个平均形状。估计模型参数B的 过程是通过最小化重构残差得到的,即最小化 $E(B) = (x - x_0)^T (x - x_0)$,其

中 x_0 为当前的重建模型。由当前参数 B_0 估计B可以通过计算dB(差向量)得到^[12],则有:

$$dB = (p^{\mathsf{T}}p)^{-1}p^{\mathsf{T}}dx \tag{2-1}$$

其中P 是正则的:

$$dB = p^{T} dx \tag{2-2}$$

这是最基本统计形状模型的参数估计公式。差向量 dx 的获得是通过对特征 点周围的区域进行搜索得到的。对于每一点局部模型是根据它的训练数据得 到的, 它被用于在搜索过程中选择候选点。

本文将介绍一种基于统计学习方法的约束形状模型,该方法将对边缘 信息较强的特征点进行边缘约束,以及对一些关键特征点利用 Gabor 特征进 行建模。

2.3 Gabor 变换原理

1946 年, Gabor 提出了可以到达时频测不准关系下界的 Gabor 函数, 它实际上是高斯函数在频域中的平移。Gabor 函数是唯一能够到达测不准关 系下界的函数,它能够同时在时域和频域中兼顾对信号分析的分辨率要求 ^[72],Gabor 函数这一特性使它在信号处理中得到广泛的应用。二维 Gabor 函 数较好地描述了哺乳动物初级视觉系统中一对简单视觉神经元的感受特性, Daugman 最早在 80 年代将其用于计算机视觉领域^[73]。

Gabor 变换是以高斯(Gaussian)函数作为窗函数的短时傅立叶变换,它的基本思想是把信号划分成许多小的时间间隔,用傅立叶变换分析每一个时间间隔,以便确定该时间间隔存在的频率。利用了 Gabor 变换原理的 Gabor 滤波器能够同时保留空域和频域信息,在弹性图束图匹配等人脸识别方法中得到应用。Gabor 变换在具体实现上就是将 Gabor 小波和灰度图像进行卷积运算,可由下面公式表示:

$$\Psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2/2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right]$$
(2-3)

其中, ||.||代表模运算, z = (x, y)表示空域中像素的位置, σ 是高斯函数沿着 x 轴和 y 轴的标准差,即在这两个轴上具有同样的标准差, $\vec{k_{u,v}} = \binom{k_{v}}{k_{y}} = \binom{k_{v}\cos\Phi_{u}}{k_{v}\sin\Phi_{u}}, k_{v} = f_{max}/2^{v^{2}}, \Phi_{u} = u\frac{\pi}{8}, v = 0,...,4, u = 0,...,7, v 表示尺$ $度, u 表示方向,通常情况下, <math>f_{max} = 2^{*}\pi$ 。此处所指的尺度与通常意义上的 尺度含义不同,它是一个用来表示频率的量。

对一个灰度分布为I(z)的图像,在给定一点z=(x,y),Gabor 变换表示为:

$$G_{u,v}(z) = \int I(z') \Psi_{u,v}(z-z') d^2 z'$$
(2-4)

通过采用多个方向和尺度的 Gabor 小波函数,对于图像上的每一个点可以得到一组系数向量,包含 40 个变换域系数(5 个尺度,8 个方向),它可以写成 $G_{u,v} = a_{u,v} \exp(i\phi_{u,v})$,其中 $a_{u,v}$ 为幅值部分,而 $\phi_{u,v}$ 为相位角。文献[9]中介绍了两种相似度计算方法,对于系数向量G和G,如果不考虑其相位差别,则其相似度公式为:

$$S_{a}(G,G) = \frac{\sum_{u,v} a_{u,v} a_{u,v}}{\sqrt{\sum_{u,v} a_{u,v}^{2} \sum_{u,v} a_{u,v}^{2}}}$$
(2-5)

*S*_a的特点是其变换是平滑的,实际上是因为相邻两个点的 Gabor 幅值的变化不会很大。另外一种是考虑相位差别的相似度计算方法,如 2-6 所示:

$$S_{\phi}(G,G') = \frac{\sum_{u,v} a_{u,v} a_{u,v} \cos\left(\phi_{u,v} - \phi_{u,v}' - \vec{d} \vec{k}_{u,v}\right)}{\sqrt{\sum_{u,v} a_{u,v}^{2} \sum_{u,v} a_{u,v}^{2}}}$$
(2-6)

其中*ā* 是系数向量*G* 和*G*之间的估计距离,具体有关它的计算可以参考 [9]。*S*, 的特点是其变化是不平滑的,这是因为离得较近两个点的相位有时 会发生较大的变化,但是可以通过估计他们之间的距离*ā* 来弥补它的不稳定 性。如果知道两个特征点处的系数向量*G* 和*G*,它们之间的距离*ā* 是可以通 过计算得到的^[9]。本文利用这个估计距离*ā* 作为启发式,来获得候选点的位 置。

2.4 基于边缘约束和 Gabor 相位估计的约束形状模型

直觉上对面部特征点进行合理的分类是为了根据其不同的特点进行局 部建模。这里把它们分成边缘点和控制点两部分。边缘点,就是人脸、眼 睛、鼻子、嘴的边缘部分。而那些控制着形状和不具有太多边缘信息的点称 为控制点,如图 2-1 所示。本节采用基于边缘约束和 Gabor 相位约束的统 计形状模型对面部特征点进行定位,下面将具体介绍一下该方法。

2.4.1 边缘点搜索方法

在每个边缘点的法线方向上取一定长度范围内的像素,然后这些像素 亮度的一阶导数进行采样,一般假定它是满足高斯分布的。利用邻域一阶导 数建模的局部模型很容易陷入局部极小,尤其是面部轮廓点,甚至经常匹配 到灰度均匀的面部或者背景上^[66]。注意到面部轮廓上的特征点在语义层面 上应该是边缘点,本文将这一特点加以利用以增强面部轮廓点的定位精度, 对于边缘点选择相应的建模方法:

$$f(g_{s}) = (c-k)(g_{s}-g)^{T} S_{g}^{-1}(g_{s}-g)$$
(2-7)

其中g是 $\{g_i\}(1 \le i \le N)$ 的平均值, g_i 是当前待搜索特征点的特征, N为样本空间的大小, S_i 为协方差矩阵, C是一个常量(本文为300), k是利用Sobel

算子计算得到的。选取候选点时根据最小化 $f(g_s)$ 的原则, 而k 使我们倾向于选择边缘信息更强的特征点。

2.4.2 控制点搜索方法

对于一幅人脸图像,我们先不考虑面部特征点的空间相对位置关系, 这里单独对每一个特征点进行调整,具体步骤如下:

1. 设置 fre=0; 表示我们计算相位估计值时仅考虑 0 频段;

2. 对于每个特征点,我们首先计算它所在位置 P_0 处的 Gabor 变换域系

数 $G_{uv}(p_0), u = 0, ..., 7, v = 0, ..., fre, 可以简写为向量形式 G(p_0);$

3. 利用一些训练样本建立人脸弹性束图^[9],第k个样本 P_0 位置的 Gabor 变换域系数构成的向量表示为 $G^*(p_0), k = 1...60$,则可以计算向量 $G(p_0)$ 和向量 $G^*(p_0)$ 之间的距离估计值 $\bar{d}_*^{[9]}$ (包括尺度从 0 到 fre);

4. 由此得到候选特征点的新位置 $P_k = P_0 + \vec{d}_k$,并计算出新位置的 Gabor 变换域系数 $G(p_k)$;

5. 利用式(2-5)来计算 $G(p_{*})$ 与 $G^{*}(p_{0})$ 的相似度,其中相似度最高(设为S)的为候选者,其位置用来更新 P_{0} 。判断S是否满足给定阈值,如不满 \mathcal{L} fre = fre + 1,转到 2。

2.4.3 面部特征点的提取算法

这一部分主要介绍面部特征点提取的方法,算法中关于仿射变换的调整可以参考 ASM 的方法。

a) 确定初始位置。利用眼睛的初始位置(人脸检测提供),将模型

 $X = \overline{X} + PB$ 经 过 平 移 、 旋 转 和 缩 放 到 人 脸 上 $X_0 = M(s_0, \theta_0)[\overline{X} + PB_0] + t_0^{[12]}$, 第一步可取 $B_0 = 0$ 。

b) 边缘点的提取。对X₀中的第 i 个点(x_i,y_i),以该点为中心沿法线方向取长度为n_s(n_s>n_p,本文取 20)个象素的线段。对于该线段上每一长度为n_p(动态的)象素的子线段计算其灰度特征的一阶导数,然后输入模型中第 i 个特征点的局部模型(式(2-7)),最佳匹配子线段的中心即为搜索到的第 i 个新点位置(x_i',y_i')。对所有边缘点进行如上搜索,得到X_s(边缘点)。

- c) 控制点的提取。利用基于 Gabor 相位特征的启发式搜索方法,对 所有控制点进行搜索,得到X_c(控制点)。
- d) X'包括边缘点和控制点两部分,则由X',X₀可求出由X₀变化到X' 时参数s,θ,t的变化1+ds,dθ,dt和形状参数B的变化dB^[12],从而可 计算出X₁ = M(s₀(1+ds),θ₀+dθ)[X + P(B₀+dB)]+t₀+dt。 计 算d(X₀,X₁),如果d(X₀,X₁)<λ,则X₁为最终搜索的结果,否则令 X₀=X₁转 b)继续迭代。

2.5 实验结果

本文采用定量方法衡量面部特征定位系统的性能,则定义了如下定位 偏差:

$$E = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{n} \sum_{j=1}^{n} dist(p_{ij}, p_{ij}) \right)$$
(2-8)

其中N为测试图片数, $dist(p_{ij}, p_{ij})$ 表示面部特征点定位系统的定位 结果 p_{ij} 和手工标注的坐标 p_{ij} 之间的欧式距离, n=103, 式(2-8)是所有点定 位偏差的均值。

本文采用一个大小为 500 的人脸数据库,230 人用来训练 ASM,从中 选择 60 个人脸图像作为控制点建模数据,其余 270 人作为测试数据,实验 结果如表 2-1 所示,如图 2-2 给出一些定位结果。

关于 Gabor 函数本身的参数对于定位偏差的影响,本文也进行了实验 验证,如表 2-2 所示,Gabor 尺度用得越多,系统的性能越高,但是复杂度 也会增加。



a)

b)



图 2-2 面部特征提取结果, a,b)具有一定光照的面部特征的提取结果 c)表情变化情况 下的面部特征的提取结果 d)正常人脸的面部特征的提取结果

Fig.2-2 Samples of facial feature extraction results a,b)Samples with illumination changes, c) Samples with expression changes d) Samples in the normal condition

Table 2-1 Experimental result comparisons						
Method Average Error Improveme						
ASM	2.10					
Edge constraint	2.03	3.33%				
CSM	1.83	12.8%				

表 2-1 实验结果对比

表 2-2 不同频段对定位精度的影响 Table 2-2 Average Error for different frequcies

fre	0	0-1	0-2	0-3	0-4	Average Error
Error						
	2.01	1.97	1.85	1.68	1.64	1.83

2.6 本章小结

本文提出的方法对面部特征进行了合理的分类,分成边缘点和控制点 两部分,对不同类型的特征点采取不同的建模和搜索策略。其中,边缘点局 部建模时考虑边缘信息的影响,而控制点则利用基于 Gabor 相位估计的方法 进行定位。这种方法的优点在于:由于采用了 Gabor 小波变换,所以它对表 情和光照具有一定的鲁棒性,而只有一部分面部特征点采取 Gabor 特征进行 建模,可以降低系统的复杂度。

需要指出的是,面部特征点定位作为人脸识别的一个关键步骤之一, 还远远不能满足人脸识别任务的需要。现在人脸识别主要依赖眼睛的位置来 对人脸的面部区域进行归一化。然而我们发现很多的人脸识别算法依赖于更 多特征点的精确定位,包括一些 3D 的人脸识别算法。而本文提出的算法融 合现有的两种成熟的技术来定位面部特征点,取得了较好定位效果,但是这 方面工作还远远不足,未来的工作将集中在如何进一步提高处理的效率,以 及如何建立更加鲁棒的局部特征模型上。

第3章 基于 Gabor 相位信息的特征提取和识别方

法

基于 2D 灰度图像的人脸识别技术在很大程度上依赖于特征表示,不同的特征表示方法极大的影响了人脸识别技术的性能。一般来说,基于灰度 图像的比较有代表性的人脸识别方法是 Eigenface, Fisherface 等等。还有一 些方法基于某种变换,能够使得某些特征得到增强,比如有人提出基于小波 变换的 Waveletface^[74]方法,侧重于多尺度信息。以上的这些方法又可以称 之为全局特征表示方法。相对的,有人提出了局部特征分析方法,比如 LFA 方法,弹性图匹配方法也可称之为局部特征方法。但是以上的人脸识 别分类方法已经不能够涵盖最新方法的本质特点。比如,近年来又涌现出了 一些基于直方图特征的人脸识别方法,该方法不同于传统的基于统计学习的 方法,它不需要训练,因而可以避免泛化分析问题,而该问题在机器学习领 域还无法解决。不同于传统人脸识别的分类方法,本文按照是否存在学习过 程对人脸识别进行了简单的回顾,这里将主要关注特征提取方法。

Gabor 特征在人脸识别领域获得了成功应用,本文基于该特征提出了 一种新的人脸特征提取和识别方法。本章的重点是研究如何利用 Gabor 象限 特征表示人脸对象,并对该特征的有效性进行初步验证。

3.1 人脸识别中特征提取的回顾

实际上,人脸识别方法的分类很大程度上依赖于特征提取方法的不同。本章将首先回顾一下基于统计学习理论的判别特征提取方法。

3.1.1 基于统计学习方法的特征提取

人脸识别由于在身份鉴定、信息安全、访问控制等方面有着广泛的应 用前景,所以吸引了大量的研究人员^[1,2,75,76]。特别是最近的二十年里,取 得了很大的进步,同时产生了大量的算法,其中大多数讨论的是如何有效地 从人脸图像中抽取和描述每个个体的特征,使之区别于其他个体。第一章已 简单回顾了一些流行的特征抽取和描述的方法,基于子空间统计学习方法就 是其中最重要的一种。所以在此我们先回顾一下人脸识别中已有的各种子空 间分析方法。通常得到的图像空间的维数都是非常高的,比如一幅100 大小 的人脸图像,它的维数就是100 维。如果把人脸图像都看成是在这样高维空 间中的点,那么表示人脸图像的点的分布可能是很不紧凑的,因而不利于分 类,并且在计算上的复杂度也非常大。为了得到人脸图象的较紧凑分布,主 元分析的子空间思想[26]被引入到人脸识别中,并获得了较大的成功。随后 子空间分析方法就引起了人们的广泛注意,从而成为了当前人脸识别的主流 方法之一。子空间分析的思想是在原空间寻找一个特征空间,使得原始样本 在该空间具有好的可分性,并把它压缩到一个低维子空间,使数据在子空间 中的分布更加紧凑,为数据的更好描述提供了手段,另外计算的复杂度也得 到了大大降低。子空间分析除了有线性和非线性空间变换之分外,根据不同 的性能目标要求,得到的子空间也是不一样的。目前在人脸识别中得到成功 应用的线性子空间分析方法有: 主元分析(Principal Component Analysis, PCA)、线性判别分析(Linear Discriminant Analysis, LDA)、独立元分析 (Independent Component Analysis, ICA)、基于核技巧的非线性子空间分析 有: 核主元分析(Kernel Principal Component Analysis, KPCA)和核Fisher判 别分析(Kernel Fisher Discriminant Analysis, KFDA)。下面本文将按照线性 和非线性,以及时间上的顺序对这些方法在特征提取方面的应用作一下简单 的回顾、比较和总结。针对于判别特征提取的判别分析方法是本文工作的一 部分,所以本章将不做详细阐述,将在第5章进行详尽的分析。

主元分析的思想来源于K-L 变换,它是图像压缩中一种最优正交变换,人们将它用于统计特征提取,在数学上,主元分析就是通过解特征值问题来对角化协方差矩阵^[26],最后选择对应前面的非零特征值对应的特征向量作为新的子空间的基向量,即主元。因而原空间的样本就可以用这些主元的线性组合来表示,其中系数就是样本在低维子空间上的投影。主元分析思想在人脸识别上的应用最早是由Kirby等^[26]提出来的,后来由Turk 等^[65]将其发展为特征脸(Eigenface)方法,用于正面的人脸识别。特征脸的思想就是从训练图像中,通过主元分析的方法学习得到一组特征脸图像(即对应的主元),那么任意给定的人脸图像都可以近似为这组特征脸图像的线性组合,组合的系数就表示为样本的特征向量。

经典的特征脸方法是采用基于Euclidean 距离的最近中心分类器,后来 比较常用的是基于Euclidean 距离的最近邻分类器,也有采用基于 Mahalanobis 距离和余旋角度量的^[77]。随着特征脸方法的成功,还出现了许 多种基于主元分析的人脸识别方法。Moghaddam 等^[29,30]后来还提出基于分

31

层的主元概率子空间方法。该方法的思想是分别用主元分析去提取与描述人 脸的类内变化和类间变化的特征,然后再结合文献[65]中的思想,用贝叶斯 决策规则来分类。实验的结果表明,结合贝叶斯分类框架和主元分析的人脸 识别方法比经典特征脸方法取得的识别效果要好,这也说明了基于贝叶斯决 策的分类误差较小。尽管主元分析在人脸识别中取得了不错的效果,但是由 于它的标准是基于所有样本的最优重建与恢复,因此对于描述不同类样本之 间的差异而言,它不能保证是最优的描述方式,从而用它来描述人脸识别的 特征可能是不充分的。

线性判别分析不同于主元分析,它是以使样本的可分性最好为目标, 因此从理论上来说,比较适用于模式识别问题。经典的线性判别分析中使用 的是Fisher 准则函数,所以线性判别分析又被称为Fisher线性判别分析 (Fisher Linear Discriminant Analysis, FLDA)。Fisher 准则对人脸识别的应用 来说,由于通常没有足够的训练样本来保证类内散度矩阵为满秩,无法直接 求解,因此需要加入一定的条件或策略来求解。最早Swets 等^[78]提出结合 主元分析的线性判别方法,即先用主元分析降维,获得原样本的最优特征表 示子空间,保证在该空间中类内散度是非奇异的,然后再在此基础上作线性 判别分析。后来Belhumeur 等把它发展为Fisher脸(Fisherface)方法,并在光 照变化、表情变化以及有无戴眼镜的情况下与特征脸方法作了实验比较,分 析的结论是在测试图像与训练图像变化不大的情况下两种方法都能得到很好 的识别效果,但是在有光照、表情等变化的情况下,Fisher 脸要比特征脸 鲁棒,另外他们还发现在特征脸方法中,对应前三个最大特征值的主元对光 照的变化比较敏感,所以建议在有光照变化时,去掉这三个主元。Zhao等^[32] 也对此方法作了深入的研究,他们称之为主元判别分析(Discriminant Analysis of Principal Component),并实现了一个实际系统。此系统在人脸识 别领域具有比较重要的影响。

同理,核Fisher 判别分析是结合核技术和线性判别分析的思想,来提取判别特征。它既利用了核技巧能有效描述数据复杂的非线性关系的特性, 又继承了Fisher 线性判别分析的优点,使类与类之间尽可能的具有可分性。因此,从理论上讲它比Fisher 线性判别分析在分析实际的模式识别问题时要有优势,在实际中也有了成功的应用。

在前人工作的基础上,本文对核Fisher方法进行改进并用来抽取和描述 人脸识别的判别特征,尤其关注基于直方图特征的核方法的研究,详细分析 参见第5章。

3.1.2 非统计学习方法判别特征提取方法

统计学习方法的优势在于它能够利用训练集合中的一些先验信息,往 往是一些由于数据的某些属性的变化引起的差异性信息。在人脸识别中,主 要体现在姿态、光照、表情、年龄等的变化,以及数据分布的变化上。而统 计学习的劣势也来源于此,也就是当它遇到与其分布不一样的数据时,系统 的性能会受到很大的影响。所以很多研究者提出一些直接提取特征的人脸识 别方法,这些方法往往不依赖于学习来获得判别特征,所以避免了泛化分析 问题,本节首先介绍弹性图匹配方法。

3.1.2.1 弹性束图匹配

与其他方法比较,弹性束图匹配方法对变化的适应性更强;向人脸库 中加入新的人脸图像时,不需要重新训练及改变已有的数据,直接加入新的 模板数据即可。但弹性束图匹配方法的计算复杂度较高;另外要达到可靠的 识别水平,最初的训练过程还需要手动生成足够多的标记图或束图。

3.1.2.2 基于空间直方图的特征提取方法

空间直方图特征^[41,42,79]在人脸识别领域的成功应用,使得人脸识别的 研究人员重新思考现有方法的前途,因为占统治地位的子空间方法的泛化分 析问题在理论上还无法解决。尤其值得一提的是,局部二值模式直方图方法 假定了人脸图像是由一系列微模式组成的,所以让我们更多的思考人脸图像 的本质特性。这些微模式包括边缘、奇异点、平坦区域等等,其位置信息能 够保留有利于识别的结构信息,这在某种程度上让模式识别领域的研究者重 新审视结构模式识别方法的未来。另外,直方图特征在图像处理领域被广泛 的研究,可以参考文献[80-84]。

3.1.3 小结

从本章回顾和分析的各种人脸识别方法可以知道,子空间分析方法由 于具有计算代价小、可分性好等特点,已成为人脸识别中的主流特征提取方 法之一。基于可分性考虑的判别分析方法在人脸识别中的应用具有一定的优 越性。但是最大的问题是它还无法解决泛化能力较弱的问题。而弹性图匹配 方法,基于LBP直方图特征的人脸识别方法无需训练。这样就存在一个非常 有意义的,同时也是非常有挑战性的问题,即如何结合这两种方法。另外, 从最近的研究发展来看,非线性分析越来越受关注。本文的研究重点之一就 是如何结合这两种人脸识别方法,从而提高人脸识别系统的性能。

基于以上对人脸识别的分类方法,本文对一些著名的人脸识别方法进行了归类。如本征脸(Eigenface),Fisherface,贝叶斯方法,全局和局部特征融合的子空间方法^[85],非线性判别分析方法等。这些方法都是基于统计学习理论提出的,即其核心的思想来源于主成分分析和线性判别分析等方法。它们都存在泛化能力较弱的问题,当测试集合和训练集合样本分布不一样时,系统的识别性能会受到很大影响,往往不能得到好的识别效果。而GFC方法^[10]是Fisherface的增强版本,同样存在这样的问题。另外,非线性判别分析中核判别分析方法在人脸识别领域取得成功的应用,已成为模式识别领域一个迅速发展的研究方向,越来越多的研究者开始关注这一领域,其中最有影响力的是,支持向量机^[14]从理论上通过VC-维和结构风险最小化原则讨论了泛化问题。但是在人脸识别应用上,泛化问题已经成为一个瓶颈。

不同于传统的统计学习方法,现在基于空间直方图特征的人脸识别方 法作为最新的人脸识别方法,已经逐渐地得到该领域的研究人员的认可。并 且该方法正逐渐与传统的人脸识别技术进行结合,为人脸识别技术走向实际 应用奠定了良好的基础。

3.2 人脸识别中 Gabor 特征表示的简述

由于Gabor函数能够同时在时域和频域中兼顾对信号分析的分辨率要 求,因此它在信号处理中得到了广泛的应用。而二维Gabor函数能较好地描 述哺乳动物初级视觉系统中一对简单视觉神经元的感受特性,所以早在80年 代^[73],Daugman将其用在计算视觉领域。弹性图匹配方法就是基于Gabor变 换的局部建模方法,它首先在人的面部区域定位一些关键点,然后对每一个 点进行Gabor变换作为特征,称之为*jet*。丁嵘等^[25]采用关键信息点来改进传 统的弹性图匹配方法取得了较好的效果。刘等^[10]提出了无需定位关键特征 点(本文称之为Gaborface)的特征表示方法,该方法的核心思路是提取整幅图 像的Gabor幅值特征构建一个扩展的向量(参见图3-1),然后进行线性判别分 析提取判别特征的方法。Gaborfaces^[10]方法的好处在于不用对原始图像进行 特征点选取,这样使得它更容易得到应用,且保留了丰富的局部和全局信 息,但是由于维数太高,需要对它进行降维。很多方法都可以解决这个问 题,比如下采样^[10],Boosting方法^[11]等等。

山等^[11]对Gaborface特征进行选择,取得非常好的识别性能,而且效率

上有进一步的提高。我们发现现有的人脸识别系统大多数都是基于幅值信息的^[1,2,10],事实上,Gabor相位信息也具有很好的识别能力,已经被成功的应用在虹膜和掌纹识别上^[87,88]。在这两个应用中,主要提取Gabor象限特征,并采取海明距离进行相似度度量。



图 3-1 Gabor 幅值特征可视化结果

Fig.3-1 Visulization of Gaborfaces(Magnitude parts)

本文提出一种完全基于Gabor相位信息的人脸识别方法,首先提出了 Gabor相位编码模式的方法,然后把它分成互不相交的多个子块从而提取空 间直方图作为最终的特征。识别的时候以直方图的交作为相似度计算公式, 利用最近邻分类器来进行识别。该方法无需训练,但是在基本FERET数据 库上取得了比LGBPHS更好的实验结果。另外,由于不同的区域对识别的影 响是不一样的,本文基于Fisher原则对每个子块进行加权,来进一步提高系 统的性能。

本章的后续内容组织如下,首先提出一种新的对象表示方法,即:基 于空间直方图的 Gabor 象限编码模式(HGPP)。然后提出基于 HGPP 的匹配 方法。最后通过实验验证本章所提出特征提取方法的有效性。

3.3 基于空间直方图的局部 Gabor 相位编码模式

这一节将首先提出一种基于Gabor相位信息的编码方法,它能够提取由 于每一点和其邻域点的象限差异性所引起的变化,然后结合空间直方图来表 示原始输入图像。由于该方法基于Gabor象限特征,所以下面将首先描述一 下如何提取该特征。

3.3.1 Gabor 相位角的象限量化

由上一节知道,对于一幅灰度图像 I(z)的 Gabor 变换结果,随着空间 位置的变化,它的幅值变化不是很明显,如图 3-1 所示。而相位部分即使 是几个像素的距离也会发生很大的变化,尽管它们代表着同一个局部特征, 如图 3-2 所示。由于这种"不稳定性"导致了前人的工作大多数是基于幅值 特征的。

然而,相位信息并不是无用的,它甚至包含更多的局部细节信息,一个典型的成功应用就是 Daugman 提出的基于象限特征的虹膜识别方法^[87]。 在[87]中,每一个点都被量化成两个比特位,本文称之为实部和虚部象限量 化方法((*P*^{*e}(*Z*),*P*ⁱⁿ(*Z*))),即:

$$P_{u,v}^{\text{Re}}(Z) = \begin{cases} 0, \text{ if } \operatorname{Re}(G_{u,v}(Z)) > 0\\ 1, \text{ if } \operatorname{Re}(G_{u,v}(Z) <= 0 \end{cases}$$
(3-1)

$$P_{u,v}^{\rm Im}(Z) = \begin{cases} 0, \text{ if } \operatorname{Im}(G_{u,v}(Z)) > 0\\ 1, \text{ if } \operatorname{Im}(G_{u,v}(Z)) <= 0 \end{cases}$$
(3-2)

这里 $Re(G_{u,v}(z))$ 和 $Im(G_{u,v}(z))$ 是Gabor特征的实部和虚部部分, $G_{u,v}(z)$ 的定义可参考上一章。

参考图 3-3,我们可以很容易把式(3-1)和式(3-2)重新形式化:

$$P_{u,v}^{\text{Re}}(Z) = \begin{cases} 0, \text{ if } \theta_{u,v}(Z) \in \{\text{I}, \text{IV}\} \\ 1, \text{ if } \theta_{u,v}(Z) \in \{\text{II}, \text{III}\} \end{cases}$$
(3-3)

$$P_{u,v}^{\text{Im}}(Z) = \begin{cases} 0, \text{ if } \theta_{u,v}(Z) \in \{\text{I}, \text{II}\} \\ 1, \text{ if } \theta_{u,v}(Z) \in \{\text{III}, \text{IV}\} \end{cases}$$
(3-4)

P^{Re}_{*u,v*}, *P*^{Im}_{*u,v*} 即所谓的象限量化(QBC),因为它们确定一个相位角所在象限。在虹膜和掌纹识别中,只是简单串接 Gabor 象限特征二值化结果,而没有考虑局部特性以及 Gabor 参数不同所带来的变化。由于局部特征在人脸识别领域获得非常成功的应用^[8,9,42],本文在 QBC 的基础之上提出抽取局部相位模式的方法。



图 3-3 Daugman的相位角的象限量化示意图 Fig.3-3 Illustration of the Quandrant Bit Coding method

3.3.2 局部 Gabor 相位编码模式

这一节将提出基于局部异或运算算子的 Gabor 相位模式的编码方法 (LGPP)。对于每一个 *u* 和 *v*, 实部的局部 Gabor 相位编码模式(Real-LGPP), 虚部的局部 Gabor 相位编码模式(Imaginary-LGPP), 分别定义如下:

$$LGPP_{u,v}^{\text{Re}}(Z_{0}) = [P_{u,v}^{\text{Re}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Re}}(Z_{1}), \qquad (3-5)$$

$$P_{u,v}^{\text{Re}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Re}}(Z_{2}), \dots, P_{u,v}^{\text{Re}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Re}}(Z_{8})]$$

$$LGPP_{u,v}^{\text{Im}}(Z_{0}) = [P_{u,v}^{\text{Im}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Im}}(Z_{1}), \qquad (3-6)$$

$$P_{u,v}^{\text{Im}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Im}}(Z_{2}), \dots, P_{u,v}^{\text{Im}}(Z_{0})_{\text{XOR}}P_{u,v}^{\text{Im}}(Z_{8})]$$

这里 Z_i , i = 1,2,...,8, 是 Z_0 的 8-邻域点, XOR 表示异或运算,本文称之为 局部异或运算算子(Local XOR Pattern, LXP),如图 3-4 所示。



图 3-4 局部异或运算算子示意图

Fig.3-4 Illustration of Local XOR Pattern Operator

8-邻域"碰巧"提供 8 个比特从而每一点得到一个字节,因而可以用一个在[0, 255]之间的整数来表示一个模式,该模式表示了该点与其邻域点之间的相位变化。简单的说,对于每一个象限量化方法我们得到 *u*v* 个 LGPP "图像",可以参考图 3-5 (a)和(b),这些图看起来像具有丰富纹理信息的指纹图。

这两种局部 Gabor 相位编码模式还等价于:

$$P_{u,v}^{\text{Re}}(Z_0)_{\text{XOR}}P_{u,v}^{\text{Re}}(Z_i) = \begin{cases} 0, \text{ if } \text{Re}(G_{u,v}(Z_0)) * \text{Re}(G_{u,v}(Z_i)) > 0\\ 1, \text{ if } \text{Re}(G_{u,v}(Z_0)) * \text{Re}(G_{u,v}(Z_i)) <= 0 \end{cases}$$
(3-7)

$$P_{u,v}^{\mathrm{im}}(Z_{0})_{\mathrm{XOR}}P_{u,v}^{\mathrm{im}}(Z_{1}) = \begin{cases} 0, \text{ if } \mathrm{Im}(\mathrm{G}_{u,v}(Z_{0})) * \mathrm{Im}(\mathrm{G}_{u,v}(Z_{i})) > 0\\ 1, \text{ if } \mathrm{Im}(\mathrm{G}_{u,v}(Z_{0})) * \mathrm{Im}(\mathrm{G}_{u,v}(Z_{i})) <= 0 \end{cases}$$
(3-8)

本文还给出了一些例子来验证该方法的有效性,从图 3-6 可以看出,不同人的 模式图是不一样的,而同一个人具有相似的模式图。另外,在对其进行特征提 取时,采用分块的方法,如图 3-7 所示。

事实上,和局部二值模式一样,局部异或模式也能反映图像的边缘、 点、平坦区域等特征,这是因为针对二值图像,异或运算反映了当前面部特 征点和周围点之间的异同。例如,当一个模式为'00000000'的时候,则表示 这个模式表示平坦区域,而当为'1111111'的时候,则表示为一个点,相应 的也可以表示成边缘等等,具体可以参考局部二值模式。



(a)



b)

图 3-5 a) 实部的局部Gabor相位编码模式(Real-LGPP) b)虚部的局部Gabor相位编码模式(Real-LGPP)

Fig.3-5 a) Real Part Local Gabor Phase Pattern b) Imaginery Part Local Gabor Phase

Pattern



图 3-6 不同的人脸图像提取的局部 Gabor 相位编码模式 Fig.3-6 LGPPs for Different faces

3.3.3 全局 Gabor 相位编码模式

全局 Gabor 相位模式实际上是提取 Gabor 参数中方向的变化所带来的 差异性,该模式被形式化表示如下:

$$GGPP_{v}^{Re}(Z_{0}) = [P_{0,v}^{Re}(Z_{0}), P_{1,v}^{Re}(Z_{0}), ..., P_{k,v}^{Re}(Z_{0})]$$
(3-9)

$$GGPP_{v}^{Im}(Z_{0}) = [P_{0,v}^{Im}(Z_{0}), P_{1,v}^{Im}(Z_{0}), ..., P_{k,v}^{Im}(Z_{0})]$$
(3-10)

很显然,8 个方向正好构成了一个字节,所以该微模式可以变成一个[0,255] 之间的整数:

$$GGPP_{\nu}^{\text{Re}}(Z_0) = P_{0,\nu}^{\text{Re}}(Z_0) * 2^k + P_{1,\nu}^{\text{Re}}(Z_0) * 2^{k-1} + \dots + P_{k,\nu}^{\text{Re}}(Z_0)) \quad (3-11)$$

$$GGPP_{\nu}^{\mathrm{Im}}(Z_0) = P_{0,\nu}^{\mathrm{Im}}(Z_0) * 2^k + P_{1,\nu}^{\mathrm{Im}}(Z_0) * 2^{k-1} + \dots + P_{k,\nu}^{\mathrm{Im}}(Z_0)) \quad (3-12)$$

3		1					2
			3	邋		$\overline{\mathbf{n}}$	1.
							儀
10							
	瀫				2		虚
Т,	2		蜀	1		濫	
j.	10			5.5	1	3	

图 3-7 模式分块方法示意图

Fig.3-7 Illustration of division method for Gabor Phase Pattern



(a)



(b)

图 3-8 a) 实部的全局Gabor相位编码模式(Real-GGPP) b)虚部的全局Gabor相位编码模式(Real-GGPP)

Fig.3-8 a) Real Part Global Gabor Phase Pattern b) Imaginery Part Global Gabor Phase Pattern

3.3.4 基于空间直方图的 Gabor 相位模式

在 Daugman 的方法中,象限量化的结果直接用来表示虹膜图像,识别的时候主要利用了海明距离进行度量。然而,在本文中,对于每一种象限量化方法,都能够得到 40 个和原始图像同样大小的图像。这些图像具有很好的纹理特征,而直方图被广泛用来提取纹理特征,所以本文进一步利用空间

直方图来提取判别特征。由于单一的直方图会失去很多结构信息,以人脸图像为例,它本身是一个拓扑图像,这就意味着空间结构信息对人脸识别是尤其重要的。本文首先把 GPP 图像分成互不相交的几个子区域,即 R₁,...,R_L, 然后针对每一块抽取局部的直方图特征(如图 3-7 所示),这样所有的直方图特征连接成为一个扩展的向量 (HGPP),即:

$$H_{GGPP}^{\text{Re}} = (H_{GGPP}^{\text{Re}}(v,l) : v = 0, \dots, 4; l = 1, \dots, L)$$

$$H_{GGPP}^{\text{Im}} = (H_{GGPP}^{\text{Im}}(v,l) : v = 0, \dots, 4; l = 1, \dots, L)$$

$$H_{LGPP}^{\text{Re}} = (H_{LGPP}^{\text{Re}}(u,v,l) : u = 0, \dots, 7; v = 0, \dots, 4; l = 1, \dots, L) \quad (3-13)$$

$$H_{LGPP}^{\text{Im}} = (H_{LGPP}^{\text{Im}}(u,v,l) : u = 0, \dots, 7; v = 0, \dots, 4; l = 1, \dots, L)$$

这里 L 是子区域的个数。



Fig.3-9 Flow chart of HGPP

3.3.5 进一步分析

至此,本文提出了基于空间直方图的局部 Gabor 相位编码模式的方法,如图 3-9 所示。该方法具有如下特征:

(1) 基于 Gabor 小波,继承了 Gabor 变换的特性,比如多尺度,多方向,以及其局部特性等等,因而保留大量判别信息。

(2)由于无需训练,因而不存在泛化能力较弱的问题,这在统计学习 里还无法克服,尤其是当训练集合和测试集合的分布有很大差别的时候,会 导致系统性能急剧下降。

(3) 计算简单,我们可以发现局部 Gabor 相位编码模式是基于移位和 异或运算得到的,而直方图特征的提取只需要加运算即可。

(4) 该方法本质上是一种局部建模方法,这是因为 Gabor 变换和空间

直方图都具有局部特征,因而该方法对表情,饰物,光照等的变化是鲁棒的。

3.4 基于 HGPP 的人脸识别

当我们采用基于直方图特征的对象表示方法时,不能够有效的利用传 统欧式距离等作为相似度计算公式。但是现在有很多现成的基于直方图特征 的相似度计算方法,比如直方图的交,所以本文主要利用它作为最近邻分类 器的相似度计算公式。

3.4.1 基于直方图的交的匹配方法

假定 H 和 S 为直方图,基于直方图的交相似度计算公式如下所示:

$$S_{HI}(H,S) = \sum_{i=1}^{B} \min(H_i, S_i)$$
(3-14)

这里 *B* 是直方图的 bins, *H_i* 表示频度,直方图的交实际上累加了两个直方 图的公共部分。基于此,两个人脸图像的直方图特征 *H*1,*H*2 之间的相似度 如下所示:

$$S(H1,H2) = S(H1_{GGPP}^{\text{Re}}, H2_{GGPP}^{\text{Re}}) + S(H1_{GGPP}^{\text{Im}}, H2_{GGPP}^{\text{Im}}) + S(H1_{LGPP}^{\text{Re}}, H2_{LGPP}^{\text{Re}}) + S(H1_{LGPP}^{\text{Im}}, H2_{LGPP}^{\text{Im}})$$
(3-15)

每一项分别计算如下所示:

$$S(H1_{GGPP}^{\text{Re}}, H2_{GGPP}^{\text{Re}}) = \sum_{\nu=0}^{4} \sum_{l=1}^{L} S_{HI}(H1_{GGPP}^{\text{Re}}(\nu, l), H2_{GGPP}^{\text{Re}}(\nu, l))$$
(3-16)

$$S(H1_{GGPP}^{Im}, H2_{GGPP}^{Im}) = \sum_{\nu=0}^{4} \sum_{l=1}^{L} S_{HI}(H1_{GGPP}^{Im}(\nu, l), H2_{GGPP}^{Im}(\nu, l))$$
(3-17)

$$S(H1_{LGPP}^{\text{Re}}, H2_{LGPP}^{\text{Re}}) = \sum_{u=0}^{7} \sum_{v=0}^{4} \sum_{l=1}^{L} S_{HI}(H1_{LGPP}^{\text{Re}}(u, v, l), H2_{LGPP}^{\text{Re}}(u, v, l))$$
(3-18)

$$S(H1_{LGPP}^{Im}, H2_{LGPP}^{Im}) = \sum_{u=0}^{7} \sum_{v=0}^{4} \sum_{l=1}^{L} S_{HI}(H1_{LGPP}^{Im}(u, v, l), H2_{LGPP}^{Im}(u, v, l))$$
(3-19)

L 表示子区域的个数。简单的说,式(3-16),(3-17),(3-18),(3-19)中的各 个子块的相似度的和,即相当于多分类器的融合,该相似度计算公式能够保 留空间结构信息,所以更适合人脸识别。

3.4.2 基于 Fisher 原则的加权方法

不同人脸区域对识别的重要性是不一样的,因而从不同区域提取的直 方图特征具有不同的判别能力,直观的来说,就是对于不同的直方图块赋予 不同的权重,则式(3-16)到 式(3-19)重新改写成如下形式:

$$S(H1_{GGPP}^{\text{Re}}, H2_{GGPP}^{\text{Re}}) = \sum_{\nu=0}^{4} \sum_{l=1}^{L} \omega_{GGPP}^{\text{Re}}(\nu, l) S_{HI}(H1_{GGPP}^{\text{Re}}(\nu, l), H2_{GGPP}^{\text{Re}}(\nu, l))$$
(3-20)

$$S(H1_{GGPP}^{Im}, H2_{GGPP}^{Im}) = \sum_{\nu=0}^{4} \sum_{l=1}^{L} \omega_{GGPP}^{Im}(\nu, l) S_{HI}(H1_{GGPP}^{Im}(\nu, l), H2_{GGPP}^{Im}(\nu, l))$$
(3-21)

$$S(H1_{LGPP}^{\text{Re}}, H2_{LGPP}^{\text{Re}}) = \sum_{u=0}^{7} \sum_{v=0}^{4} \sum_{l=1}^{L} \omega_{LGPP}^{\text{Re}}(u, v, l) S_{HI}(H1_{LGPP}^{\text{Re}}(u, v, l), H2_{LGPP}^{\text{Re}}(u, v, l))$$
(3-22)

$$S(H1_{LGPP}^{Im}, H2_{LGPP}^{Im}) = \sum_{u=0}^{7} \sum_{v=0}^{4} \sum_{l=1}^{L} \omega_{LGPP}^{Im}(u, v, l) S_{HI}(H1_{LGPP}^{Im}(u, v, l), H2_{LGPP}^{Im}(u, v, l))$$
(3-23)

这里 @ 表示不同的局部直方图具有不同的权重。

本文主要利用 Fisher 原则来计算权重*ω*,给定一个集合包含了*C*个类别,两个对应的子块直方图分别来自或者属于同一类,或者是不同类别,分别构成了类内,类间子空间。对于第*p*个直方图 *H_p*,类内均值和方差可以用如下公式来计算^[89]:

$$m_{I}^{p} = \frac{1}{C} \sum_{i=1}^{C} \frac{2}{N_{i}(N_{i}-1)} \sum_{k=2}^{N_{i}} \sum_{j=1}^{k-1} S_{HI} \left(H_{ij}^{p}, H_{ik}^{p} \right)$$
(3-24)

$$\sigma_{I}^{P} = sqrt\left(\sum_{i=1}^{C}\sum_{k=2}^{N_{i}}\sum_{j=1}^{k-1} \left(S_{HI}\left(H_{i,j}^{P}, H_{i,k}^{P}\right) - m_{I}^{P}\right)^{2}\right)$$
(3-25)

H^p_{ij}表示 H_p 直方图块是来自第 *i*-类的第 *j*-个样本, N_i表示第 *i*-类的样本个数, 类似的, 类间的均值和方差按如下公式计算:

$$m_E^p = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{j=i+1}^{C} \frac{1}{N_i N_j} \sum_{k=1}^{N_j} \sum_{l=1}^{N_j} S_{Hl} \left(H_{ik}^p, H_{j,l}^p \right)$$
(3-26)

$$\sigma_{E}^{p} = sqrt\left(\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\sum_{k=1}^{N_{i}}\sum_{l=1}^{N_{i}}\left(S_{HI}\left(H_{i,k}^{p}, H_{j,l}^{p}\right) - m_{E}^{p}\right)^{2}\right)$$
(3-27)

最后, H_p 的权重ω_p可以按如下公式来求:

$$\omega_{p} = \frac{\left(m_{l}^{p} - m_{E}^{p}\right)^{2}}{\left(\sigma_{l}^{p}\right)^{2} + \left(\sigma_{E}^{p}\right)^{2}}$$
(3-28)

以上权重设置的过程实际上利用每一个直方图子块的判别能力,也就 是大的权重赋予给了具有强判别能力的区域直方图。尽管权重的计算过程似 乎违背了本文提出方法的非统计学习特征,然而通过实验我们发现没有权重 的识别方法已经让人很满意了。

3.5 实验分析

这一节,基本 FERET 数据库 和 CASPEAL-R1 数据库被用来验证本 章提出方法的性能。它们是目前规模较大的两个数据库,能够有效的验证各 种方法的性能。

3.5.1 基本 FERET 测试

基本 FERET 数据库被广泛地用来测试人脸识别算法的性能,在本文的实验中,首先根据眼睛的位置把图像归一化成 128*128 大小的图像,此时 Gabor 核的参数 $f_{\text{max}} = \pi/2$ 。



图 3-10 FERET 数据库的样本 Fig. 3-10 Samples in the FERET database

根据 FERET 测试协议,本文采取同样的 gallery 和 probe 数据库,以及 同样的训练集合。其中 Fa (1196)作为 gallery 数据库, Fb (1195), Fc (194), Dup I (722) 和 Dup II (234) 作为 probe 集合。

3.5.1.1 不同图像大小对识别率的影响

如图 3-11 所示,不同的图片大小和 f_{max} 是相互关联的,这是因为 f_{max} 越大,核的有效积分区域就越小,则大的 f_{max} 对应于一个小尺寸的人脸图像,而小的 f_{max} 对应于大的人脸图像,才能够提取合适的空间结构信息。



c)

图 3-11 Gabor 核的幅值部分 f_{max} . a) $f_{\text{max}} = \pi/2$; b) $f_{\text{max}} = \sqrt{2\pi}/2$, and c)

 $f_{\rm max} = \pi$

Fig.3-11 Magnitude parts of Gabor kernels. a) $f_{\text{max}} = \pi / 2$; b) $f_{\text{max}} = \sqrt{2\pi} / 2$, and c)

 $f_{\rm max} = \pi$

不同图片大小实际上提供的信息量是不一样的,当使用大的图片时,通 常会得到更多有利于识别的信息,但是系统的复杂度也会提高。

由表 3-1 可以很容易知道,大的人脸图像和合适的 f_{max} 能取得比小的图像更好的识别效果。下面测试中,都采用 128x128 大小的归一化图像。参考 LGBPHS 的经验,直方图的 bins 选为 128,以及 64 个子块。

	Fb (%)	Fc (%)	Dup I(%)	Dup II(%)			
64x64, $f_{\text{max}} = \pi$	95.1	97.4	74.9	72.2			
88x88, $f_{\rm max} = \sqrt{2}\pi / 2$	97.2	98.5	77.1	75.6			
128x128, $f_{\rm max} = \pi / 2$	97.6	98.9	77.7	76.1			

表 3-1 不同图像大小对识别率的影响

Table 3-1 Recogntion rates for different image sizes for basic HGPP

3.5.1.2 对旋转和平移的鲁棒性测试

直方图特征忽略位置信息,所以对于旋转和平移是鲁棒的,对于空间 直方图,它同样具有这样的鲁棒性。山等^[90,91]人验证了基于像素点的人脸识 别方法对于误配准问题是不鲁棒的。如下的实验能够验证基于分块的人脸识 别算法更加适合于实际的应用。下面首先描述一下实验中所用到的人脸图像 处理情况,图 3-12 为正面图像,而图 3-13 和 3-14 则为对原始图像的旋转 和平移之后的结果。



图 3-12 正面图像 Fig.3-12 Frontal face





图 3-13 第一行人脸图像向左旋转 2,6,10 度,第二行表示人脸图像向右旋转 2,6,10 度

Fig.3-13 The first row rotates with 2,6,10 degrees, the second row with 2,6,10 degrees



图 3-14 前两行表示人脸图像在 Y 左平移 2,3,4,第三行和第四行表示人脸图像在 X 轴上进行平移 2,3,4 像素

Fig.3-14 The top two rows with X-translation 2,3,4 pixels, the tow bottom rows with Y-translation 2,3,4 pixels



图 3-15 在 FB 测试集合上验证 HGPP 对旋转的鲁棒性 (2 表示向左旋转 2 度, -2 表示向右旋转 2 度)

Fig.3-15 Robustness of HGPP against rotation on FB probe(2 means 2 degress Left rotation, -2 means 2 degress Right rotation)



图 3-16 在 FB 测试集合上验证 HGPP 对 X 轴上的平移的鲁棒性(例如 3 表示向左平移 3 个像素, -3 表示向右平移 3 个像素)

Fig.3-16 Robustness of HGPP against X-translation on FB probe(3 means 2 pixels Left translation, -3 means 3 pixels Right translation in the X-direction)



图 3-17 在 FB 测试集合上验证 HGPP 对 Y 轴上的平移的鲁棒性(例如 3 表示向左平移 3 个像素, -3 表示向右平移 3 个像素)

Fig.3-17 Robustness of HGPP against Y-translation on the FB probe(3 means 3 pixels Left translation, -3 means 3 pixels Right translation in the Y-direction)

从上面的测试可以知道,本文提出的方法对平移和旋转的鲁棒性是可以接受的。特别地,对比[90,91],可以看到本文的方法是非常鲁棒的,其原因在于 HGPP 是基于区域直方图特征的人脸识别方法,而 Fisherface 和 GFC 方法都是基于像素匹配的。

直方图特征不考虑位置信息,所以当小的平移,或者旋转存在时,大 部分像素点都在本来的子块中,所以即使是基于空间直方图的方法仍然对误 匹配具有鲁棒性。

3.5.1.3 降维分析

本文提出的方法有一个很大的问题就是特征维数很高,通过分析发现特征维数取决于子块的数目和直方图的 bins 数,所以调整这两个参数能够解决维数过高的问题。首先,由于不同的子块数决定了保留空间信息的多少,所以我们首先通过实验来衡量子块数目的变化对识别率的影响。



图 3-18 子块的个数对识别率的影响

Fig.3-18 Recogntion rates for different number of sub-regions

从图 3-18 我们可以看出,随着子块数的增加,复杂度变大,但识别率 会提高,当子块数为 4*4,8*8,和 16*8,即每一个子块的大小分别为 32*32,16*16,和 8*16时,128 和 64 块的时候要比 16 块的时候识别率有明 显的提高。实际上,这也说明了人脸图像的空间结构信息对识别是非常有利 的,但是随着子块数的增加,系统的复杂度也会增加。如图 3-19 所示的实 验中采用了 64 个子块。

除此以外,直方图的 bins 数也会影响到系统的识别性能以及复杂度。 从图 3-19 可以看出直方图的 bins 数越大,识别率就越高,但是相应提取的 特征长度也会增加。我们可以看出系统的复杂度问题已经得到了很好地解 决,因为提取特征的长度在大幅度下降以后,系统的识别性能下降的幅度 并不是很明显。

通过以上两个实验的观察可以知道, 直方图 bins 和子块数的选择范围

100 95 95 90 90 85 80 Fb Fb Fc 100 90 128 64 32 16

是很宽的。事实上,LGBPHS和LBP方法已经给我们提供了一些经验。

图 3-19 不同的直方图 Bins 数对识别率的影响

Fig.3-19 Recogntion rates for different number of histogram bins

3.5.1.4 各种模型性能对比

从 3.3 节可以看到本文提出的方法是由多个模型组成的,包括基于实部和 虚部的全局模型、局部模型。由于基于不同的特征,以及不同的模式提取方 式,可以知道不同模型能够提供不同的判别信息,下面的实验用来验证各种模 型的识别率,以及它们之间的互补性。

表 3-2 各种局部 Gabor 相位编码模式的首选识别率对比

Table 3-2 Rank-1 recognition rate comparasitions for different kinds of Gabor Phase

Patterns							
	Fb(%) Fc(%) Dup I(%) Dup II(%)						
Re_GGPP	91.9	87.6	64.3	59.4			
Im_GGPP	91.2	86.1	64.4	61.1			
Re_LGPP	97.0	97.9	74.9	72.6			
Im_LGPP	96.8	97.9	74.7	74.4			
Re_Im_GGPP	92.8	89.2	67.2	62			
Re_Im_LGPP	97.1	98.5	77	75.6			
Re_GGPP_LGPP	97.4	98.5	75.4	74.8			
Im_GGPP_LGPP	97.0	98.5	76.1	75.2			
GPP (all)	97.6	98.9	77.7	76.1			

从表 3-2 可以看出, Re_LGPP 和 Im_LGPP 取得很相似的性能, 但是 却远远好于 Re_GGPP 和 Im_GGPP。需要说明的是, GGPP 的复杂度要远 远小于 LGPP, 它的存储空间是 LGPP 的 1/8。另外, 实部和虚部, 以及 GGPP 和 LGPP 的结合都提高了系统的性能。而 GPP 将所有模型结合在一 起取得了最高的性能,所以进一步说明以上这些模型是互补的。

3.5.1.5 和各种著名的人脸识别系统的对比结果

FERET 作为一个著名的人脸数据库,很多算法在该数据库上的测试结果都能够在公开发表的文献中得到,所以可以进行充分对比试验。在这一节,主要可以获得如下方法的实验结果并进行对比,包括 LGBPHS, LBP, LDA, EBMG, GFC 等。其中大部分方法都是基于 Gabor 幅值特征的,其对比结果如表 3-3 所示。

Methods		FERET probe sets				
	Data Source	Fb(%)	Fc(%)	Dup I(%)	Dup II(%)	
HGPP_Weighted	This paper	97.5	99.5	79.5	77.8	
HGPP	This paper	97.6	98.9	77.7	76.1	
LGBPHS_Weighted	ICCV2005 [56]	98.0	97.0	74.0	71.0	
LGBPHS	ICCV2005 [56]	94.0	97.0	68.0	53.0	
LBP_Weighted*	ECCV2004 [57]	97	79.0	66	64.0	
LBP	ECCV2004 [57]	93.0	51.0	61.0	50.0	
UMD LDA	TPAMI2000 [22]	96.2	58.8	47.2	20.9	
USC EBGM	TPAMI2000 [22]	95.0	82.0	59.1	52.1	
GFC	This paper	97.2	79.9	68.3	46.6	

表 3-3 和各种著名的人脸识别系统的对比结果

Table 3-3 Recognition rate comparasitions with other well-known face recognition systems

需要说明的是,加权的 HGPP(HGPP_Weighted)方法利用 FERET 提供的 训练集合(共 1002 个样本)。LGBP_Weighted 方法发布最好的识别结果,它同 样是作者所在课题组的一个研究成果,该方法主要是利用 Gabor 幅值信息和 LBP 方法进行结合。从表 3-3 可以看到本文的方法取得了比其它方法更好的识 别性能,但是在 Dup I和 Dup II集合上,可以看到现有的这些方法在这两个集 合上仍然不是很理想。

3.5.2 CAS-PEAL-R1 数据库

CAS-PEAL 数据库是一个大型人脸数据库^[92,93],包括 1040 人以及 99594 张图片,涉及到姿态,表情,饰物,老化等变化,主要是用来衡量基 于东方人的人脸识别系统的性能。本文选取它的一个子集 CAS-PEAL-R1, 在文献[20]中给出了该子集的测试协议。Gallery 数据库中每人一张标准照片 (1040 张图片),并且选择如下测试子集合:饰物(Accessory),背景 (Background),距离(Distance),表情(Expression),老化(Aging),光照 (lighting)来检验算法的识别性能。训练集合和以上的集合没有交集,共包括 通过随机选择的 1200 张图片来训练识别模型。有关该数据库的详细信息可 以从网址 <u>http://www.jdl.ac.cn/peal/index.html</u>得到。



图 3-20 CAS-PEAL 数据库上样本 Fig.3-20 Samples of the CAS-PEAL database 表 3-4 CAS-PEAL-R1 数据库首选识别率的对比

Table 3-4 Rank-1 recognition rate comparasition on the CAS-PEAL-R1 database

Methods Probe sets	Eigenface	Fisherface	GFC	LGBPHS	HGPP	HGPP_Weighted
Accessory	37.1	61.0	85.1	86.8	91.9	92.5
Lighting	8.2	21.8	44.3	51	61.7	62.9
Expression	53.7	71.3	92.9	95.2	96.4	96.8
Distance	74.2	93.5	1.00	98.9	99.6	99.6
Background	80.5	94.4	98.9	98.7	99.6	99.8
Aging	50	72.7	93.9	1	96.9	98.4

从表 3-4 我们可以看到,本文提出的方法要远远好于其他的方法,尤 其在饰物,光照以及表情等大的测试集合上取得非常好的结果。其他的对比 方法的实验结果都来源于[21]。其中 GFC 利用的是 Gabor 幅值信息,与其 相比,本文提出的基于 Gabor 相位信息的方法得到更加鲁棒的识别性能,所 以如果正确的使用相位信息是能够得到好的结果的。

3.6 本章小结

本章首次提出了一种基于 Gabor 象限信息的人脸识别方法,该方法最 大的优点在于它无需训练,从而避免了泛化分析问题。同时,通过大量实验 验证了 Gabor 相位信息更具有鲁棒性,取得了远远好于基于 Gabor 幅值特征 的各种经典人脸识别方法的性能。基于生物学知识,一个人的眼睛、鼻子等 部位对识别是非常有利的,所以本文利用 Fisher 原则对人脸图像进行加 权,从而提高 HGPP 方法的性能。针对该方法特征维数较高的问题,本节 采用基于均匀量化的方法,实验表明该方法是非常有效的。另外,正如前文 所述,一个好的方法不仅仅要泛化能力强,而且还要能够利用给定的训练集 合所包含的先验信息,因为在条件控制情况下的人脸识别技术更符合目前人 脸识别技术的现状。第5章将着重探讨如何利用统计学习技术来提高 HGPP 的识别性能。
第4章 局部异或算子的一般性扩展

由上一章可以知道局部异或算子(LXP)算子仅仅适用于二值图像,那么 是否LXP不能应用到一般灰度图像呢?本章提出了一种新的方法,使得LXP 在识别领域更具适用性。该方法的核心思想在于如何把一幅图像变成一幅或 者多幅二值图像,从而可以直接利用LXP算子。

鉴于梯度信息可以用来反映像素点之间的大小关系,并且通过结合局 部信息可以获得图像的纹理描述^[94],本文基于梯度方向量化的方法把一幅 灰度图像变成了4幅二值图像。在此基础上,利用LXP算子来提取局部异或 模式,实际上反映了相邻像素点梯度方向的一致性。本章将首先回顾一下 LBP算子。

4.1 局部二值模式(LBP)

LBP算子是一种强大的纹理描述工具,每一个像素点和其邻域点进行 大小对比,然后对这个结果进行二值量化(给定阈值),则得到一种纹理描述。基本LBP的邻域范围是3x3,后来通过指定半径,LBP算子可以衡量不 同数目的邻域关系。图4-1给出了一个基本LBP算子的运算过程,它可以看 作是对梯度方向的一种量化,以及对邻域关系的描述^[94]。



图 4-1 局部二值模式(LBP)算子的示意图 Fig.4-1 Ilumination of Local Binary Pattern(LBP) operator

在文献[41]中,张文超结合 Gabor 幅值特征和 LBP 算子,取得了非常好的识别性能。该方法的主要思想是利用 Gabor 变换来增强特征,然后利用 LBP 算子提取微模式(micro-pattern),最后用空间直方图来表示人脸图像。

4.2 局部异或模式(LXP)

P(*Z*)是一个二值的图像或者数据,则本文定义 LXP 微模式:

 $LXP(Z_0) = [P(Z_0) \text{xor} P(Z_1), P(Z_0) \text{xor} P(Z_2), \dots, P(Z_0) \text{xor} P(Z_8)]$ (4-1)

这里 Z_i, i = 1,2,...,8, 是 Z₀ 的 8-邻域点, XOR 表示异或运算,这个运算本 文称之为局部异或模式算子(LXP),详见上一章内容。8-邻域"碰巧"提供 8 个比特从而每一点得到一个字节,因而可以用一个在[0,255]之间的整数来 表示一个模式。如果 1 或 0 代表了某种特性,LXP 实际上表示邻域点是否 具有不同的特性。那么是否 LXP 算子仅仅适用于二值图像呢?是否可以像 LBP 一样更具一般性。下一节将提出一种泛化 LXP 的方法,即 LXP 能够适 用于一般的灰度图像。

4.3 基于梯度方向二值化的 LXP 模式

这一节将首先对输入图像沿着不同的方向计算它们的梯度方向信息, 然后对他们的结果进行量化,本文称之为梯度方向二值化。主要采用4个方 向对梯度方向进行二值化,其定义分别如下所示:

对于0度方向,图像I梯度方向二值化的结果分别表示成如下的形式:

$$GZ_{0^{\circ}}(i,j) = \begin{cases} 0, & \text{if } I(i,j) > I(i+1,j) \\ 1, & \text{if } I(i,j) <= I(i+1,j) \end{cases}$$
(4-2)

对于 45 度方向,图像I梯度方向二值化的结果分别表示成如下的形式:

$$GZ_{45^{0}}(i,j) = \begin{cases} 0, & \text{if } I(i,j) > I(i-1,j+1) \\ 1, & \text{if } I(i,j) <= I(i-1,j+1) \end{cases}$$
(4-3)

对于 90 度方向,图像I梯度方向二值化的结果分别表示成如下的形式:

$$GZ_{_{90^{\circ}}}(i,j) = \begin{cases} 0, & \text{if } I(i,j) > I(i,j+1) \\ 1, & \text{if } I(i,j) <= I(i,j+1) \end{cases}$$
(4-4)

对于 135 度方向,图像I梯度方向二值化的结果分别表示成如下的形式:

$$GZ_{_{135^{\circ}}}(i,j) = \begin{cases} 0, & \text{if } I(i,j) > I(i+1,j-1) \\ 1, & \text{if } I(i,j) <= I(i+1,j-1) \end{cases}$$
(4-5)

对于任何一个梯度方向二值化的结果即*GZ*,我们都可以采用 LXP 算 子来获得一种纹理描述。实际上是衡量了相邻点的梯度方向是否具有一致 性,而 LBP 是对梯度方向量化结果邻域关系的直接描述。下面解释一下该 运算对于原始图像的含义, 针对 0 度方向(水平方向)连续三个点,LXP 的 1 和 0 分别表示如下含义,如图 4-2 所示。





图 4-2 LXP 的含义示意图 a) '1'表示的含义(LXP) 0 表示的含义

Fig.4-2 Illustration of LXP, a) the meaning for '1' b) the meaning for '0'

从以上的图示可以看到 LXP 运算能够衡量 3 个点之间梯度方向的一致 性。而梯度方向的一致性可能表示边缘的存在,LBP 能够告诉我们边缘附 近的梯度方向(可以理解为图 4-2 b) 所示的两个结果)。而实际上这个方向是 不稳定的,它很容易受到光照等因素的影响,但是无论其梯度方向如何,这 个边缘信息是相对稳定的。所以在这一点上 LXP 是有利的,但是还不足以 说明其优越性。LXP 还能够衡量是否存在拐点,不关心它是否是一个亮点 还是暗点,这种特性诚然会有一定的信息丢失,但是也会带来对光照等因素 的鲁棒性。我们还发现在人脸识别或者是纹理分析等领域得到广泛应用的 Gabor 特征同样具有这样的特性,对于一个亮点或者是暗点其 Gabor 幅值的 结果是一样的,因为它可以看作是一种特征增强的有力工具。从另一方面说 明了在人脸识别领域,LXP 是具有适用性的。从直觉上来说,本文提出的 LXP 算子是可以很好的与 Gabor 特征进行结合的,原因如下:

(1) Gabor 作为一种特征增强工具,不同于原始灰度图像,增强特征 附近的梯度方向可以不考虑,而只关心其分布即可。所以 Gabor 的幅值可以 和 LXP 结合,从而刻画边缘,孤立点等位置信息。而且 Gabor 特征对孤立 点的建模和 LXP 具有一致性。

(2) LBP 是对梯度方向量化的结果进行简单的联接,而 LXP 能够很好的描述相邻像素点之间梯度方向的相互关系。

如图 4-3 所示,首先对原始图像进行 Gabor 变换,对于幅值部分计算 4 个梯度方向的量化结果,然后利用 LXP 算子得到可视化结果图。



a)

		3557		
WIN .	K)	12		

b)(续)

		家		

c)



d)

图 4-3 基于 Gabor(5 个尺度 8 个方向)幅值特征的 LXP 示意图 a)表示梯度方向是 0 度时 LXP 结果 b)表示梯度方向是 45 度时 LXP 结果 c)表示梯度方向是 90 度时 LXP 结果 d)表示梯度方向是 135 度时 LXP 结果

Fig.4-3 Illustraiton of Gabor Magnitude based LXP results (5 scales, 8 orientations), a) LXP results for 0-degree gradient direction b) LXP results for 45-degree gradient direction c)

LXP results for 90-degree gradient direction d) LXP results for 135-degree gradient

direction

对于 Gabor 相位信息, LXP 也同样可以获得一种人脸图像的表示。相位信息往往能提供更多的细节信息,但是正如[8]所述,它具有不稳定性。 所以不能直接利用 Gabor 相位信息,本节同样利用了其梯度量化的结果,然 后再计算 LXP 特征。后续的实验表明,该方法是非常有效的,也证明了梯 度方向量化的结果是具有稳定性的,而且在人脸识别领域具有适应性。下面 我们给出其可视化的结果,如图 4-4 所示。

		The sea	Č, .	

	a)		
	· · · ·	nië The sin The sine	5	

b) (续)

					Meeting Alexandre	
			酸			
		C	c)			
	115					
					S.	
Ż			教			
				a and a second s		

d)

图 4-4 基于 Gabor(5 个尺度 8 个方向)相位特征的 LXP 示意图 a)表示梯度方向是 0 度时 LXP 结果 b)表示梯度方向是 45 度时 LXP 结果 c)表示梯度方向是 90 度时 LXP 结果 d)表示梯度方向是 135 度时 LXP 结果

Fig.4-4 Illustraiton of Gabor Phase based LXP results(5 scales, 8 orientations), a) LXP results for 0-degree gradient direction b) LXP results for 45-degree gradient direction c) LXP results for 90-degree gradient direction d) LXP results for 135-degree gradient

direction

4.4 基于 LXP 的空间直方图

人脸识别的误配准问题使得我们考虑利用基于区域的方法进行建模。类 似于 LBP, LXP 也是一种模式刻画算子,它可以反映图像边缘,孤立点等的位 置。而且这些图像具有很好的纹理特征,直方图被广泛用来提取纹理特征。本 文将进一步利用空间直方图来提取判别特征。由于单一的直方图会失去很多结 构信息,以人脸图像为例,它本身是一个拓扑图像,这就意味着空间结构信息 对人脸识别是尤其重要的。本文首先把 LXP 图像分成互不相交的几个子区 域,即 *R*₁,...,*R*_L,然后针对每一块抽取局部的直方图特征,这样所有的直方 图特征连接成为一个扩展的向量 (HGLXP),即:

 $H_{LXP} = (H_{LXP}(u, v, l) : u = 0, ..., 7; v = 0, ..., 4; l = 1, ..., L)$ (4-6)

这里 L 是子区域的个数。

4.5 进一步分析

至此,本文提出了基于 Gabor 特征的 LXP 空间直方图方法。同样的, 该方法基于 Gabor 小波,所以它继承 Gabor 变换的特性,比如多尺度,多方 向,以及其局部特性等等,因而保留大量判别信息。该方法也可以避免由于 训练引起的泛化能力较弱的问题,这在统计学习里还无法克服。

尽管该方法的计算很简单,但是它的特征维数很高,所以复杂度相对 来说也比较高。本节仍然采用均匀量化方法来解决这个问题。除此以外,本 章还验证了相位的量化结果能够提供相对稳定的信息,所以这在某种程度上 解决了相位信息在人脸识别领域应用时不稳定的问题。

4.6 基于 LXP 的人脸识别

本章仍然采用直方图的交作为最近邻分类器的相似度计算公式。基于此,两个人脸图像的直方图特征 *H*1,*H*2之间的相似度如下所示:

$$S(H1_{LXP}, H2_{LXP}) = \sum_{u=0}^{7} \sum_{v=0}^{4} \sum_{l=1}^{L} S_{Hl} (H1_{LXP} (u, v, l), H2_{LXP} (u, v, l))$$
(4-7)

从下一节实验可以看到利用上面相似度计算公式,识别性能已经很 好,但是不能够利用训练集合中先验信息,而传统的统计学习方法尽管存在 泛化能力较弱的问题却可以通过学习获取先验知识,所以在下一章提出另外 一种策略来进一步提高系统的识别性能。它说明了本文提出的方法能够和传 统的统计学习理论进行结合。

4.7 实验分析

这一节,基本 FERET 数据库被用来验证各种方法的性能。该数据库被 广泛地用来测试人脸识别算法的性能,在本文实验中,首先根据眼睛的位置把 图像归一化。根据 FERET 测试协议,实验中采取同样代表集合(gallery)和测试 (probe)数据库。首先基于灰度图像对两种局部模式算子的性能进行了比较, 然后基于 Gabor 特征进一步验证本文提出方法的性能。

4.7.1 基于灰度图像的 LBP 和 LXP 的实验对比

为了保留更多空间结构信息,对于 88*88 大小的图像,两种方法的子 区域大小都选择为 8x11,都采用均匀量化方法对直方图特征进行降维,本 文采用 16 个 Bins。事实上参数的选择有很多的经验可以借鉴,比如 LBP, HGPP 和 LGBPHS 方法。

如表 4-1 所示,基于原始灰度图像的实验表明本文方法和 LBP 具有相近的性能,但是在 Dup II 数据上,LXP 取得相对较好的性能。Dup II 数据反映了人的年龄变化的数据集合(一年以上),实际上也包括了光照、姿态等变化。

基于灰度图像边缘部分的实验表明 LXP 更适合对边缘信息进行建模, 首先对原始输入图像利用 Sobel 算子来提取边缘信息,然后分别利用 LBP 和 LXP 算子提取判别特征,其对比结果如表 4-2 所示:

表 4-1 基于灰度图像的 LBP 和 LXP 结果对比(16 Bins)

Table 4-1 Comparison results for LBP and LXP based on Grey-level images (16 bins)

>+	FERET probe sets					
力伝	Fb(%)	Fc(%)	Dup I(%)	Dup II(%)		
Image_LXP	92.9	59.8	52.1	35.9		
Image_LBP	92.9	50	49.7	25.2		

从表 4-1 和表 4-2 可以看出, Sobel 算子之所以能够提高这两种方法的 识别率,是因为这两种算子可以很好的对边缘信息进行建模,尤其 LXP 算 子的优势更明显。当对 Sobel 边缘提取后的结果进行建模时, LXP 的效果要 明显好于 LBP,这说明了 LXP 能够更好地反映边缘等信息的分布。

表 4-2 基于 Sobel 算子的 LBP 和 LXP 结果对比

		FERET probe sets				
力法	Fb(%)	Fc(%)	Dup I(%)	Dup II(%)		
Image_LXP_Sobel	91.5	79.9	56.9	53.8		
Image_LBP_Sobel	87.8	76.8	52.4	47.4		

Table 4-2 Comparison results of LBP and LXP based the Sobel operator

4.7.2 基于 Gabor 特征的 LBP 和 LXP 的实验对比

为了进一步验证 LXP 算法的性能,本文利用 Gabor 特征对原始图像进行特征增强,然后对它进行建模。实验结果表明 LXP 的性能要远远好于 LBP 的性能,尤其在 Dup 数据集合上,如表 4-3 所示,所以可以知道本文 的方法是一种很好的对增强信息进行建模的工具。

表 4-3 基于 Gabor 特征的 LBP 和 LXP 结果对比

Tuble 1 5 Comparison Testatis of EET and ETT bused on Subor Feduare							
Probes	Fb(%)	Fc(%)	Dup I(%)	Dup II(%)			
Models							
GMLXP	96.7	97.9	71.6	71.4			
LGBPHS*	94.1	97	58.8	49			
GPLXP	97.4	98.5	75	75.2			
Phase_LBP	96.2	97	70.4	67.9			
HGLXP	97.7	98.9	77	76.4			

Table 4-3 Comparison results of LBP and LXP based on Gabor feature

*表示是作者得到的实验结果

LGBPHS 是 Gabor 幅值特征和 LBP 相结合,而 HGLXP 是 Gabor 幅值特征和 LXP 相结合。需要说明的是 LGBPHS 的结果和[16]略有差别,是因为实验条件不一样。GPLXP 和 GMLXP 分别表示基于 Gabor 相位和幅值信息的人脸识别方法,而 Phase_LBP 表示利用相位信息量化结果在人脸识别领域是非常稳定的,远远好于 LGBPHS 人脸识别方法。HGLXP 表示 Gabor 相位和幅值信息都与 LXP 特征相结合的方法。

下面这一组实验则是表明本文提出方法并非是由于模板数是 LGBPHS 4 倍而导致较好的识别性能。如表 4-4 所示,其中 0°,45°,90°,135°表示 只有一个梯度方向使用时的识别率。

表 4-4 基于 Gabor 幅值部分的 LXP 特征(单一方向)和 LGBPHS 的对比 Table 4-4 Comparison results of GMLXP(sigle direciton) and LGBPHS based on Gabor

	Fb(%)	Fc(%)	Dup I(%)	Dup II(%)
LGBPHS	94.1	97	58.8	49
GMLXP(0°)	95.2	95.4	65.9	63.7
$GMLXP(45^{\circ})$	95.8	96.4	67.9	68.3
GMLXP(90°)	95.3	97.4	70.1	65.4
GMLXP(135 [°])	95.3	97.4	69.2	66.7

Magnitude

4.7.3 空间直方图特征的降维

正如前文所述,本文方法存在一个缺点,即特征维数过高。通常来 说,针对均匀量化方法,可以经验选择直方图的 Bins 数来解决这个问题。该 方法虽然能够提高系统对噪声的鲁棒性,但是会损失一些信息,这一小节通 过实验观测从而得出一些结论,将观察一下均匀量化方法是如何影响系统性 能的。

从表 4-5 和表 4-6 可以看出,当直方图的 bins 数大幅度减少以后,识别率并没有受到太大的影响。其中表 4-5 是对基于灰度图像的系统识别率变化的观测,而表 4-6 是对基于 Gabor 特征的系统识别率变化的观测。

表 4-5 Image_LXP 的直方图 Bins 数和识别率的关系

Table 4-5 The relationship between the number of histogram bins and recognition rates for Image LXP

Probes Bins	Fb (%)	Fc (%)
32	91	63.4
16	91	63.9
8	90.4	61.9.

表 4-6 HGLXP 的直方图 Bins 数和识别率的关系

Table 4-6 The relationship between the number of histogram bins and recognition rates for

HGLXP

Probes Bins	Fb (%)	Fc (%)
32	96.8	98.5
16	96.7	97.4
8	96.7	97.4

4.7.4 基于基本 FERET 测试的实验对比

另外,表4-7中的对比实验结果大部分来自于已发表的论文,其中包括基于 Gabor 特征的几种著名人脸识别系统,如 USC_EGBM, LGBPHS 和 GFC。 另外,除了 GFC 方法,其他方法都是基于空间直方图特征的人脸识别方法。 本文提出的方法也是基于 Gabor 特征和直方图特征的方法。所以在大型数据库 上进行性能对比能够很好地衡量本节提出方法的性能。

LGBPHS^[56]方法在基本 FERET 测试中取得非常好的识别效果,该方法主要是利用 Gabor 幅值信息和 LBP 方法进行结合,最终提取空间直方图作为人脸图像的特征表示。而本文提出的方法基于局部异或算子,并且结合 Gabor 幅值和相位信息。从表 4-7 可以看到本文提出的 HGLXP 方法取得了比其它方法

更好的识别性能,而且无需训练从而避免了泛化分析问题。 表 4-7 和各种著名的人脸识别系统的对比结果(FERET)

Table 4-7 Comparison results with other well-known face recognition systems(FERET)

>+			FERE	Γ probe sets	
万法		Fb(%)	Fc(%)	Dup I(%)	Dup II(%)
HGLXP	This paper	97.7	98.9	77	76.4
LGBPHS	ICCV2005 [57]	94.0	97.0	68.0	53.0
LBP_Weighted*	ECCV2004 [56]	97	79	66	64
GFC	*	97.2	79.9	68.3	46.6
USC EGBM	TPAMI2000 [22]	95	82	59.1	52.1

4.7.5 基于 CAS-PEAL-R1 实验对比

本节在大型数据库 CAS-PEAL-R1 上对 HGLXP 方法进行了验证,如表 4-8 所示,实验结果表明该方法具有非常好的识别性能。

表 4-8 和各种著名的人脸识别系统的对比结果(CAS-PEAL-R1)

Table 4-8 Comparison results with other well-known face recognition systems

Methods Probe sets	Eigenface	Fisherface	GFC	LGBPHS	HGLXP
Accessory	37.1	61.0	85.1	86.8	91.9
Lighting	8.2	21.8	44.3	51	59.7
Expression	53.7	71.3	92.9	95.2	96.1

(CAS-PEAL-R1)

4.8 本章小结

本章提出一种新的纹理图像描述算子LXP,该算子是一种对梯度方向 相互关系的一种刻画。对于原始图像的刻画时,不考虑梯度的方向,而只反 应边缘,线条等的分布信息。所以对基于Sobel算子的结果进行建模的时 候,LXP的效果更好。对于原始图像Gabor变换结果进行建模的时候,本文的方法也表现出更大的优势,这是因为LXP对于边缘或增强特征的结果进行 建模的时候,可以不考虑边缘附近的梯度方向信息。在识别的过程中,主要 依赖于以直方图的交作为相似度计算公式,采用最近邻分类器进行决策。

不同于传统的基于 Gabor 幅值特征的人脸识别方法,本章提出了对 Gabor 幅值和相位信息同时进行编码的方法。首先对局部异或运算算子进行 泛化处理,使得该算子更具适用性,最后结合空间直方图特征得到最终的对 象表示方法。在识别过程中,主要依赖直方图的交作为相似度计算公式,并 采用最近邻分类器进行决策。在基本 FERET 和 CAS-PEAL-R1 数据库上实 验表明该方法是非常有效和鲁棒的。总的来说,本章提出了一种把 LXP 算子 进行一般性扩展的方法,并结合 Gabor 特征应用到人脸识别领域。

尽管本文的方法取得了非常好的识别性能,但是仍然有改进的空间, 比如特征维数较高,所以可以进一步提出特征选择方法来解决这个问题,另 外其它的判别分析方法也能够用来提高该系统的性能。

下一步工作可以尝试应用到其他物体的识别上,而下一章的内容主要 集中在如何利用统计学习方法提高本节提出方法的性能上。

第5章 基于核学习的人脸识别方法

核学习方法在人脸识别领域取得了非常成功的应用。但是在其应用中 一个核心的问题就是核函数选择问题。在[15,16,38]中,研究者进行了很多 尝试,试图构建一个好的核函数,从而提高系统的性能。在人脸识别领域, 核判别分析方法属于子空间方法,本章将简单阐述子空间方法的技术细节, 指出其不足与优势,从而为本章的工作奠定了理论基础。本章尤其关注支持 向量机和核 Fisher 之间的内在联系,并给出证明,在此基础上提出基于最 近邻样本空间的判别分析方法。另外,文献[16,17]提出了基于直方图特征的 核函数,于是启发我们利用核学习方法结合直方图特征,从而使得本文提出 的 HGPP 和 HGLXP 能够利用训练集合中的先验信息,解决了它们如何进行 训练的难题。

5.1 基于子空间人脸识别方法

子空间分析方法是统计模式识别方法中一类重要的方法,本质上是一种 特征提取和选择的方法。它的思想就是根据一定的性能目标来寻找一个线性 或非线性的空间变换,把原始信号数据压缩到一个低维子空间,使数据在子 空间中的分布更加紧凑,为数据的更好描述提供了手段,另外,计算的复杂 度也得到了大大降低。本章将详细介绍几种重要的子空间分析方法以及它们 在人脸识别中的应用。

为了方便讨论,先给出一些重要定义和基本假设:

假设已知*C*个不同的人,每人分别有*N_i*(*i*=1,2,…,*C*)幅人脸图像,则 总共有*N* = $\sum_{i=1}^{c} N_i$ 幅图像,即有*C*类*N*个已知样本。*N*个样本可以简单的表 示为 **x**_i(*i*=1,2,…,*N*)。另一种表示方法 **x**_i(*i*=1,2,…,*C*;*j*=1,2,…,*N*)包含 了类的信息在内,其中*i*和*j*表示该样本是第*i*类中的第*j*个样本。**x**为所有 样本的总体平均, **x**_i = $\frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{x}_{ij}$ 为第*i*(*i*=1,2,…,*C*)类样本的类内平均。 任意待识别的人脸图像 x,识别问题即判断 x 属于 C类中的哪一类,是 C类分类问题;确认问题即判断 x 与某一给定的 x_i (i=1,2,...,N)是否属于 同一类,是两类分类问题。子空间分析方法的目的就是寻找一个变换 W, 将原样本从 n 维样本空间投影到一个 m(m < n) 维的子空间(特征空间), 即: $y_i = W^T x_i$ (i=1,2,...,N),并且把待识别样本也投影到该子空间,即 $y = W^T x$,然后在子空间上利用 y_i (i=1,2,...,N)对 y 进行分类。

5.1.1 主元分析方法

主元分析的思想来源于K-L 变换(Karhunen Loeve Expansion),目的是 找一组最优的单位正交向量,即所谓的主元,作为子空间的基,那么就可以 用它们的线性组合来重建原样本,并使重建后的样本和原样本的误差最小。 首先考虑对N个已知的n维样本向量 \mathbf{x}_i (i=1,2,...,N)的描述,显然它们可 以表示为n个单位正交向量 \mathbf{u}_i (j=1,2,...,n)的线性组合,主元分析的目的 就是寻找一组能够最好地描述这N个已知的n维样本向量 \mathbf{x}_i (i=1,2,...,N) 的分布的单位正交向量,即所谓的主元(Principal Component),使得用这些 主元来表示原向量误差最小。

已知样本的协方差矩阵为:

$$\mathbf{S}_{i} = \frac{1}{N} \sum_{i=1}^{N} \widetilde{\mathbf{x}}_{i} \widetilde{\mathbf{x}}_{i}^{T}$$
(5-1)

其中 $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \bar{\mathbf{x}}(i = 1, 2, \dots, N)$ 。则对散度矩阵 \mathbf{S}_i 进行主成分分析,如式(5-2) 所示:

$$\mathbf{S}_{\mathbf{V}} = \mathbf{V}\Lambda \tag{5-2}$$

其中V的列向量是 S_i 的特征向量,V称为 S_i 的特征向量矩阵, Λ 是对

角矩阵,其对角线元素是 $\mathbf{S}_{,}$ 的特征值, Λ 称为 $\mathbf{S}_{,}$ 的特征值矩阵。

主元分析基于信号表示的最优化原则,提取的特征能够保证最有效的 表示原信号(均方误差最小),所以又叫最优表示特征(Most Expressive Feature/MEF)。但是从分类角度来看,最优的表示并不一定能够保证最优的 可分性^[78],为了更好地分类,我们需要的是最优判别特征(Most Discriminating Feature/MDF)^[78]。下面将介绍线性判别分析方法。

5.1.2 线性判别分析方法

线性判别分析,尤其是 Fisher 原则在模式识别领域具有非常重要的地位,不同于主成分分析着重于信息的完整性,Fisher 原则是用来衡量特征的可区分性,所以更加适合于分类问题。

5.1.2.1 原理

线性判决分析以使样本的可分性最好为目标,寻找一组最优的线性变换使每类的类内离散度最小,并且使各类间的离散度达到最大,因此从理论上来说,比较适用于模式识别问题。经典的线性判别分析中使用的是 Fisher 准则函数,所以线性判决分析又被称为 Fisher 线性判别分析(Fisher Linear Discriminant Analysis, FLDA)^[97]。

为了定义样本的可分性,首先定义两个散度矩阵:

类间散度矩阵(Between-class Scatter Matrix):

$$\mathbf{S}_{b} = \sum_{i=1}^{C} p(\boldsymbol{\sigma}_{i})(\overline{\mathbf{x}}_{i} - \overline{\mathbf{x}})(\overline{\mathbf{x}}_{i} - \overline{\mathbf{x}})$$
(5-3)

类内散度矩阵(Within-class Scatter Matrix):

$$\mathbf{S}_{w} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i}) E\{((\boldsymbol{x}_{i} - \overline{\mathbf{x}}_{i})(\boldsymbol{x}_{i} - \overline{\mathbf{x}}_{i})^{T}) \mid \boldsymbol{\varpi}_{i}\}$$
(5-4)

 $p(\boldsymbol{\omega}_i)$ 是先验概率,容易证明: $\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_w^{[99]}$ 。所以样本协方差矩阵 \mathbf{S}_t 又称为总体散度矩阵(Total Scatter Matrix)。在不需要加以区别的时候,我们把 \mathbf{S}_b , \mathbf{S}_w 和 \mathbf{S}_t 统称为散度矩阵。线性判别分析的目的是寻找一个子空间使类间散度尽可能大的同时类内散度尽可能小。为了构造判别准则,首先必须对以上散度矩阵进行定量度量,通常用矩阵的行列式(Determinant)或者迹

(Trace)来度量,这里我们使用行列式。所谓线性判别分析就是寻找一个最优 线性变换 W:

$$\mathbf{W} = \arg \max_{\mathbf{W}} \frac{\left| \mathbf{W}^{T} \mathbf{S}_{b} \mathbf{W} \right|}{\left| \mathbf{W}^{T} \mathbf{S}_{w} \mathbf{W} \right|}$$
(5-4)

W可以通过解广义特征值问题 \mathbf{S}_{b} W= \mathbf{S}_{w} WA求得。可以证明,当 \mathbf{S}_{w} 非奇异

时,最优线性变换W的列向量即为 $\mathbf{S}_{w}^{-1}\mathbf{S}_{b}$ 的特征向量。这组向量也称为最优 判别向量集。

5.1.2.2 计算问题: 解决小样本问题

传统的线性判别分析方法需要计算矩阵 **S**⁻¹_w**S**^b 的特征值和特征向量,由于矩阵求逆运算存在不稳定性,所以我们希望尽量避免直接的求逆运算,并且由于 **S**⁻¹_w**S**^b 未必是对称矩阵,其特征值分解运算也存在不稳定性,因此我们可以采用同时对角化(Simultaneous Diagonalization)的方法来产生一个稳定的计算过程^[77]。下面首先介绍一下增强的判别分析方法^[10,98]。

1. 增强的判别分析方法

首先对S_w进行特征值分解:

$$\mathbf{S}_{\mathbf{w}}\mathbf{U}^{\mathbf{F}} = \mathbf{U}^{\mathbf{F}}\boldsymbol{\Lambda} \tag{5-5}$$

其中, U是S_w的特征向量矩阵, A是S_w的特征值矩阵, 需要说明的是U对 应的特征值要大于 0。则对(UA^{-1/2})^TS_b(UA^{-1/2})进行特征值分解, V是 $(UA^{-1/2})^{T}S_{b}(UA^{-1/2})$ 的特征向量矩阵; $\Sigma \in (UA^{-1/2})^{T}S_{b}(UA^{-1/2})$ 的特征值矩阵。 很容易知道, $(UA^{-1/2}V)^{T} 和 UA^{-1/2}V 把S_{b} 和 S_{w} 分别对角化为 \Sigma 和 I, 这就是所$ 谓的同时对角化。则最终的变化矩阵W:

$$\mathbf{W} = \mathbf{U} \Lambda^{-1/2} \mathbf{V} \tag{5-6}$$

本文称该方法为增强的判别分析方法^[98]。

2. PCA+LDA

另一种广泛应用的方法就是结合主元分析和线性判别分析的方法,即 在 PCA+LDA 框架下计算判别空间。这种方法用于图像检索(Image Retrieval) 和人脸识别最早见于文献[78],并由 Belhumeur 等人发展为 Fisher 脸 (Fisherface)方法,W. Zhao 等人也对该方法进行了大量研究,并完成了实际 系统。他们的系统在 FERET 测试中取得了很好的表现,也是目前最为成功 的人脸识别系统之一。这种方法的思想是通过主元分析,首先获得原样本的 最优表示特征,在此基础上通过线性判别分析,获得原样本的最优表示特征 的最优判别特征。具体来说就是首先应用主元分析,将原样本维数降到小于

或等于 \mathbf{S}_{w} 的秩,通常是N-C,这样计算得到的样本类内散度矩阵就是非奇

异的,可以进一步采用标准的线性判别分析方法求得最终的子空间。需要指出的是,主元分析和线性判别分析遵循两种不同的准则,在第一步主元分析的时候丢掉的信息从信号表示来说是最无用的信息,但从样本分类的角度来说则未必是无用的。事实上,通过降维使 S_w 满秩,已经丢掉了 S_w 的零空

间,而S_w的零空间对于分类具有重要意义。

3. 零空间方法

Chen 等人在文献[33]中提出一种新的方法来解决线性判别分析中的小 样本问题,直接利用 S_w 的零空间来寻找最佳判别向量集,这里称为零空间 法。当 S_w 奇异时,必然存在非零向量q使得 $q^T S_w q=0$ 。事实上,如果存在 向量q使得 $q^T S_w q=0$ 且 $q^T S_b q \neq 0$,则判别准则 $|q^T S_b q|/|q^T S_w q|$ 必然达到最大 值,这样的q就存在于 S_w 的零空间上。传统的方法针对 S_w 奇异性,往往通 过直接或间接地去掉 S_w 的零空间,零空间法并不试图求解 $S_w^{-1}S_b$ 的特征向 量,相反恰好利用了被 S_w 变为零向量的向量集(S_w 的零空间)作为求解最优 判别向量集的基础。所以可以先求得 S_w的零空间,将原样本投影到 S_w的零空间上,这样投影后的样本的类内散度为零,再求使类间散度最大化的变换,即可得到最优判别向量集。则首先对 S_w进行特征值分解得到^[100]:

$$\mathbf{S}_{w}\mathbf{U} = \mathbf{U}\Lambda\tag{5-7}$$

由于 \mathbf{S}_{w} 奇异,即存在零特征值,去掉 Λ 中非零特征值所在的行和列,以及 U中对应于零特征值的特征向量所在的列,用得到 U_{1} 改变类间散度矩阵 $\mathbf{S}_{b}^{T} = \mathbf{U}_{1}^{T}\mathbf{S}_{b}\mathbf{U}_{1}$,然后计算 \mathbf{S}_{b}^{T} 的特征向量 \mathbf{V} 。当然,此时 $\mathbf{S}_{w}^{-1} = \mathbf{U}_{1}\Lambda_{1}^{-1}\mathbf{U}_{1}^{T}$ 不再成 立,但最后求得的变换 $\mathbf{U}_{1}\mathbf{V}$ 仍然满足:

$$(\mathbf{U}_{\mathbf{V}}\mathbf{V})^{\mathsf{T}}\mathbf{S}_{\mathsf{h}}(\mathbf{U}_{\mathbf{V}}\mathbf{V}) = \Sigma, (\mathbf{U}_{\mathbf{V}}\mathbf{V})^{\mathsf{T}}\mathbf{S}_{\mathsf{w}}(\mathbf{U}_{\mathbf{V}}\mathbf{V}) = 0$$
(5-8)

线性判别分析的根本目的是在最大化类间散度的同时最小化类内散度,由式(5-8)可知U₁V能够达到这一目的的。而上一小节介绍的增强判别分析方法恰好相当于在S_w零空间的正交补空间中应用线性判别分析。

4. 其他的零空间法

针对零空间法计算复杂的问题,刘等^[99]提出了一种改进方法。如前所述,如果存在向量q使得q^TS_wq=0且q^TS_bq≠0,则判别准则 |q^TS_bq|/|q^TS_wq|必然达到最大值,这样的向量q对分类非常有意义的。但 如果q^TS_wq=0且q^TS_bq=0,这样的向量q对分类却没有用处。可以首先找到 这个交集并把它去掉,可以很容易知道这个交集恰好是总体散度矩阵S_t的 零空间^[99]。可以首先去掉S_t的零空间,然后在一个较低维数的子空间中求 降维之后的类内散度矩阵的零空间,并在此基础上求最优判别向量集。实际 上说明了并不是整个S_w的零空间都对分类有意义,也就是S_w的零空间和S_b

的零空间的交集对分类就没有用处。该算法的计算复杂度和降维法相似,而 远低于直接的零空间法。至于算法的有效性,从理论上讲与零空间法是一样 的。关于该算法的具体细节可参考[99]。

5.1.2.3 类间散度矩阵的秩带来的问题

由前文可知类间散度矩阵的秩由于计算的对称性而与类别数相关。事 实上,这样会带来一些问题,首先会导致判别特征个数上的限制,其次在利 用所求的判别空间进行判别的时候,所需要的判别特征与训练过程中的类别 数不相关^[101,102]。有很多的研究成果来解决这个问题,其中有非参数方法, 作差法^[101,102]等等。基于决策边界矩阵的方法是其中一个比较典型的非参数 建模方法,该方法主要估计边界的概率密度,然后计算最小判别子空间用来 进行人脸识别。邱^[102]提出基于作差法的非参数判别分析方法,很好地解决 了类间散度矩阵对于判别分析的限制问题。

5.1.3 基于核技术的子空间分析方法

近年来,核方法已成为模式识别领域一个迅速发展的研究方向。它的 主要思想是由 Vapnik 提出并应用于 SVM。其后,Mika 和 Baudat 把核方法 应用到判别分析领域,将 Fisher 线性判别分析方法进一步拓广到非线性情 形,提出了核 Fisher 方法。不同于前面介绍的方法在原样本空间进行子空 间分析,基于核技术的子空间分析方法则是通过某种非线性映射**Φ**,把原 样本映射到一个高维特征空间上再进行子空间分析。利用核技术使我们可以 在不进行明确的非线性映射情况下在任意高维空间上进行子空间分析,因此 也就可以对原样本的高阶统计特性加以考虑。

5.1.3.1 基于核技术的主元分析方法

基于核技术的主元分析(Kernel Principal Component Analasis, KPCA), 以下简称核主元分析,是将核技术与主成分分析相结合的一种非线性子空间 分析方法。

首先将原样本映射到一个高维的特征空间中,这个特征空间可以是任 意高维甚至无限维的,通常用一个隐函数Φ来实现从原始空间到高维空间的 映射。在实现的时候,无需知道Φ的具体表示形式。下面就需要在这个特征 空间中应用主元分析。首先需要对数据进行中心化,由于我们并不能明确知 道Φ是怎样一个映射,所以无法直接计算高维空间中的样本平均,但我们可 以通过间接的方式实现高维空间上的中心化^[103,104]。这里我们先假设数据已 经中心化,则特征空间中协方差矩阵为:

$$\mathbf{S}_{i}^{\bullet} = \frac{1}{N} \sum_{i=1}^{N} \Phi(\mathbf{x}_{i}) \Phi(\mathbf{x}_{i})^{T}$$
(5-9)

这里 $\Phi(\mathbf{x})$ 已经中心化,则相应的特征值问题为

$$\lambda \mathbf{v} = \mathbf{S}_t^{\Phi} \mathbf{v} \tag{5-10}$$

这里,
$$\mathbf{v} = \sum_{i=1}^{N} a_i \Phi(\mathbf{x}_i)$$
, 结合上述两式得到:
 $\lambda \sum_{j=1}^{N} a_j (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)) = \frac{1}{N} \sum_{j=1}^{N} a_j \sum_{k=1}^{N} (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_k)) (\Phi(\mathbf{x}_k) \cdot \Phi(\mathbf{x}_j))$ (5-11)

定义一个 $N \times N$ 的矩阵 $\mathbf{K} = (K_{ij})_{N \times N}, K_{ij} = (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)),$ 则我们只要解特征值问题,即:

$$N\lambda \mathbf{a} = \mathbf{K}\mathbf{a} \tag{5-12}$$

相应的特征向量为**a**₁,**a**₂,…,**a**_N,保留**m**个特征向量,为了进行特征度量, 我们需要对原样本进行两次投影,第一次用一个隐函数Φ来实现从原始空间 到高维空间的映射,第二次再将高维空间的样本投影到特征空间中的主元 上,即:

$$(\mathbf{v}_k \cdot \Phi(\mathbf{x})) = \sum_{i=1}^N a_i^k (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x})) \quad k = 1, 2, \cdots, m$$
(5-13)

关于高维空间中点积(Φ(**x**)·Φ(**y**))的计算,根据文献[14],我们采用核 函数来代替,即:

$$\mathbf{k}(\mathbf{x}, \mathbf{y}) = (\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})) \tag{5-14}$$

这样,我们可以在不进行第一步映射的情况下计算高维空间的点积。 在这里我们采用了多项式(Polynomial)核函数:

$$\mathbf{k}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^d = \left(\sum_{i=1}^n x_i \cdot y_i\right)^d$$
(5-15)

其中 d 为多项式次数。当然还有其他的核函数,例如高斯(Gaussian)核函数:

$$\mathbf{k}(\mathbf{x}, \mathbf{y}) = \exp(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2})$$
(5-16)

在不同的应用中如何选择核函数本身还是一个有待于研究的问题,这是因为 不同的核函数代表了不同高维空间。

核主元分析被应用于人脸识别,取得了一定的效果^[92]。基于这个结果,我们推断基于核技术的线性判别分析方法应该能在人脸识别上获得更好的效果,因此我们对这个问题进行了研究。

5.1.3.2 基于核技术的线性判别分析方法

类似于核主元分析,基于核技术的线性判别分析,简称为核判别分 析,就是将核技术与线性判别分析相结合的一种非线性子空间分析方法。

与核主元分析相同,核 Fisher 的目标是把样本映射到一个更高维空间 然后进行 Fisher 判别分析^[99]。用一个隐函数Φ来实现从原始空间到高维空 间的映射。类间散度矩阵S₄和类内散度矩阵S₂的定义如下所示:

$$\mathbf{S}_{b} = \sum_{i=1}^{C} \sum_{j=1, j \neq i}^{C} p(\boldsymbol{\varpi}_{i}) p(\boldsymbol{\varpi}_{j}) (u_{i} - u_{j}) (u_{i} - u_{j})^{T} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i}) (u_{i} - u) (u_{i} - u)^{T}$$
(5-17)

$$\mathbf{S}_{w} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i}) E\{((\Phi(x_{i}) - u_{i})(\Phi(x_{i}) - u_{i})^{T}) \mid \boldsymbol{\varpi}_{i}\}$$
(5-18)

这里 $u_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \phi(x_{ij})$ 表示类别i的均值向量, u表示所有样本的均值向量,

p(*σ_i*) 是先验概率。为了实现在高维空间 *F* 中判别分析,等价于最大化下面的表达式:

$$J(\mathbf{w}) = \frac{|\mathbf{w}^T \mathbf{S}_b \mathbf{w}|}{|\mathbf{w}^T \mathbf{S}_w \mathbf{w}|}$$
(5-19)

w = $\sum_{i=1}^{n} \alpha_{i} \phi(x_{i})$, 其中 α_{i} 是权向量,则核 Fisher 的优化准则变成如下表达式:

$$J(\boldsymbol{\alpha}) = \frac{|\boldsymbol{\alpha}^{T} \mathbf{K}_{b} \boldsymbol{\alpha}|}{|\boldsymbol{\alpha}^{T} \mathbf{K}_{w} \boldsymbol{\alpha}|}$$
(5-20)

其中核矩阵K_w和K_b的定义如下所示:

$$\mathbf{K}_{w} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i}) E(\boldsymbol{\eta}_{j} - \boldsymbol{m}_{i}) (\boldsymbol{\eta}_{j} - \boldsymbol{m}_{i})^{T}$$
(5-21)

$$\mathbf{K}_{b} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i})(\boldsymbol{m}_{i} - \boldsymbol{m})(\boldsymbol{m}_{i} - \boldsymbol{m})^{T}$$
(5-22)

其中 $\eta_j = (k(x_1, x_j), k(x_2, x_j), \dots, k(x_n, x_j))^T$, η_j 的类别标号与 x_j 相同, m_i 表示 类别i的 η_i 的均值向量, \bar{m} 是全部 η_i 的均值向量。

类似于线性判别分析,这个问题能够通过求解矩阵 $\mathbf{K}_{w}^{-1}\mathbf{K}_{b}$ 的特征向量 得到变换空间(本文称之为 GKFD)。利用求伪逆的方法来计算 $\mathbf{K}_{w}^{-1}\mathbf{K}_{b}$,然后 求解它的主成分构成的变换空间 $\boldsymbol{\alpha}$,这样我们可以用下面的表达式来计算核 判别特征:

$$v = (\mathbf{w}.\Phi(x)) = \sum_{i=1}^{n} \boldsymbol{\alpha}_{i} k(x_{i}, x)$$
(5-23)

需要说明的是,前面几节中介绍的解决小样本问题的方法在核判别分析中也适用,另外,核 Fisher 也可以基于 KPCA+FLD 框架来实现,具体可以参考[100]。该方法具有更快的速度,而且不会降低识别性能。后续章节将会对其进行进一步的扩展。

5.1.4 独立元分析方法

独立元分析(Independent Component Analysis/ICA)是对主元分析的非线 性推广^{[20][21]}。独立元分析原来多应用于盲源分离(Blind Source Separation/BSS)问题,后来被M. Bartlett用于人脸识别^[52,53]。两者的区别在 于主元分析的目标是最小化重建误差,而独立元分析的目标是最小化子空间 的基向量,也就是所谓的独立元(Independent Component)之间的统计相关 性;主元分析中的主元相互正交,而独立元分析中的独立元相互独立;从统 计意义上来讲,主元分析考虑的是二阶统计量,可以消除数据的相关性;而 独立元分析考虑的是高阶统计量,更进一步使数据独立,所以可以更加紧凑 的描述数据,事实上近几年来该方法获得越来越多的关注^[52,53]。

文献[53]对独立元分析、主元分析和线性判别分析应用于人脸识别的效

果做了比较,前者略好于后两者。而 K. Baek 等人最近在文献[105]中则对以 往的独立元分析在人脸识别上好于主元分析的评价提出了质疑,认为在适当 选取距离度量而使两种方法都达到最佳表现时,独立元分析不如主元分析有 效,而且在计算复杂度上独立元分析与主元分析相比也明显处于劣势(小时 数量级对秒数量级)。这个问题还有待于我们进一步研究探讨。

5.1.5 总结

需要强调的是,主元分析、线性判别分析和独立元分析都是线性的子 空间分析方法^[107],而核主元分析和核判别分析则是非线性的子空间分析方 法,原因就在于基于核技术的方法中存在非线性映射Φ。尽管我们并没有 进行显性的映射,而且在隐空间上我们采用了主元分析或线性判别分析这样 的线性子空间分析方法,但对于原样本来说,最终求得的子空间是非线性 的。

基于核技术的方法除了具有统计意义上的优势之外,其复杂度要高于 线性判别分析方法。但是在具体实现上,该方法并不复杂,因为我们通过核 函数可以有效地计算高维空间上的点积,而其他计算(例如特征值分解等)基 本上都是在 N×N 维的点积矩阵上进行的,因此没有实现上的问题。

5.2 支持向量机的零空间特性

统计学习人脸识别方法的框架包括特征提取及选择,判别分析和分类 器设计。其中判别分析是最为关键的步骤。核 Fisher 判别分析在模式识别 领域取得了成功的应用,它主要利用类均值向量的差进行判别分析^[106]。由 于 SVM 分类器是线性形式的,分类时等价于样本在 SVM 的法向量(即分类 面的法线方向,图 5-1 中的*h*(*x*))上进行投影,根据结构风险理论^[14],*h*(*x*) 是具有强判别分析能力的投影方向,该投影方向不依赖于原始类均值向量的 差。一般情况下,原始空间的样本投影到 SVM 的法向量形成一维空间,具

有非常好的可区分性。

除了经典的 Fisher 等判别分析方法之外,还有一些重要的判别分析方法,比如 Lee^[106]提出决策边界(Decision Boundary)理论,它的一个主要贡献 是通过构建决策边界特征矩阵(Decision Boundary Feature Matrix, DBFM), 然后对 DBFM 进行主成分分析来获取和原始特征空间具有同样判别能力的 最小维数的子空间^[106]。

本节的主要贡献是在高维空间利用 SVM 的法向量把 DBFM 扩展到非 线性情形,从而构建核化的 DBFM,然后结合核 Fisher 方法进行判别分 析。基于 SVM 的法向量具有零空间性质的这一特性,本文提出了在多类问 题情况下利用 SVM 的法向量构建核化的 DBFM 的方法,最后利用基于零空 间核化费舍尔判别分析方法来计算投影空间。为了检验这种方法的有效性, 本文把它应用到人脸识别领域。

本节的后续内容组织如下,首先证明 SVM 法向量的一些特性。然后提出基于支持向量的非线性判别分析方法。



图 5-1 x_1, x_2 是支持向量, h'(x)是法向量, Margin 是类问距

Fig.5-1 x_1, x_2 are support vectors, h'(x) is the normal vector, Margin is the distance between different classes

5.2.1 支持向量机

SVM 是统计学习理论中最年轻的部分, 1995 年才完成其主要的理论内

容,现在仍然在不断的发展中。SVM 的基本思想是根据 Vapnik 提出的结构 风险最小化(Structure Risk Minimization)原理,通过最大化分类间隔或边缘 (Margin)尽可能的提高学习机的泛化能力。它是从线性可分情况下的最优分 类面(Optimal Hyperplane)提出的,而在其定义空间线性不可分的情况下, SVM 则将其映射到一个更高维的空间变成一个线性可分的问题。考虑到求 解最优分类面算法的性质,在这个高维空间中,我们只需进行内积运算即 可,所以可以利用 kernel 技术,这样即使高维空间的维数增加很多,在其 中求解最优分类面的问题并没有增加多少计算复杂度。

SVM 本质上是一种二分类方法,而在现实世界中的大多数分类问题都 是多类问题,如本文讨论的人脸识别问题就是一个典型的多类问题,因此 SVM 具有很大的局限性,必须寻找一种 SVM 多分类方法,才能使 SVM 具 有实用价值。应用较多的是一对一(One-against-One)方法和一对多(Oneagainst-Rest)方法。假定样本集合包含 k 个类别,对于 One-against-One 方法 而言,需要建立 k(k-1)/2 个 SVM 分类器,对于 One-against-Rest 方法而 言,需要建立 k 个 SVM 分类器。本文主要采用了复杂度较小的 Oneagainst-Rest 方法计算支持向量集合和法向量。

支持向量集的性质是 SVM 理论非常重要的一部分,在文献[108,109] 中,Bernhard 等人仅仅利用支持向量集来训练分类器,就取得了与基于全部 训练样本得到的分类器同样或略好的性能。Rik 等人^[110]提出优化支持向量 机类间距(图 5-1 中的 Margin)的梯度的方法,并且成功地把这种方法应用到 人脸检测中,取得很好的实验结果。

5.2.2 支持向量机的零空间特性

为了简化问题,本节假定考虑的是两类问题,样本集合为 $\{(x_i, y_i) | x_i \in R^n, y_i \in \{-1, 1\}\}$,该集合分成正例集合和反例集合,SVM 分类面函数表示为:

$$h(x) = w^T x + b \tag{5-24}$$

其中 $w = \sum_{i}^{n} \alpha_{i} y_{i} x_{i}$, $\sum_{i}^{n} \alpha_{i} y_{i} = 0, \alpha_{i} > 0$, x_{i} 是支持向量, b是阈值。 则最大化两类之间的类间距等价于最小化如下表达式^[14]:

$$\underset{w}{\operatorname{Min}} \quad \frac{1}{2} w^{T} w \tag{5-25}$$

约束条件:
$$y_i(w^T x_i + b) - 1 \ge 0$$
 (5-26)

本文假定Φ(x_i)是x_i的高维空间表示,Φ是一个隐函数,无需知道函数具体 形式,则有如下目标函数:

$$\underset{w}{\operatorname{Min}} \quad \frac{1}{2} w^{T} w \tag{5-27}$$

约束条件:
$$y_i(w^T \Phi(x_i) + b) - 1 \ge 0$$
 (5-28)

其中 $w = \sum_{i=1}^{n} \alpha_i y_i \Phi(x_i)$ 。

若 $\Phi(x_i)$ 不是支持向量,则 $\alpha_i = 0$,否则 $\alpha_i \neq 0$,且式(5-28)的等号成立 [14],即:

$$y_i(w^T \Phi(x_i) + b) - 1 = 0$$
(5-29)

下面定义两个由支持向量构成的集合:

$$S_{1} = \{\Phi(x_{i}) \mid y_{i} = 1, \Phi(x_{i}) \text{ is the support vector } \}$$
(5-30)

$$S_{2} = \{\Phi(x_{i}) \mid y_{i} = -1, \Phi(x_{i}) \text{ is the support vector } \}$$
(5-31)

对任意 $\Phi(x_i) \in S_1$,有:

$$U_{1}^{\Phi} = \frac{1}{|s_{1}|} \sum_{\Phi(x) \in S_{1}} \Phi(x)$$
(5-32)

同理对任意 $\Phi(x_i) \in S_2$,有:

$$U_{2}^{\Phi} = \frac{1}{|s_{2}|} \sum_{\Phi(x) \in S_{2}} \Phi(x)$$
(5-33)

在两类的情况下,基于支持向量的类内散度矩阵定义如下所示:

$$\mathbf{S}'_{w} = \sum_{\Phi(x_{i})\in S_{1}} (\Phi(x_{i}) - U_{1}^{\Phi}) (\Phi(x_{i}) - U_{1}^{\Phi})^{T} + \sum_{\Phi(x_{i})\in S_{2}} (\Phi(x_{i}) - U_{2}^{\Phi}) (\Phi(x_{i}) - U_{2}^{\Phi})^{T}$$
(5-34)

则有:

$$w^{T}\mathbf{S}_{w}' w = \sum_{\Phi(x_{i})\in S_{1}} w^{T} (\Phi(x_{i}) - U_{1}^{\Phi}) (w^{T} (\Phi(x_{i}) - U_{1}^{\Phi}))^{T} + \sum_{\Phi(x_{i})\in S_{2}} w^{T} (\Phi(x_{i}) - U_{2}^{\Phi})) (w^{T} (\Phi(x_{i}) - U_{2}^{\Phi}))^{T}$$

$$= \sum_{\Phi(x_{i})\in S_{1}} (w^{T}\Phi(x_{i}) - w^{T}U_{1}^{\Phi}) (w^{T}\Phi(x_{i}) - w^{T}U_{1}^{\Phi}))^{T} + \sum_{\Phi(x_{i})\in S_{2}} (w^{T}\Phi(x_{i}) - w^{T}U_{2}^{\Phi}) (w^{T}\Phi(x_{i}) - w^{T}U_{2}^{\Phi})^{T}$$
(5-35)

由式(5-32),(5-33),(5-34),和(5-35),则有:

$$w^T \mathbf{S}'_w w = 0 \tag{5-36}$$

可以知道 SVM 的法向量在基于支持向量的类内散度矩阵前提下,具有零空间特性。在高维空间, SVM 的分类面函数表示为:

$$h(\Phi(x)) = w^T \Phi(x) + b \tag{5-37}$$

则它对 $\Phi(x)$ 的导数 $h'(\Phi(x))$ 是分类面的法向量即: $h'(\Phi(x)) = w$,则:

$$h'(\Phi(x)) = \sum_{i=1}^{n} \alpha_i y_i \Phi(x_i)$$
 (5-38)

其中 $\Phi(x_i)$ 是支持向量。由于 $y_i \in \{-1,1\}$,故 $h'(\Phi(x))$ 可以理解为以 α_i 为权重 计算支持向量集中属于不同类别的两个子集 S_1 和 S_2 的均值向量,然后作 差。假定两类样本的原始均值向量为 u_1 和 u_2 ,Fisher 的投影方向依赖于 $u_1 - u_2$ ^[107],且它被用来构建类间散度矩阵,所以,采用以差向量为表示形 式的 $h'(\Phi(x))$ 做判别分析是很自然的选择。

5.3 基于 SVM 的零空间核判别分析方法

本节首先简述 Decision Boundary 理论,该理论的具体细节部分可参考

[5]。然后提出构建 DBFM 和基于支持向量定义类内散度矩阵的方法,最后利用零空间判别分析方法计算投影空间。该方法实际上是在最近临样本空间进行判别分析。

5.3.1 决策边界特征矩阵

Decision Boundary 理论的一个最主要贡献是给出决策边界特征矩阵(D BFM)的定义和性质^[106]:

DBFM 定义: *S* 是 Decision Boundary 或分类面, N(x)是 Decision Boundary 上*x*点的法向量, p(x)是概率密度函数。则 DBFM 的定义如下所示:

$$\sum_{\text{DBFM}} = \int_{S} N(x) N^{T}(x) p(x) dx$$
(5-39)

DBFM 性质: DBFM 大于零的特征值对应的特征向量构成最本质的投影子空间,该子空间能够保留原始空间全部判别信息并且维数是最小的空间。在多类问题情形,DBFM 是基于 One-against-One 策略定义的。由于 $u_i - u_j$ 决定 Fisher 的投影方向,在 Fisher 原则中定义的类间散度矩阵(即式 (5-22))可以认为是利用类均值向量(等权重)的差构建的 DBFM。

在高维空间 F 中,我们利用 SVM 计算分类面,则其任意一点的法向量都是相同的而且是可计算的,基于此,下一节将提出构建核化的决策边界特征矩阵的方法。

5.3.2 基于 SVM 的零空间判别分析

本节提出在 F 空间以及多类问题情形定义核化的决策边界特征矩阵和 类内散度矩阵的方法。为了计算上的方便,本文采用 One-against-Rest 方 法。

$$\sum_{\text{KDBFM}} = \sum_{i=1}^{C} p(\boldsymbol{\sigma}_{i}) h_{i}^{T} (\Phi(x)) (h_{i}^{T} (\Phi(x)))^{T}$$
(5-40)

$$\mathbf{S}_{w}^{'} = \sum_{i=1}^{c} p(\boldsymbol{\sigma}_{i}) \sum_{k=0}^{n_{i}} p(\boldsymbol{\sigma}_{k} \mid \boldsymbol{\sigma}_{i}) E((\Phi(\boldsymbol{x}_{k}^{'}) - \boldsymbol{u}_{ik}^{'})(\Phi(\boldsymbol{x}_{k}^{'}) - \boldsymbol{u}_{ik}^{'})^{T})$$
(5-41)

 \sum_{DRM} 即决策边界特征矩阵。假定 S_i 表示在 C_i 构成正例集以及 $C_{i,ij}$ 构成反例

集时,利用两类 SVM 算法计算得到的所有支持向量构成的集合, $h_i(\Phi(x))$ 是分类面的法向量。 S_i 被分成两个子集合,分别代表来自不同 Margin 上的 样本构成的集合, $S_a \subset C_i$,这里 u'_{i0} 表示 S_a 的均值向量,而剩下的支持向量 构成集合 S_{i2} 。我们知道 S_{i2} 中的样本属于多个类别,假定其类别数为 n_i ,针 对每一个类别计算它们的类均值向量分别表示为 u'_{ik} , $k = 1,...,n_i$ (如果只有 一个样本,它自身就是该类别的中心)。这样我们可以利用多个均值向量来 表示 S_i 为 $(u'_{i0},u'_{i1},...,u'_{in})$, x'_k 代表以 u'_k 为中心的支持向量, $p(\sigma_i)$ 是先验概 率密度。因此,式(5-41)不同于传统的类内散度矩阵定义方法,每一类是由 一个类中心向量来表示。而本文考虑的是支持向量集合,我们利用多个支持 向量均值来表示它。需要说明的是:在 S_i 集合中,如果某个类别的支持向 量的个数大于 1,那么该类别对类内散度矩阵 S'_w 的计算是有贡献的。而且在 实验中我们发现在一对多的策略下, C_i 类中几乎全部的样本都包含在 S_i 中,这样就保证式(5-41)的计算是基于整个样本空间的,而且这种基于多个 均值向量定义类内散度矩阵的方法可以利用尽可能多的类内信息。

在 One-against-Rest 策略下,反例集合的样本偏多,为了解决这个问题,把反例集合分成 *k*-subpart 份,正例集合保持不变,则可以得到 *k*-subpart 个 SVM 的法向量和支持向量集来计算 DBFM 和基于支持向量类 内散度矩阵。

根据 式(5-40)以及第 5.1 节和 5.2 节的讨论可知, (u_i - u_{i0})也可用来构 建特殊的 DBFM。由前文的阐述,我们知道(u_i - u_{i0})和h_i(Φ(x))具有一致性, 即它们都是均值向量的差向量,但是它们提供不同类型的判别信息。据此结 合式(5-40)定义扩展决策边界特征矩阵(Ex-DBFM)。

85

$$\sum_{Ex-DBFM} = \sum_{i=1}^{C} p(\boldsymbol{\sigma}_{i})((u_{i}^{T}-u_{i0}^{T})(u_{i}^{T}-u_{i0}^{T})^{T} + h_{i}^{T}(\Phi(x))(h_{i}^{T}(\Phi(x)))^{T})^{(5-42)}$$

需要说明的是, $(u_i - u_{i0})$ 和 $h_i(\Phi(x))$ 都需要利用自身的模进行归一化。

根据 Decision Boundary 理论, DBFM 的大于零特征值对应的特征向量 w^* 满足:

$$(w^*)^T \sum_{\text{KDBFM}} w^* > 0$$
 (5-43)

 $<u></u>则 J(w[*]) = \frac{(w[*])^{T} \sum_{\text{DBFM}} w^{*}}{(w[*])^{T} \mathbf{S}_{w}^{*} w^{*}} = +\infty , 这些特征向量构成的投影空$

间具有最好的识别性能,所以可以利用基于零空间核化费舍尔判别分析方法 来计算得到^[99],则使得 $(w^*)^T \mathbf{S}_w w^* = \mathbf{0}$,这样可以尽可能使 w^* 满足结构风险 最小化理论。

在 F 空间, w的定义在整个样本空间,则很容易计算核化的决策边界特征矩阵以及 \mathbf{S}'_w 的核矩阵 \mathbf{K}'_w 。为了表述方便,设 \mathbf{K}'_b 为 DBFM 或者 Ex-DBFM,则利用零空间方法计算投影空间。另外,当 \mathbf{K}'_b 为 DBFM 时,称之 为 KDBA,而 \mathbf{K}'_b 为 Ex-DBFM 时,称之为 SV-KFD。

这一小节主要探讨了基于支持向量的判别分析方法,该方法实质是通 过支持向量机在高维空间寻找最近邻样本,后续的实验表明该方法是合理 的,能够提高系统的识别性能。针对 KPCA+FLD 框架核学习方法,下一节 将提出一个容易实现的基于最近邻样本空间的核判别分析方法。实际上在人 脸识别领域,容易实现的方法始终是研究者追求的目标。

5.4 基于最近邻样本空间的 Fisher 判别分析方法

实现核 Fisher 的另一个有效途径是 KPCA 和 FLD 相结合,本节基于这个框架来提高现有系统的效率,由于上一节利用离得最近的样本进行判别分析能够取得很好的识别效果,所以本文提出了 Nearest Neighbor-Fisher

Linear Discriminant analysis(NN-FLD)方法,如图 5-2 所示。当它与 KPCA 相结合时,本文称之为 Nearest Neighbor-Kernel Fisher Linear Discriminant analysis(NN-KFLD)方法。

▼ 为其中一类均值向量,而Y^N表示其余类别中离该均值向量最近的 那些样本(欧式距离),为了提高系统的效率我们只取 10 个最近邻样本。同 时,类内散度矩阵保持不变,而重新定义类间散度矩阵的形式如 式(5-44) 所示:

$$S_{b} = \sum_{i=1}^{C} p(\boldsymbol{\varpi}_{i}) \sum_{j=1}^{10} ((\bar{X}_{i} - Y_{ij}^{N})(\bar{X}_{i} - Y_{ij}^{N})^{T})$$
(5-44)

线性判别分析^[86]是基于欧式空间的,所以针对线性判别分析需要采用同样的机制(欧式距离)来选择最近邻样本,然后得到重新定义的类间散度矩阵。 该思想与 SV-KFD 一致,后者通过计算支持向量获得高维空间最近邻样本。



图 5-2 NN-FLD 的示意图 Fig.5-2 Illustration of NN-FLD

5.5 基于核费舍尔判别分析和直方图特征的人脸识别方法

正如前文所述,核方法在支持向量机和判别分析领域都得到了非常成功的应用。本节将主要利用它的思想来进一步提高本文提出的两种方法识别

性能,即 HGPP 和 HGLXP。首先本节讨论直方图特征和核方法相结合的 一些背景知识,然后提出新的人脸识别方法。

在文献[15,16]中,很多研究者把直方图特征和核学习(支持向量机)进行 了结合,取得了非常好的效果,另外,核 Fisher 判别分析方法是核技巧的 一个经典应用,而且已经证明一些基于直方图特征的相似度函数是正定的, 所以启发我们用它来提高以直方图为特征的人脸识别系统的性能。

对于直方图特征,直方图的 Chi-square 被广泛地用来计算相似度,假定 x 和 y 为直方图特征,则有:

$$S_{GW-chi}(x,y) = \sum_{i=1}^{B} \frac{(x_i - y_i)^2}{x_i + y_i}$$
(5-45)

$$K_{_{GW-chi}}(x,y) = \exp(-r * S_{_{GW-chi}}(x,y))$$
(5-46)

这里 B 是直方图的 Bins 数, x_i, y_i 表示频度, $S_{GW-chi}(x,y)$ 表示 Chisquare 相似度计算公式, 在文献[16]中, 已经证明加权高斯函数 $K_{GW-chi}(x,y)$ 是满足 Mercer 条件的核函数。

基于直方图特征的核 Fisher 判别分析方法实际上是一种通过核学习的 方式利用训练数据库中先验知识的方法,从而提高系统性能的统计学习方 法。式(5-47)中给出了一种形式化的表示方法,子区域 R_i 上的变换矩阵 W_i 是利用加权高斯函数 $K_{GW-cht}(x,y)$ 作为核函数的核 Fisher 判别分析方法计算得 到的, v_i 为在区域 R_i ,i = 1,...,L上获得的判别特征:

$$v_i = \mathbf{w}_i \Phi(x) = \sum_{j=1}^n \alpha_i^{j} k(x_i^{j}, x)$$
 (5-47)

其中, x_i^j 是第 j 个人脸图像在区域 $R_i, i = 1, ..., L$ 上的空间直方图特征, 而 v_i^1, v_i^2 是对应两个人 P_1, P_2 的判别特征, 那么相似度计算公式如 式(5-48)所示:

$$d(P_1, P_2) = \sum_{i=1}^{L} \frac{v_i^1 \cdot v_i^2}{\|v_i^1\| \cdot \|v_i^2\|}$$
(5-48)

从上面的相似度计算公式可以知道,本文提出的方法实际上是多分类 器融合。而且更为重要的是,该方法可以利用图像本身的结构信息,从而更 加有利于提高人脸识别系统的性能。

5.6 实验分析

本节在原始灰度图像上进行对比实验,根据手工标定人眼睛的位置把 原始图像归一化为 64*64 大小的图片。本文利用的是多项式核函数, $k(x,y) = (\frac{x \cdot y}{|x| \cdot |y|} + 1)^r$,其中r是一个常整数。

5.6.1 基于灰度图像性能对比

本节首先利用大型数据库 CAS-PEAL-R1 来验证所提出方法的识别性能, 该数据库正在逐渐被人脸识别研究人员所接受,而且它的规模要比基本 FERET 数据库大得多。CAS-PEAL-R1 训练集合每个人都有 4 张图片,而 FERET 数据库的训练集合中很多人包含的图片只有 2 张,不利于求解支持 向量。有关该数据库的详细信息可以参考[62],在该数据库上进行测试得到 的结果如表 5-1 所示。实验环境为 PIV 3.2G, 2G Ram, KPCA 维数为 550, *k*-subpart = 10。

表 5-1 CAS-PEAL-R1 数据库的识别结果(r=2)

	PCA	Fisherface	KNullspace	GKFD	KDBA	NN_KFLD	SV-KFD
Accessory	37.1	61	61.5	58.7	64	64	64.1
Lighting	8.2	21.8	16.9	17.5	17.1	19.2	19.8
Expression	53.7	71.3	77.5	78.2	80	81	81.2

Table 5-1 Experimental results for the CAS-PEAL-R1 database(r=2)

需要说明的是, PCA 和 Fisherface 方法来源于参考文献[62], 而且与本 文提出的方法采取了同样的预处理方法。从表 5-1 可以看出,本文提出的方 法都好于其它算法,在数据量较大的 CAS-PEAL-R1 数据库上也有一定幅度 的提高,可以让我们相信该方法具有很强的识别能力。在 CAS-PEAL-R1 数 据库上,Fisherface 的识别率有时候高于 GKFD,可能的原因是 GKFD 利用 求伪逆的方法保留一些噪声信息,所以使得识别性能不够稳定。从以上实验 也可以看到 KNullspace 方法取得了相对稳定的性能,并且避免了伪逆运 算,所以也是很实用的判别分析方法。NN_KFLD 的性能接近 SV-KFD,但 是它的训练时间却要远远小于 SV-KFD,SV-KFD 的时间约为 85 分钟,而 NN-KFLD 约为 30 分钟,训练时间对于核学习方法是很重要的,参考文献 [105]知道核方法训练时间和样本数是非线性递增的关系。尽管 SV-KFD 方 法复杂度较高,但是基于灰度图像人脸识别还是很有优势的,在第6章,本 文将实现一个基于 SV-KFD 方法的原型系统。

本文提出的基于支持向量的判别分析方法试图区分不同类中离得最近的样本,该思想最早起源于支持向量机结构风险最小化原理。我们成功把这一思想应用到判别分析领域。性能最好的 SV-KFD 方法的核心部分是用 SVM 的法向量和支持向量集合来构建 DBFM 和类内散度矩阵,然后利用 SVM 本身所具有的零空间特性来进行判别分析。主要贡献包括如下几方面:1)利用法向量定义的 DBFM,根据结构风险最小化理论,在 SVM 的法向量上投影构成一维空间,具有最好的可分性。2)尽管仅仅利用支持向量似 乎减少了样本,但是在实验中发现几乎全部正例集合中的样本都在支持向 量集中,所以可以知道在一对多策略情况下,能保留尽可能多类内散度信息。3)基于支持向量的均值向量的差向量进一步用来构建扩展的决策边界特征矩阵,实验表明它能提高系统的性能。

基于最近临样本空间的核判别分析方法也可以基于 KPCA+FLD 框架来 实现,其实质也是寻找边缘上的样本并进行区分。该方法的好处在于复杂 度较低,但是性能还可以接受。

5.6.2 基于 HGPP 的核学习方法性能测试

本节使用 CAS-PEAL-R1 的 3 个最大测试集合,即饰物(Accessory),表情(Expression),光照(Lighting)来验证本文提出方法的性能。

5.6.2.1 CAS-PEAL-R1 数据库上的性能对比

本节选择的基准对比方法为 Ensemble-KFA 和 HGPP, 其中 Ensemble-KFA 是基于 HGPP 和原始散度矩阵的集成核学习方法。在这个实验中, 图 像块数为 64, 直方图 Bins 数选择为 32。

如图 5-3 所示, K-HGPP 方法基于 NN_KFLD, 取得非常好的识别性
能。尤其在最大测试集合(光照)上,取得了比 HGPP 好得多的识别性能,在 饰物、表情集合上也都所提高,这是因为测试集合和训练集合之间有类似的 变化,而 K-HGPP 能够学习训练集合的分布作为先验知识,从而提高系统 的性能。另外,K-HGPP 也取得了比 Ensemble-KFA 更稳定识别性能,并且 共有约 6,000 个测试样本用来进行对比试验,所以本文后续实验都采用 K-HGPP 方法,KPCA 维数为 550。



图 5-3 CAS-PEAL-R1 数据库上的性能对比

Fig.5-3 Comparison of the performance on the CAS-PEAL-R1 database

下面的实验用来观测子块数和 K-HGPP 方法识别性能之间的相互关系,从图 5-4 可以看到子块数越多,识别性能越高,从而说明空间结构信息 有利于人脸识别。





Fig.5-4 Recognition rates for different numbers of sub-regions(CAS-PEAL-R1)

5.6.2.2 FERET 数据库上的性能对比

本文进一步利用所有基本 FERET 测试数据库来测试 K-HGPP 方法的性能,其中子块的数目为 64 块。从图 5-5 中可以看到,该方法在 FERET 数据库上取得了比其它方法更好的识别性能。

上面的实验验证了 K-HGPP 方法的有效性。同时,本节提出的方法实际 上解决了基于非统计学习方法的空间直方图特征是否也可以训练的问题,这意 味着在某种可控制的情况下,即训练集合和测试集合的样本涉及的变化比较一 致的时候,可以通过利用训练集合中的先验知识来提高系统的识别性能。





Fig.5-5 Comparison of the performance on the FERET database

5.6.3 基于 HGLXP 的核学习方法性能测试

由上一章可以知道,HGLXP 的特征维数比较高,为了解决这一问题,本 节仍然采用均匀量化的方法来降低特征维数,首先验证在 64 个子区域的情况 下不同的直方图 Bins 数对识别率的影响。

图 5-6 所示的 K-HGLXP 是基于 HGLXP 特征和 NN_KFLD 的核学习人 脸识别方法,从实验结果可以看到直方图的 Bins 数对识别率的影响是很小 的,所以后续的实验中主要利用 16 个 Bins 数。这也说明了均匀量化的方法 在保持系统性能没有大幅度下降的情况下,可以极大地降低 HGLXP 的复杂 度。

为了验证所提出方法的有效性,本文在 CAS-PEAL-R1 和基本 FERET

数据库上对 K-HGPP, K-HGLXP, HGLXP 等方法进行对比实验。由上一个 实验可知,子块数对识别率的影响是很大的,所以后续的实验选择 64 个子 块。需要说明的是,仍然采用 CAS-PEAL-R1 数据库的最大几个测试集合, 即表情、饰物、光照测试集,其对比结果如图 5-7 所示。



图 5-6 直方图 Bins 数对识别率的影响(K-HGLXP)

Fig.5-6 Recognition rates for different number of histogram bins(K-HGLXP)





Fig.5-7 Comparison of the performance on the CAS-PEAL-R1 database 基本 FERET 数据库作为一个人脸识别领域最著名的测试平台之一,可以使得我们更加相信该方法具有很好的适用性,而且该数据库提供的训练集



合相对较少,所以适合于验证统计学习方法的泛化能力,在该数据库上的对比结果如图 5-8 所示。

Fig.5-8 Comparison of the performance on the FERET database

从以上两个实验可以看到 K-HGLXP 取得了比原有方法都要好的识别 性能,但是也看到人脸识别仍然是一个巨大的挑战,尤其涉及到年龄阶段以 及光照的剧烈变化等等。需要说明的是通过观察 FERET 的 Dup 数据集可以 看到它的数据涉及到光照、姿态等的变化,所以如果年龄的变化限定在某个 范围,那么光照、姿态问题仍然是实现鲁棒的人脸识别系统的最为关键的问 题。在第6章,本文将实现一个基于 K-HGLXP 方法的原型系统。

5.7 本章小结

由上一章可知,非统计学习方法,即基于空间直方图方法,取得了非常好的识别性能,然而它无法利用训练集合中的先验知识。本章提出了一种基于集成核学习方法来进一步提高 HGPP 和 HGLXP 的性能。集成学习策略在保留空间结构信息方面具有优势,从而更加有利于人脸识别,本节通过大量的实验进一步验证所提出方法的有效性。另外,基于 Gabor 幅值和相位特征的人脸识别方法取得了最好的识别性能,这说明了这两种特征之间具有互补性。

尽管现有系统取得非常好的识别性能,但是还存在一些问题,如复杂

图 5-8 FERET 数据库上的性能对比

度高等,解决这个问题的一个可行的思路是进行特征选择,比如 AdaBoost 方法。本文提出的方法也可以进一步得到提高,比如通过加权的策略。

第6章 面部特征定位和基于人脸识别的原型系统

人脸识别技术发展到今天还没有完全成熟,所以还不能广泛的应用到 实际系统中,但是可以选择一些合适的场合来应用现有的技术。本章基于前 面章节提出的核心技术和关键问题解决方案设计了几个原型系统,实现了面 部特征点定位系统、照片的半自动标注系统、以及基于人脸识别技术的考勤 系统。

6.1 面部特征定位系统

面部特征提取作为人脸识别技术的基础部分,它的相关研究还在起步 阶段,属于最薄弱环节之一。本节基于约束形状模型,构建了一个演示系 统。该系统的主要任务是定位照片中的人脸图像的面部特征,并计算与基准 定位之间的误差。事实上,不仅仅识别过程中需要定位面部特征点,还有许 多其它的应用是基于面部特征精确定位,比如三维恢复、卡通图像的生成等 等。



图 6-1 面部特征定位系统 Fig.6-1 Facial feature extraction system

随着计算机硬件、存储设备、软件处理技术的发展,视频数据在实际 应用中变得越来越重要。相对于静态图片,视频能够提供更多的信息,所以 动态的定位面部特征也是一个非常有价值的课题,本节提供了一个基于视频 的面部特征提取算法。

轟 无标题 - 约束形状模型的测试系统	
文件 (E) 编辑 (E) 查看 (V) 帮助 (E)	
	Function Set 文件 3.09302 消減 3.09302 方差 2.366792 时间 265 mg 结果文件 D:\FeaturePosT27

图 6-2 基于视频的面部特征定位系统



需要说明的是,约束形状模型需要首先进行人脸检测,所以增强人脸 检测能力可以进一步提高系统的性能。尽管人脸检测技术现在已经逐渐的成 熟,好的情况下能够达到 95%的检测率,但是在实际应用系统中,尤其是 室外环境下还需要进一步提高系统的性能。一般来说,通过新采集的数据重 新训练人脸检测分类器可以提高系统的适应性。

6.2 照片的半自动标注系统

在人脸识别技术尚未达到完全成熟阶段的今天,选择一个合适的应用 是一个很有意义的课题。由于 K-HGLXP 在 FERET 和 CAS-PEAL-R1 数据 库上都取得了最好的识别性能,本节主要实现了一个基于 K-HGLXP 人脸识 别方法的照片半自动标注系统。从前面的阐述中,我们发现有训练过程参与 的人脸识别方法虽然有局限性,但是对于人脸识别来说,数据中涉及到的变 化还是可以通过训练集的构建来获得的,所以本文基于 K-HGLXP 实现了一个照片的标注系统,其中训练集合包括 FERET 和 CAS-PEAL-R1 的训练数 据集合。该系统可以进行单张标注,也可自动运行达到标注的目的。实际 上,照片(正面的)的半自动系统可以通过有限的标注完成对照片的标注,而 对人脸图像进行聚类,可以完成对照片的自动标注。

下面首先介绍如何进行单张标注,所谓单张标注就是对一张照片进行 标注。需要说明的是,在系统运行过程中,对于一个人的所有照片,只需对 第一张进行标注即可。

<mark>3</mark> 照片半自动标注系统	
选择路径 聚类 浏览	单张浏览标注
bczhang1.bmp	YisualGaborFeatureExtract 建议第一次标注

图 6-3 单张标注图片

Fig.6-3 Annotating the picture one by one

如图 6-3 所示, 演示系统提示第一张照片需要标注, 这是因为标注的 过程也是一个识别的过程, 所以需要选定一个或多个代表图像。事实上, 代 表集合的选择是人脸识别中的一个比较关键的问题,本文由于是通过一个交 互过程来完成的, 所以可以设定多个代表图像。

如图 6-4 所示,在弹出的对话框中输入照片所对应人的姓名,就完成 了设定代表图像的任务。

第 原并半自动标注系统 选择路径 聚类 浏览	 单张浏览 标注
bczhang1.bmp	Ŷ → 1 計 2 他尝 3 报偿 4 包括 5 包 6 約 7 报 (·) OK

图 6-4 标注第一张图片 Fig.6-4 Annotation of the first picture

后续的照片,系统会自动的标注,如图 6-5 所示。需要说明的是,对 于每一个个体,用户都需要一个交互的过程来完成第一张照片的标定。



图 6-5 自动标注后续图片

Fig.6-5 Automatic annotating other pictures

★ 照片半自动标注系统 选择路径 選業 浏览	单张浏览	
bczhang1	YisunlGaborFeatureExtract 文 承英結果 6t t t 2 li	
lin2 lin3	lind	

该系统还可以批处理进行聚类,以便于对照片进行归类,如图 6-6 所示。

图 6-6 自动标注(聚类)

Fig.6-6 Automatic annotating of the pictures (Clustering)

生活中的照片对于计算机来说是非常难识别的,很多情况下,由于光 照、表情、姿态的变化导致识别问题变得非常困难。但是如果有一个交互的 过程可能会使得问题变得相对简单一些,本节的半自动照片标注系统可以看 作是人脸识别技术的一个很好的应用。

人脸识别系统的一个瓶颈就是泛化能力不强的问题,但是在实际系统 中,设计者不得不面对这一问题。一个可行的方案是,利用最新采集的图片 重新训练分类器,从而提高识别系统的适应能力。这种在线学习系统还没有 被深入地研究,所以这个问题是非常有意义的课题,不仅仅具有实际应用价 值,而且具有理论意义。

6.3 基于人脸识别的考勤系统

人脸识别在人机交互领域具有最好的可接受性,所以成为研究者的研 究热点。本节实现了一个基于人脸识别的考勤系统,需要说明的是,人数规 模比较小,而且可以保留数据作为事后证据的情况下,考勤系统还是可以应 用的,但是还需要专业人员进行维护,因为现有的基于统计学习的人脸识别 系统泛化能力较弱。另外,该系统利用的是基于灰度图像的 SV-KFD 人脸识 别方法。实际运行过程中要确保连续 5 张照片被正确的识别,并满足阈值要 求,并且保留当前的识别图片以便将来发生问题时用来查询以便确认事实。 另外,这些保存下来的测试数据还可以用来训练分类器,从而提高人脸识别 系统的适应性,但是这一过程需要专业人员来完成,所以未来的工作可以集 中在如何通过在线学习的方法来提高现有系统的识别性能上。



图 6-7 基于人脸识别的考勤系统

Fig.6-7 Time Card system based on Face recognition

考勤系统是人脸识别技术的一个非常好的应用,但是由于各种因素的 影响还不能够满足需要,原因之一就是需要专业研究人员进行维护,而本质 原因是因为人脸识别技术的泛化能力还不强,所以当对不同的人群进行识别 时,需要重新进行训练。但是,随着人脸识别技术的发展,我们有理由相信 这些问题能够得到很好地解决,从而使得基于人脸识别技术的考勤系统能够 获得广泛的应用。

相比于现在比较成熟的基于指纹、掌形的考勤系统,人脸识别还在不 断完善之中,但是它潜在的应用前景已经成为共识,所以人脸识别的研究在 很长一段时间内仍然是一个热点课题。

6.4 本章小结

本节对于人脸识别在实际应用中存在的共性问题进行了探讨并提出一 些可能的方案。现有的大多数识别算法都在一定程度上依赖训练集合,也就 是说,多数识别算法的泛化能力并没有理论上的保障,这对于基于经验风险 最小化的分类学习算法而言问题更为突出,即使是对于所谓基于结构风险最 小化的学习算法,其实际系统的泛化能力已经成为一个瓶颈。因此,保障识 别算法性能的最好办法还是要构建一个与"测试集"(实际运行时需要识别的 人脸图像的集合实际上是不可能得到的,但我们姑且认为该集合是存在的) 属性尽量相近的学习集。一些需要考虑的关键点包括:

(1)训练集合与测试集的一致性。需要考虑种族、年龄、性别等特性的一致性。例如,在东方人群的学习集上学习到的识别模型直接应用于西方人就不是一种明智的选择。

(2)训练集合与测试集成像条件应尽量保持一致。这包括摄像设备、 光照条件、表情、姿态条件等。如果"测试集"变化不可预测,那么学习集中 就应该尽可能覆盖可能出现的各种变化条件。一般来说,在不同条件下采集 的数据其差别是很大的,所以当训练集合和测试集合成像条件不同时,统计 学习方法所学到的分布不能够预测测试集合的变化,所以导致系统性能的下 降。

(3)统计学习近几年来发展得很快,其中支持向量机能够定量的解决 泛化能力问题。人脸识别的成功依赖于机器学习算法的不断进步,可以断 言,学习集的规模要尽可能的大,研究者提出一个强分类器能够覆盖所有的 数据分布,可以解决泛化分析问题。一个小的学习集往往会导致大的学习偏 差,甚至不能正确收敛。对于很多可以良好控制的应用而言,例如门禁或考 勤系统,定期或者不定期的根据前一阶段的系统运行记录进行模型更新是提 高识别系统性能的捷径之一。这包括以下几种可以考虑的可能性:

(1)重新训练法。将全部原型图像集及其累计记录的"测试集"加入到 学习集中,重新训练识别模型。

(2)基于增量学习的识别模型更新。重新训练法固然有效,但并不经济,而且需要保留全部历史学习集,增量学习法是一种更为经济有效的方法。目前研究者已经提出了多种对子空间进行增量学习的策略可以借鉴使用,但这方面的研究还需要进一步深入,尤其是对于判别分析模型的增量式

学习方法。增量学习可以在尽可能少的用户交户的情况下完成。理想情况下,可以通过自动在线学习的方式对识别模型进行联机更新,但是目前尚未 找到该问题的成熟的解决方案。

(3)更新原型图像集。更新原型图像集的意义在于可以在一定程度上 解决老化问题,尤其是对于需要长期运行的门禁和考勤系统,无论是短期老 化还是长期老化,都可以通过更新原型图像集来解决。

泛化能力较弱已经成为人脸识别技术发展的一个瓶颈,这一问题的解决 不仅仅依赖于技术的进步,还取决于如何选择一个合适应用,以及如何设计 应用系统。本节初步探讨这个问题,并设计了面部特征定位、照片的半自动 标注系统、以及考勤系统。

结论

人脸识别技术具有重要的理论意义和研究价值,经过几十年的发展已 经获得了长足的进步,但是实践证明它还没有完全成熟。本文重点探讨了实 现鲁棒人脸识别系统的一些核心技术和关键问题,对面部特征精确配准、基 于 Gabor 相位特征的人脸表示和识别方法、非线性判别分析方法进行了专 门研究。

本文提出基于约束形状模型(CSM)的方法解决面部特征精确配准问题。它是实现鲁棒实用的人脸识别系统的基本前提。该方法首先对面部特征 进行分类,对不同类型的特征点采取不同的建模和搜索方法,提出一种约束 形状模型。该模型把面部特征合理地分成边缘点,控制点两部分。对于边缘 点采取改进的 ASM 的建模策略,即考虑了边缘强度信息,而对于控制点采 用 Gabor 特征进行建模,以及基于启发式的搜索策略。实验表明,CSM 能 够更精确的定位面部关键特征。

本文提出一种基于 Gabor 相位信息的特征提取和识别方法—HGPP 方法。该方法不同于传统的人脸识别方法,它主要采用 Gabor 相位特征,并且 首次把 Gabor 象限信息应用到人脸识别领域。同时该方法还基于一种新的局 部异或算子(LXP)。另外一个更重要的特点是无需训练,能够直接提取直方 图特征,因而避免了统计学习的泛化分析问题。

本文提出的局部异或算子(LXP)具有一般性,也就是可以适用于一般的 灰度图像。本文根据不同的梯度方向对人脸图像进行二值化,使得原始的灰 度图像转化成多个二值图像,而 LXP 算子被用来反映邻域点之间的梯度方 向的一致性。最后,LXP 算子成功的和 Gabor 特征进行结合,提取直方图 特征,取得非常好的识别性能。

本文证明了核费舍尔和支持向量机之间的内在联系,即支持向量机的 法向量具有零空间特性。基于该特性,本文提出了基于支持向量的判别分析 方法。该方法实际上是在最近邻样本空间进行判别分析,支持向量即为高维 空间离得最近的样本。

本文提出集成核学习方法提高基于空间直方图特征人脸识别方法的性能。空间直方图特征在识别的时候,可以直接利用最近邻分类器,而且无需训练,其优点是泛化能力较强,但是不能够利用训练集合中的先验知识。由于直方图特征相似度函数及相关函数被证明是正定的,所以启发我们利用核

费舍尔方法提高人脸识别系统的性能。

人脸识别仍然是一个非常有挑战性的工作,根据以往的研究经验,如 下工作需要进一步进行深入研究:

(1)面部特征精确定位仍然是一个最基础的工作,它作为人脸识别的 最关键步骤之一,决定了系统的性能。特征建模是面部特征点定位一个核心 问题,所以下一步工作是利用本文提出的 LXP 特征对若干特征点建模,从 而提高定位精度。

(2) 基于 Gabor 相位信息的特征提取方法还存在特征维数过高的问题,下一步工作是可以进一步利用特征选择的方法选择一些更具判别能力的特征,从而提高系统的效率。对于局部异或算子,还可以研究它的 n-阶算子,以及旋转不变特性等等。另外一个方向就是研究它在其他领域的应用。

(3)关于非线性判别分析一个需要解决的问题是核函数的选择问题, 针对具体的数据,设计一个最好的核函数,是一个非常有价值的研究课题。 另外一个问题是如何提高系统的效率,可行的思路是尽量降低原始输入特征的维数。

最后需要指出的是,人脸识别仍然是非常有挑战性的研究课题。它有 巨大的应用背景,所以也面临着许多实际问题,比如光照,姿态变化等等, 因此,尝试不同的信息采集方式也许是未来解决该问题的最关键一步,比如 红外特征,3D特征等等。

105

参考文献

- 1. R.Chellappa, C.L.Wilson, and S.Sirohey, *Human and machine recognition of faces: A survey*, in Proc. IEEE, 1995, 83(5):705-740
- Wen-Yi Zhao, Rama Chellappa, P.J. Jonathon Phillips, and Azriel Rosenfeld, Face Recognition: A Literature Survey, ACM Computing Survey, 2003, 35 (4):399-458
- 3. J. D. Woodward, *Biometrics: Privacy's Foe or Privacy's Friend*?, in Proc. IEEE,1997, 85,(9):1480-1492
- 4. A. K. Jain, S. Pankanti and R. Bolle (eds.), *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999.
- S. Pankanti, R. M. Bolle and A. K. Jain, *Biometrics: The Future of Identification*, in Special issue of IEEE Computer on Biometrics, 2000, 33(2):46-80
- H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie and T. S. Huang, (eds.), Face Recognition: From Theory to Applications, Springer-Verlag, Berlin, 1998
- 7. S. Gong, S. J. McKenna and A. Psarrou, *Dynamic Vision: From Images to Face Recognition*, Imperial College Press, 2000
- M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz ,and W. Konen, *Distortion invariant object recognition in the dynamic link architecture*, IEEE Trans. Computers, 1993, 42(3):300-311
- L. Wiskott, J.M. Fellous, N. Krüger, and C. von der Malsburg, *Face Recognition by Elastic Bunch Graph Matching*, Technical Report IR-INI 96-08, Institut f
 ür Neuroinformatik, Ruhr Universit
 ät Bochum, D-44780, Bochum, Germany, 1996
- C. Liu and H. Wechsler, Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, IEEE Trans. on Image Processing, 2002, 11(4): 467-476

- S. Shan, P. Yang, X. Chen, W. Gao, *AdaBoost Gabor Fisher Classifier for Face Recognition*, in Proc. of IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2005:278-291
- T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, Active shape models their training and application, Computer Vision and Image understanding, 1995, 61(1):38-59
- S. Mika, G. Ratsch, J. Weston, B. Schölkopf and K. R. Müller, Fisher Discriminant Analysis with Kernels, in Proceedings of IEEE Neural Networks for Signal Processing Workshop, 1999:41-48
- C.J.C. Burges. A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining, 1998, 2(2):121-167
- 15. A. Barla1, F. Odone, A. Verri. *Histogram Intersection kernel for image classification*. in proc. of ICIP, 2003:161-164
- Belongie S., Fowlkes C., Chung F.N, and Malik J.. Spectral partitioning with indefinite kernels using the nystorm extensions. in proc. of ECCV, 2002:531-542
- O. Chapelle, P. Hallner, and V.N. Vapnik. Support vector machines for histogram-based image classification. IEEE Transactions on Neural Networks, 1999, 10(5):1055--1064.
- 18. H. Chan and W.W. Bledsoe, *A man-machine facial recognition system: some preliminary results*. Technical report, Panoramic Research Inc., Cal, 1965.
- 19. A.J.Goldstein, L.D.Harmon, and A.B.Lesk. *Identification of human faces*. Proceedings of the IEEE, 1971, 59(5): 748~760
- 20. T. Kanade. *Picture processing system by computer and recognition of human faces*. Ph.D Dissertation. Kyoto Kyoto University, 1973.
- R. Brunelli and T. Poggio, *Face Recognition: Features versus Templates*. in IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(10):1042-1052
- P.J.Phillips, H.Moon, etc. *The FERET Evaluation Methodology for Face-Recognition Algorithms*. IEEE Transactions on PAMI, 2000, 22(10):1090-1104
- 23. Abdi Valentin, O'Toole, Cottrell, *Connectionist Models of Face Processing: A Survey*. Pattern Recognition, 1994, 27:1209-1230

- 24. A. Tefas, C. Kotropoulos and I. Pitas, Using Support Vector Machines to Enhance the Performance of Elastic Matching for Frontal Face Authentication. IEEE Trans. Pattern Analysis and Machine Intelligence, 2001, 23(7):735-745
- 25. 丁嵘,苏光大,林行刚. 使用关键点信息改进弹性匹配人脸识别算法.电子 学报,2002:1292-1294
- 26. M. Kirby and L. Sirovich, Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces. in IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(1):103-108
- 27. Alex Pentland, Baback Moghaddam, Thad starner. View-Based and Modular Eigen-spaces for Face Recognition MIT Media Lab perceptual Computing Section. IEEE Conference on Computer Vision & Pattern Recognition, 1994:84-91
- P.Belhumeur, J.Hespanha, and D.Kriegman. *Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection.* in Proceedings of Fourth European Conference on Computer Vision, 1996:45~56
- 29. B. Moghaddam and A. Pentland. *Probabilistic Visual Learning for Object representation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 20(7):696-710
- 30. Baback Moghaddam, Tony Jebara, Alex Pentland. *Bayesian Face Recognition*. Pattern Recognition, 2000, 33:1771-1782
- 31. C.Liu. *Statistical and evolutionary approaches for face recognition*. George Mason University, Ph.d Dissertation, 1999.
- 32. W.Zhao, R. Chellappa, and A.Krishnaswamy. Discriminant Analysis of Principal Components for Face Recognition. Proc. of Inter. Conf. On Auto. Face and Gesture Recognition, 1998:336-341
- 33. L.F.Chen, H.M.Liao, J.C.Lin, M.T.Ko, and G.J.Yu. *A New LDA-based Face Recognition System Which Can Solve the Small Sample Size Problem.* Pattern Recognition, 2000, 33(10):1713-1726
- 34. J.Yang, Y.Yu, W.Kunz. *An Efficient LDA Algorithm for Face Recognition*. Proc. Int'l Conf. Control, Automation, Robotics and Vision, 2000:817-829

- 35. H. Yu aand J. Yang. A Direct LDA Algorithm for High-dimensional Data with Application to Face Recognition. Pattern Recognition, 2001, 34:2067-2070
- 36. M.H.Yang, N. Ahuja, D. Kriegman. *Face Recognition Using Kernel Eigenfaces*. In the Conf. on Image Processing, 2000, 1:37-40
- M.H.Yang. Kernel Eigenfaces vs Kernel Fisherfaces: Face Recognition Using Kernel Methods. Proc. Int'l. Conf. Automatic Face and Gesture Recognition, 2002:215-210
- 38. 刘青山. 人脸跟踪与识别的研究. 中科院自动化所博士学位论文, 2003
- 39. A. Hyvärinen. *Survey on Independent Component Analysis*. in Neural Computing Surveys, 1999, 2:94-128
- 40. M.S. Bartlett and T. J. Sejnowski. *Independent Components of Face Images: A Representation for Face Recognition.* in Proc. of the Fourth Annual Joint Symposium on Neural Computation, 1997:3-10
- M.S. Bartlett, H. M. Lades and T. J. Sejnowski. *Independent Component Representations for Face Recognition*. in Proceedings of the SPIE, Conference on Human Vision and Electronic Imaging III,1998, 3299:528-539
- 42. C. Liu. Capitalize on Dimensionality IncreasingTechniques for Improving Face RecognitionGrand Challenge Performance. IEEE Trans. Pattern Analysis and Machine Intelligence, 2005:725-737
- 43. C. Liu. Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition. IEEE Trans. Pattern Analysis and Machine Intelligence, 2004, 26(5): 572-581
- 44. F. S. Samaria. *Face Recognition Using Hidden Markov Model*. Ph.D. Thesis, 1994
- 45. A. Nefian. A Hidden Markov Model-Based Approach for Face Detection and Recognition. Ph.D. Thesis, 1999
- 46. 薛斌党,欧宗瑛. E-HMM/ANN混合网络人脸识别. 计算机辅助设计与图 形学学报, 2002, 14(11):1070~1073
- P.J. Phillips. Support vector machines applied to face recognition. In Advances in Neural Information Processing Systems, pp.803. Editors: M.C.Mozer, M.I.Jordan, and T. Petsche, MIT Press, 1998:803-809

- 48. K.Jonsson, J.Matas, J.Kittler, Y.P.Li. *Learning Support Vectors for Face Verification and Recognition*. Proceeding of the 4th International Conference on Face and Gesture Recognition, Grenoble, France, 2000:208-213
- G.Guo, S.Z.Li and K.Chan. *Face Recognition by Support Vector Machines*. Proc. of the 4th Int. Conf. on Auto. Face and Gesture Recognition, 2000:196-201
- P.Penev and J.Atick. Local Feature Analysis: A General Statistical Theory for Object Representation. Network: Computation in Neural Systems, 1996, 7:477-500
- 51. T.F.Cootes, G.J.Edwards, C.J.Taylor. *Active Appearance Models*. Proc. European Conf. Computer Vision, 1998, 2:484-498
- 52. Peng Yang, Shiguang Shan, Wen Gao, Stan Li, Dong Zhang. Face Recognition UsingAda-Boosted Gabor Features. Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, Korea, 2004:356-361
- 53. Li, S.Z., Lun Zhang, ShengCai Liao, XiangXin Zhu, RuFeng Chu, Meng Ao, Ran He. A Near-infrared Image Based Face Recognition System. Automatic Face and Gesture Recognition, 2006:455 – 460
- 54. B. Moghaddam. *Principal manifolds and probabilistic subspaces for visual recognition*. IEEE Trans. PAMI, 2002 ,24(6):780-788
- 55. Shuicheng Yan, HongJiang Zhang, Yuxiao Hu, Benyu Zhang, QianSheng Cheng. *Discriminant Analysis on Embedded Manifold*. ECCV (1), 2004: 121-132
- 56. W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang. Local Gabor Binary Pattern Histogram Sequence(LGBPHS): a Novel Non-Statistical Model for Face Representation and Recognition. in proc. 10th IEEE International Conference on Computer Vision, 2005: 786-791
- 57. T. Ahonen, A. Hadid, and M. Pietikainen. *Face Recognition with Local Binary pattern*. in *proc*. 8th European Conference on Computer Vision, 2004:469-481
- 58. P.J. Philips, P.J. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, and J.M. Bone. Face Recognition Vendor Test 2002: Evaluation report, Technical Report, NISTIR6965. National Institute of Standards and Technology, 2003. http://www.frvt.org

- 59. P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, William Worek. *Overview of the Face Recognition Grand Challenge*. IEEE Conference on Computer Vision and Pattern Recognition, 2005:947-954
- Terence Sim, Simon Baker, and Maan Bsat. *The CMU Pose, Illumination and Expression Database*. IEEE Transactions on Pattern Analysis and MachineIntelligence, 2003, 25(12):1615 1618
- 61. E. Bailly-Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J.Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz, and J.-P. Thiran. *The BANCAdatabase and evaluation protocol.* in Audio- and Video-Based Biometric PersonAuthentication (AVBPA), 2003:625–638
- 62. Wen Gao, Bo Cao, Shiguang Shan, Delong Zhou, Xiaohua Zhang, Debin Zhao. *TheCAS-PEAL Large-Scale Chinese Face Database and Evaluation Protocols*. TechnicalReport No. JDL_TR_04_FR_001, Joint Research & Development Laboratory, CAS, 2004. http://www.jdl.ac.cn
- 63. A.R. Martinez and R. Benavente. *The AR face database*. Technical Report 24,Computer Vision Center(CVC) Technical Report. Barcelona, Spain, 1998.
- 64. K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In Second International Conference on Audio and Video-basedBiometric Person Authentication,1999:72-77
- 65. M.Turk, A.Pentland. *Eigen-faces for Recognition*. Journal of cognitive neuroscience.1991,3(1):71-86
- 66. 山世光. 人脸识别中若干关键问题的研究. 中国科学院计算技术研究 所,博士学位论文, 2004
- 67. M. Kass, A. Witkin, D. Terzopoulos *Snakes: active contour models*. International Journal of Computer Vision ,1988:321--332
- 68. Cootes TF, Wheeler GV, Walker KN, Taylor CJ. *View-Based active appearance models*. In: Proc. of the 4th IEEE Int'l Conf. on Automatic Face and Gesture Recognition. Washington: IEEE Computer Society,2000:227-238
- 69. X.Hou, S.Z. Li, H. Zhang, Q. Cheng. *Direct appearance models*. In: Proc. of CVPR, 2001:828-833

- 70. Shiguang, Shan, Wen Gao, Wei Wang, Debin Zhao. Enhanced Active Shape Models with Global Texture Constraints for Face Image Analysis. Fourteenth International Symposium On Methodologies For Intelligent Systems, N.Zhong et al., 2003:593-597
- 71. Yi Zhou, Lie Gu, HongJiang Zhang. Bayesian Tangent Shape Model: Estimating Shape and Pose Parameters via Bayesian Inference. CVPR (1),2003:109-118
- 72. D. Gabor. *Theory of communication*. Journal of the Institution of Electrical Engineers, part III, 1946, 93(26): 429--457
- 73. J.G. Daugman. Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters. J. Optical Soc. Amer., 1985, 2(7):160-169
- 74. Jen-Tzung Chien, Chia-Chen Wu. *Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002:1644-1649
- 75. 周杰, 卢春雨, 张长水, 李衍达.人脸自动识别方法综述.电子学报, No.4, 2000
- 76. 张翠平,苏光大. 人脸识别技术综述. 中国图形图像学报, 2000,5:885-894
- 77. 边肇祺,张学工. 模式识别. 北京,清华大学出版社, 2000
- 78. D. L. Swets and J. Weng. Using Discriminant Eigenfeatures for Image Retrieval. in IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8):831-836
- 79. H. Zhang, W. Gao, X. Chen, D. Zhao. Learning Informative Features for Spatial Histogram-Based Object Detection. in Proc. of International Joint Conference on Neural Networks, 2005:1806-1811
- 80. M. J. Swain and D. H. Ballard. Color indexing. International Journal of Computer Vision, 1991, 7(1):11-32
- C. Schmid and R. Mohr. Local Greyvalue Invariants for Image Retrieval. IEEE Trans. Pattern Analysis and Machine Intelligence, 1997, 19(5): 530-534

- M. Stricker and A. Dimai. *Color Indexing with weak spatial constraints*. in proc. SPIE, Storage and Retrieval for Still Image and Video Database IV, 1996, 2670: 29-40
- E. Hadjidemetriou, M.D. Grossberg, S. K. Nayar. *Multiresolution Histograms and Their Use for Recognition*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004, 26(7): 831-847
- 84. S. Ravela and R. Manmatha. *Retrieving images by appearance*. in proc. 6th IEEE International Conference on Computer Vision, 1998:608-613
- 85. 王蕴红,范伟,谭铁牛.融合全局和局部特征的子空间人脸识别算法. 计算机学报,2005:1659-1663
- 86. Jian Yang, Alejandro F. Frangi, Jing-Yu Yang, David Zhang, Zhong Jin: KPCA Plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition. IEEE Trans. Pattern Anal. Mach. Intell.,2005, 27(2): 230-244
- 87. J.G. Daugman. *High confidence visual recognition of persons by a test of statistical independence*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993, 15(11):1148–1161
- D. Zhang, W.K. Kong, J. You, M. Wong. Online Palmprint Identification. IEEE Trans. Pattern Analysis and Machine Intelligence, 2003, 25(9): 1041-1050
- 89. R. Mitchell Parry, I. Essa. *Feature weighting for segmentation*. in proc. 5th international conference on music information retrieval, 2004:116-119
- 90. 山世光,高文,唱轶钲,曹波,陈熙霖. 人脸识别中的"误配准灾难"问题研究. 计算机学报,2005,28(5):782-791
- 91. S. Shan, W. Gao, Y. Chang, B. Cao, Peng Yang. Review the Strength of Gabor Features for Recognition from the Angle of its Robustness to Misalignment. in Proc. of International Conference on Pattern Recognition, 2004:338-341
- 92. W. Gao, B. Cao, S. Shan. The CAS-PEAL Large-Scale Face Database and Evaluation Protocols. Technical Report No. JDL_TR_04_FR_001, JDL, CAS, 2004

- 93. 张晓华,山世光,曹波,高文,周德龙,赵德斌. CAS-PEAL大规模中国人 脸图像数据库及其基本评测介绍. 计算机辅助设计与图形学学报, 2005, 17(1):9-17
- 94. T. Ojala, M. Pietikainen, T. Maenpaa. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971–987
- 95. M. Pietikäinen, T. Ojala, and Z. Xu. *Rotation-invariant texture classification using feature distributions*. Pattern Recognition, 2000, 33:43-52
- 96. 洪子泉, 杨静宇. 基于奇异值特征和统计模型的人像识别算法. 计算机研 究与发展, 1994, 31(3):60~65
- 97. R. A. Fisher. *The Use of Multiple Measurements in Taxonomic Problems*. in Annals of Eugenics, 1936, 7
- Chengjun Liu, Harry Wechsler. A Gabor Feature Classifier for Face Recognition. In proc. of ICCV, 2001: 270-275
- 99. Liu Wei, Wang Yunhong. *Null space based kernel fisher discriminant analysis for face recognition*. In: Proc. 6th International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004:369-374
- 100. J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, Z. Jin. KPCA Plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence,2005:230-244
- 101. Xipeng Qiu, Lide Wu. Stepwise Nearest Neighbor Discriminant Analysis. IJCAI, 2005:829-834
- 102. Xipeng Qiu, Lide Wu. Face Recognition by Stepwise Nonparametric Margin Maximum Criterion. In proc. of ICCV, 2005:1567-1572
- 103. B. Schölkopf, A. Smola and K. Müller. *Nonlinear Component Analysis as a Kernel Eigenvalue Problem*. in Neural Computation, 1998, 10(5):1299-1319
- 104. M. Yang, N. Ahuja and D. Kriegman. Face Recognition Using Kernel Eigenfaces. in Proceedings of the International Conference on Image Processing, 2000:37-40
- 105. K. Baek, B. A. Draper, J. R. Beveridge and K. She. PCA vs.ICA: A Comparison on the FERET Data Set. in Proceedings of International

Conference on Computer Vision, Pattern Recognition and Image Processing in conjunction with the Sixth JCIS, 2001:824-827

- 106. C. Lee and D. A. Landgrebe. Feature Selection Based on Decision Boundaries. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993, 15(1):388-400
- 107. K. Fukunaga. Introduction to Statistical Pattern Recognition. Academic Press, 1990
- 108. Bernhard Scholkopf, Chris Burges. *Extracting Support Data for a Given Task.* KDD, 1995:252-257
- 109. Bernhard Scholkopf. Learning with Kernels. MIT Press, 2002
- 110. Rik Fransens, Jan De Pris. SVM-based Nonparametric Discriminant Analysis, An application to a Face Detection. In proc. of ICCV, 2003:1289-1296

攻读学位期间发表的学术论文

- 1. 张宝昌,陈熙霖,山世光,高文.基于支持向量的 kernel 判别分析方法. 计算机学报 (录用待发表)
- Baochang Zhang, Shiguang Shan, Xilin Chen, Wen Gao. Histogram of Gabor Phase Patterns: A novel object representation for face recognition. IEEE Trans. on Image Processing (accept, SCI impact factor 2.42)
- Baochang Zhang, Xilin Chen, Shiguang Shan, and Wen Gao. Nonlinear face recognition based maximum average margin criterion, IEEE International Conference on Computer Vision and Pattern Recognition, Vol. 2, San Diego, CA, 2005:1036~1041[SCI]
- Baochang Zhang, Shiguang Shan, Wei Wang. Constrainted Shape Model Using Edge Constraint And Gabor Wavelet Based Search. International Conference on Audio-and Video-Based Biometric Person Authentication, 2003, Guildford,UK., 2003:52-61[SCI]
- Baochang Zhang, Gabor-Kernel Fisher analysis for Face Recognition, Proc. Of Pacific Conference on Multimedia, 2004:802-809 [SCI]
- Baochang Zhang, Wen Gao, Shiguan Shan, Peng Yang. Discriminanting Gaborface and Support Vector Machine Classifier for Face Recognition, Asian Conference on Computer Vision, 2004:37-42
- Baochang Zhang, Xilin Chen, Wen Gao, Discriminant Analysis based on Kernelized Decision Boundary for Face Recognition, International Conference on Audio-and Video-Based Biometric Person Authentication, 2005:966-976[SCI]
- Bingpeng Ma, Fei Yang, Wen Gao, Baochang Zhang, The application of extended geodesic distance in head poses estimation, International Conference on Biometric Authentification, 2006:192-198[SCI]
- Yi Li, Baochang Zhang, Shiguang Shan, Xilin Chen, Wen Gao, bagging based efficient kernel fisher analysis for face recognition, International Conference on Pattern Recognition, 2006:523-526
- 专利:

一种基于 Gabor 相位模式的图像识别方法,申请号:200510068027.5

哈尔滨工业大学博士学位论文原创性声明

本人郑重声明:此处所提交的博士学位论文《人脸特征提取和非线性识 别方法的研究》,是本人在导师指导下,在哈尔滨工业大学攻读博士学位期 间独立进行研究工作所取得的成果。据本人所知,论文中除已注明部分外不 包含他人已发表或撰写过的研究成果。对本文的研究工作做出重要贡献的个 人和集体,均已在文中以明确方式注明。本声明的法律结果将完全由本人承 担。

作者签字: 日期: 年月日

哈尔滨工业大学博士学位论文使用授权书

《人脸特征提取和非线性识别方法的研究》系本人在哈尔滨工业大学 攻读博士学位期间在导师指导下完成的博士学位论文。本论文的研究成果归 哈尔滨工业大学所有,本论文的研究内容不得以其它单位的名义发表。本人 完全了解哈尔滨工业大学关于保存、使用学位论文的规定,同意学校保留并 向有关部门送交论文的复印件和电子版本,允许论文被查阅和借阅。本人授 权哈尔滨工业大学,可以采用影印、缩印或其他复制手段保存论文,可以公 布论文的全部或部分内容。

保密□,在 年解密后适用本授权书。

本学位论文属于

不保密□。

作者签名:	日期:	年 月	日
导师签名:	日期:	年 月	日

致谢

值此论文即将完成之际,在此由衷感谢在本人攻读博士期间所有关心、帮助、支持我的老师、同学和家人们。

我首先要感谢我尊敬的导师高文教授。他是一位高瞻远瞩的科学家,在布 满荆棘的道路上始终将正确的路标放在我前进的方向上。正是因为高老师艰苦 的创业才使我有机会在一个国内甚至国际上最好的环境下进行科研工作。高老 师清晰的思路总能使我在纷繁芜杂中找到最主要的问题,然后去解决它。高老 师以他广博的知识、严谨的治学态度、高标准的研究水平、丰富的研究经验随 时影响和教育着我,必将使学生受益终身。

感谢实验室的陈熙霖研究员。陈老师严谨的科学作风、忘我的工作精神、 无私的胸怀一直是我钦佩和学习的榜样。博士期间的工作一直受到他的悉心指 导,在与陈老师的学习和交流中,使我终生受益,让我感受到科研的乐趣。同 时感谢陈老师在百忙之中给我以国际期刊撰写方面的大力帮助。感谢实验室的 赵德彬、姚鸿勋教授。两位老师忘我的工作精神和丰硕的科研成果,一直为我 所钦佩和向往。两位老师给我的无私帮助和热情关心,使我终生难忘。

感谢山世光副研究员。他严谨的工作作风,渊博的知识都使我受益匪浅。 作为课题组组长,总能在关键的时候给我以支持,是我最为感谢的。他在论文 撰写方面的大力帮助,使我学到很多写科技论文的技巧。

感谢王亚东教授对我一直以来的鼓励。感谢凌晓峰教授指点我如何修改国 际期刊论文,使得它顺利的被接受。

感谢王巍、赵三强、崔国勤、焦锋、苏煜、杨非、张洪明、杨澎等给我提 供了很多源代码,为我节约了很多时间。感谢王瑞平、张文超、曹波、陈杰、 苗军、卿来云、柴秀娟、李祎、钟必能、王海晶等非常有意义的讨论,使我学 到很多知识。感谢工程师王乐珩在程序设计方面的帮助。感谢自动化的刘青山 博士给我提供了很多非常好的意见。感谢我的师弟刘心博士生,他在北京那段 时期,是我博士期间最快乐的时光之一。

感谢我的父母,他们无私的爱是我前进的动力,在我最困难的时候,他们 给我的鼓励和信任使我重新收起沮丧的情绪,以饱满的热情投入到科研中。

感谢我的哥哥,他的顽强和责任感是我一辈子都学习的榜样,他为了我的 学业放弃自己的理想,这是我工作的最大动力。希望他的将来更加幸福,此论 文能够让他能够得到一点安慰。

感谢我的叔叔张文生博士在精神和物质上的鼓励和帮助激励我一直坚持我 的学业。

感谢我的妻子林娜娜,在我们相爱的九年来,她给我的是亲情般的感情, 使我始终能够勇于面对一切困难和挑战。在我攻读博士学位期间,她的感受可 能比我还艰难,但是她的坚强深深感染了我,使得我始终以最大的热情去工作。

感谢所有帮助过我的人。