

分类号 TP3

密级

UDC

编号

中国科学院研究生院

博士学位论文

人脸识别中若干关键问题的研究

山世光

指导教师 高文教授

中科院计算所

申请学位级别 工学博士 学科专业名称 计算机应用

论文提交日期 2004.05 论文答辩日期 2004.07

培养单位 中国科学院计算技术研究所

学位授予单位 中国科学院研究生院

答辩委员会主席

声 明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。就我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：

日期：

关于论文使用授权的说明

中国科学院计算技术研究所有权处理、保留送交论文的复印件，允许论文被查阅和借阅；并可以公布论文的全部或部分内容，可以采用影印、缩印或其它复制手段保存该论文。

作者签名：

导师签名：

日期：

摘 要

自动人脸识别 (AFR) 研究试图赋予计算机根据面孔辨别人物身份的能力。该研究具有重要的科学意义和巨大的应用价值。从学科建设与发展的角度看, AFR 作为一个科学问题, 是一个典型的图像模式分析、理解与分类计算问题, 涉及模式识别, 计算机视觉, 智能人机交互, 图形学, 认知科学等多个学科。同时, 作为生物特征识别关键技术之一的 AFR 技术则在公共安全、信息安全、金融等领域具有潜在的应用前景。

经过三十多年的发展, AFR 技术取得了长足的进步, 目前最好的 AFR 系统在理想情况下已经能够取得可以接受的识别性能。但测试和实践经验表明: 非理想条件下的人脸识别技术还远未成熟! 要开发出真正鲁棒、实用的 AFR 应用系统还需要解决大量的关键问题, 尤其需要研究: (1) 作为识别必要前提条件的面部关键特征精确定位问题; (2) 高效的人脸描述特征及其相应的高精度核心识别算法; (3) 如何提高 AFR 系统对不可避免的配准错误的鲁棒性问题。另外, 对开发鲁棒实用的 AFR 系统而言, 研究应用系统设计层面的诸多工程技术问题同样至关重要。以设计开发鲁棒、实用的 AFR 系统为目标, 本文重点探讨了人脸识别中的上述关键问题。本论文的主要贡献总结如下:

1. 全面综述了人脸识别研究的历史和现状

人脸识别研究国内外综述文章最晚也是 2000 年发表的, 实际综述内容大多是 1999 年前的文献情况。本文给出的 AFR 综述首先探讨了 AFR 的一般计算模型, 并按照 AFR 发展特点, 将 AFR 研究划分为三个历史阶段, 对每个阶段研究的特点以及代表性方法进行了较为详尽的总结, 并从面部特征自动定位、人脸表示模型、分类和识别算法三个角度对人脸识别领域的主要方法进行了分类整理。在此基础上, 结合近年来主要的人脸识别评测, 对人脸识别研究的现状进行了阐述。此外, 还总结了现有的主要商业人脸识别系统以及主要的公共人脸图像库的情况。最后讨论了人脸识别领域目前仍然面临的主要开放问题并简单分析了 AFR 领域的主要技术趋势。

2. 研究了特征精确配准问题, 提出了局部纹理约束的主动表观模型 LTC-AAM

面部特征精确配准是鲁棒实用的人脸识别系统的基本前提。主动形状模型 (ASM) 和主动表观模型 (AAM) 是目前解决该问题的主流方法。在对它们的优缺点进行综合分析的基础上, 本文对二者进行集成, 提出了一种局部纹理约束的主动表观模型 LTC-AAM。该模型的要点包括: (1) 继承了 ASM 的局部纹理匹配模型, 并针对面部轮廓点采用了边缘增强的局部纹理匹配模型; (2) 在 ASM 形状调整时引入了 AAM 中度量当前纹理与模型纹理匹配度的目标函数, 从而将 ASM 形状参数调整规范到了与 AAM 统一的参数优化框架中; (3) 保留了 AAM 利用当前纹理预测表观模型参数变化的启发式参数优化过程。这些特点使得 LTC-AAM 融合了 ASM 局部搜索模型、AAM 纹理约束以及启发式的参数优化过程, 从而在较大程度上继承了二者的优势。实验表明, LTC-AAM 能够更精确的定位面部关键特征, 为后续的人脸识别打下了良好的配准基础。

3. 研究了 Gabor 特征人脸描述方法, 提出了一种 AdaGaDA 人脸识别新方法

良好的人脸表示是鲁棒高效的人脸识别算法的关键因素。Gabor 小波因具有优良的空间局部性和方向选择性, 能够提取图像局部区域内多个方向的空间频率和局部结构特征, 是一种良好的人脸描述方法。弹性图匹配方法 (EGM) 和 Gabor 特征 Fisher 判别分

类方法 (GFC) 是应用 Gabor 特征的典型代表。实用系统中 EGM 需要解决关键特征点的定位问题, 其速度和精度难以提高; 而 GFC 对高维 Gabor 特征直接简单下采样的策略尽管避免了特征定位难题, 但却遗漏了大量判别特征。针对这一问题, 本文考虑了如何对 Gabor 特征进行有效降维的问题, 将 Boosting 策略创新性的应用于 Gabor 特征的优化选择以产生维数较低的 AdaGabor 特征, 并最终通过对 AdaGabor 特征的判别分析实现人脸识别——即所谓的 AdaGaDA 人脸识别新方法。在 CAS-PEAL 和 FERET 两个大型人脸图像库上的对比实验表明: AdaGaDA 方法不但可以有效降低 Gabor 特征维数, 而且识别精度也比传统方法有了较大提高。

4. 首次提出并系统研究了 AFR 中的“误配准灾难”问题, 给出了 AFR 算法误配准鲁棒性能的量化评价方法, 提出了一种对误配准进行强化训练的 E-Fisherface 方法, 同时还分析了 AdaGaDA 等基于 Gabor 特征的 AFR 方法的误配准鲁棒性

本文首次系统研究了面部特征配准精度对人脸识别算法性能的影响问题, 发现 Fisherface 的识别率会随着误配准的增大而急剧下降——本文称这一现象为“误配准灾难”(COMA) 问题。在分析产生原因的基础上, 探讨了 COMA 问题的四类解决方案: 精确再配准方法、不变特征方法、误配准学习方法和误配准不变距离度量方法。为了量化比较不同识别算法对误配准的鲁棒性, 本文提出了一套识别算法误配准鲁棒性能量化评估方法。针对 Fisherface 的“误配准灾难”问题, 提出了一种对误配准进行强化训练的 E-Fisherface 方法。还进一步考察了所提出的 AdaGaDA 等基于 Gabor 特征的 AFR 方法对误配准的鲁棒性。这些方法在 FERET 和 CAS-PEAL 人脸图像数据库上的对比实验表明: Gabor 特征比单纯的灰度特征具有更强的鲁棒性; 对误配准进行强化训练的 E-Fisherface 方法可以大大提高识别系统的鲁棒性。

5. 初步探讨了 AFR 应用系统设计中的共性关键问题, 实现了三种应用原型系统

核心识别算法的性能当然是实用人脸识别应用系统最终识别性能的决定性因素, 但系统设计是否合适同样在很大程度上影响着应用系统的成败, 尤其在人脸识别技术尚未完全成熟的今天, 良好的系统设计可能起到事半功倍的效果, 而拙劣的系统设计可能会葬送最好的核心识别算法的前途。鉴于此, 本文初步探讨了人脸识别应用系统设计中需要考虑的一些共性关键问题, 如现场环境设计、摄像设备选择与安装、注册原型图像的挑选、核心算法选择标准、学习集选择、模型在线更新等。最后, 基于本文提出的核心技术和关键问题解决方案, 给出了人脸照片比对系统、模拟的 XP 用户自动登录系统和机场登机旅客身份验证系统三个典型 AFR 应用原型系统的设计方案。

本论文的上述研究内容不但在人脸识别理论层面具有一定的参考价值, 更关键的是本文的研究结果对于设计开发鲁棒、实用的 AFR 系统具有一定的借鉴意义, 所提出的若干关键技术已经获得了实际应用。

关键词: 自动人脸识别 (AFR), 生物特征识别, 面部特征配准, 主动形状模型 (ASM), 主动表观模型 (AAM), Gabor 小波, 主成分分析, Eigenface, 线性判别分析, Fisherface, AdaBoost, 重采样技术, AdaGabor 特征, AdaGaDA 方法, 误配准灾难 (COMA), 误配准鲁棒性

Study on Some Key Issues in Face Recognition

SHAN Shiguang (Computer Application Technology)

Supervised by Prof. GAO Wen

Automatic Face Recognition (AFR) aims at endowing computers with the ability to identify different human beings according to his/her face image. Such a research has both significant theoretic values and wide potential applications. As a scientific issue, AFR is a typical pattern analysis, understanding and classification problem, closely related to many disciplines such as Pattern Recognition, Computer Vision, Intelligent Human-Computer Interaction, Computer Graphics, and Cognitive Psychology etc. Its research achievements would greatly contribute to the development of these disciplines. While as one of the key technologies in Biometrics, AFR techniques are believed having a great deal of potential applications in public security, law enforcement, information security, and financial security.

After more than 30 years' development, AFR has made great progress especially in the past ten years. The state-of-the-art AFR system can perform identification successfully under well-controlled environment. However, evaluation results and practical experience have shown that AFR technologies are currently far from mature. A great number of challenges are to be solved before one can implement a robust practical AFR application. The following key issues are especially pivotal: (1) the accurate facial feature location problem, which is the prerequisite for sequent feature exaction and classification; (2) efficient face representation and corresponding classifier with high accuracy; (3) how to improve the robustness of AFR to inevitable mis-alignment of the facial feature. In addition, elaborate system design is also as important for developing robust and practical AFR systems. In this thesis, the above-mentioned key issues are studied, aiming at robust and practical AFR systems. And the main contribution of this thesis includes:

1. Provided a thorough survey of the AFR history and the state-of-the-art

The latest AFR survey was published in the year 2000, which in fact surveyed the AFR researches before 1999. This thesis has provided a more recent overview of the AFR research and development. The survey starts from the discussion of the general computational model of AFR methods. Then, the AFR researches are partitioned to three historical stages according to the characteristics of technical mainstreams. For each stage, its technical characteristics are summarized and typical milestone methods are mentioned. After that, AFR methods are further categorized according to facial feature extraction, face representation, and classification separately. We also survey the main public face databases and performance evaluations protocols, based on which the state-of-the-arts of AFR are summarized. Main commercial systems are introduced as well with their technical sources. Finally, the challenges and technical trends in AFR fields are discussed.

2. Studied facial feature alignment problem, and proposed a Local Texture Constraint Active Appearance Model.

Accurate facial feature alignment is the prerequisite of an AFR system. Active Shape Model (ASM) and Active Appearance Model (AAM) are the main methods for this problem.

On the basis of analyzing their merits and demerits, LTC-AAM is proposed by combining the ASM and AAM. The main points of LTC-AAM include: (1) Inherit the local texture matching model in ASM and enhanced the local texture model for facial contour landmarks by exploiting explicitly the edge constraints; (2) As in AAM, the energy function measuring the matching degree between the model texture and present texture are acted on active shape searching, therefore the shape parameter adjusting in ASM is unified into the same framework of variable optimization in AAM; (3) Linear prediction of the appearance parameter variance using the present texture as in AAM is retained for the heuristic search in the variable optimization procedure. These points have made LTC-AAM integrate the strengths of ASM's local searching model, AAM's global texture constraints and AAM's heuristic parameter optimization. Therefore, our experiments have illustrated the better performance of LTC-AAM on face alignment.

3. Investigated the Gabor wavelet face representation method, and proposed a AdaGaDA-based AFR approach.

Efficient face representation is the key for a good AFR method. For its eminent characteristics in spatial local feature exaction and orientation selection, Gabor-based face representation has attracted much attention. Elastic Graph Matching (EGM) and Gabor feature Fisher Classification (GFC) are typical approaches using Gabor features. However, EGM needs to accurately localize tens of facial landmarks, which would slow down the recognition speed and accuracy. While the GFC simply down-samples the high-dimensional Gabor features evenly, many discriminating features may be lost, though this has avoided the challenging accurate feature alignment. Against this problem, this thesis innovatively proposes an AdaBoost-based Gabor feature dimension reduction method. Boosting techniques are exploited to select low-dimensional AdaGabor features with much discriminating power. Those selected AdaGabor features are classified by Fisher linear discriminant analysis to achieve final identification, which is named by AdaGaDA face recognition approach. Experiments on FERET and CAS-PEAL face database have illustrated that AdaGaDA has efficiently reduced the dimension of the Gabor representations, and the recognition performance has also greatly improved compared with traditional methods.

4. Initially proposed and investigated the “Curse of mis-alignment” problem; Presented a set of quantitative measurements evaluating AFR method’s robustness to mis-alignment; Proposed E-Fisherface to enhance the robustness of Fisherface to mis-alignment by perturbation learning in the model training stage; The robustness of AdaGaDA and GFC are also analyzed theoretically and evaluated experimentally.

This paper initially studies how the facial feature alignment accuracy influences the face recognition performance. We discover that the performance of Fisherface degrades abruptly with the increase of mis-alignment degree. This paper defines this phenomenon as “Curse of Mis-alignment” (COMA) problem. Based on the analysis of its cause, four categories of solutions to COMA are discussed, that is, alignment-retuning method, mis-alignment robust feature-based method, mis-alignment learning method, and mis-alignment invariant distance measurement method. To quantitatively measure and compare the robustness of different AFR approaches, a set of quantitative measurements evaluating AFR method’s robustness to

mis-alignment are proposed. Afterwards, aiming at the COMA problem of Fisherface, an E-Fisherface method is proposed to enhance the robustness of Fisherface to mis-alignment by perturbation learning in the model training stage. The robustness of AdaGaDA and GFC to mis-alignment is also analyzed theoretically and evaluated experimentally. The experimental evaluations on FERET and CAS-PEAL face database have shown that Gabor feature are more robust to mis-alignment compared with the pure image gray-level intensity feature, and the proposed E-Fisherface can greatly improve the robustness of the Fisherface methods.

5. Primarily studied on some common key issues in AFR system design, and implemented three prototype AFR systems.

The overall performance of AFR systems certainly depends mainly on the kernel AFR algorithm, however, system design strategies would further impact the overall performance of AFR systems. This point is even more reasonable nowadays, given that most present AFR technologies are far from mature. Favorable system design would greatly facilitate the recognition procedure, while a bad one may degrade terribly a best AFR technique. Therefore, some common key issues in AFR system design are briefly discussed, such as the workplace environment design and rebuild, imaging equipment selection and installation, selection of registration prototype images, selection of kernel recognition algorithm, learning set construction, model updating etc. Finally, based on the techniques proposed in this thesis, three prototype AFR systems, facial photo matching system, simulated XP face logon system, and airport passenger verification system, are implemented for demonstration purpose.

The research achievements of this thesis would not only contribute to the AFR computation theory, but also facilitate the development of robust and practical AFR applications. Some of these technologies have been successfully applied to practice.

Keywords : Automatic face recognition (AFR), Biometrics, face alignment, Active Shape Models, Active Appearance Models, Gabor wavelet, Principal Component Analysis (PCA), Eigenface, Linear Discriminant Analysis, Fisherface, AdaBoost, Re-sampling technologies, AdaGabor features, AdaGaDA face recognition method, Curse of mis-alignment, Robustness to mis-alignment

目 录

摘 要.....	I
目 录.....	VII
第一章 绪论.....	1
1.1 研究背景与项目情况.....	1
1.2 自动人脸识别问题描述.....	1
1.3 人脸识别的研究意义与典型应用.....	2
1.4 人脸作为生物特征识别方法的优势与不足.....	3
1.4.1 人脸识别的技术优势.....	4
1.4.2 人脸识别的弱点.....	5
1.5 人类视觉识别系统特性简介及其借鉴意义.....	5
1.6 问题的提出及本文主要贡献.....	8
1.6.1 问题的提出.....	8
1.6.2 本文主要贡献.....	10
1.7 论文的组织结构.....	11
第二章 人脸识别研究综述.....	13
2.1 人脸识别的一般计算模型探讨.....	13
2.2 人脸识别研究历史、现状概述.....	14
2.3 人脸识别的国内研究状况.....	19
2.4 人脸识别主要技术方法归类.....	20
2.5 国内外公用人脸图像数据库及其性能评测情况.....	22
2.5.1 国内外主要公共人脸图像数据库.....	22
2.5.2 FERET 测试.....	23
2.5.3 FRVT2000 和 FRVT2002 测试.....	23
2.6 国内外主要人脸识别商业系统.....	25
2.7 人脸识别领域的若干开放问题和技术发展趋势.....	26
2.7.1 若干开放问题.....	26
2.7.2 可能的技术趋势浅析.....	28
2.8 人脸识别研究发展现状总结.....	28
第三章 局部纹理约束的主动表观模型.....	29
3.1 问题的提出.....	29
3.2 前人工作简介.....	30
3.2.1 主动形状模型.....	30
3.2.2 主动表观模型 (AAM).....	32
3.3 局部纹理约束的主动表观模型 LTC-AAM.....	34
3.3.1 边缘约束局部纹理模型.....	35

3.3.2 LTC-AAM 基本思想	35
3.3.3 LTC-AAM 算法流程	36
3.3.4 主要特点总结	37
3.4 实验与分析	37
3.5 本章小结	40
第四章 基于 AdaGabor 特征判别分析的人脸识别方法.....	41
4.1 问题的提出	41
4.2 前人相关工作介绍	42
4.2.1 Gabor 特征提取与性质	42
4.2.2 AdaBoost 分类方法简介	45
4.3 AdaGaDA 人脸识别方法	46
4.3.1 多类问题的二类化	46
4.3.2 基于 Boosting 技术的 AdaGabor 特征选择	47
4.3.3 基于 AdaGabor 特征的子空间判别分析	49
4.4 实验与分析	50
4.4.1 测试数据库情况介绍	50
4.4.2 AdaGabor 特征的统计特性分析	53
4.4.3 对比算法设计	54
4.4.4 识别性能对比分析	55
4.5 本章小结	56
第五章 人脸识别中的“误配准灾难”问题研究.....	57
5.1 引言	57
5.2 “误配准灾难”问题的提出	58
5.2.1 实验设计——Fisherface 在 FERET 人脸库上的测试	58
5.2.2 “误配准灾难”问题	59
5.3 “误配准灾难”产生原因分析及解决方案归类	61
5.3.1 精确再配准方法	61
5.3.2 误配准不变的鲁棒特征方法	62
5.3.3 误配准学习方法	62
5.3.4 变换不变距离度量方法	62
5.4 考虑误配准鲁棒性的识别性能评价方法	62
5.5 基于配准误差学习的增强判别分析方法 (E-Fisherface)	65
5.6 Gabor 特征的误配准鲁棒性分析	66
5.7 实验评估与分析	67
5.7.1 在 FERET 上的对比实验	67
5.7.2 在 CAS-PEAL 上的对比实验	70
5.8 本章小结	71
第六章 人脸识别应用系统设计.....	73
6.1 问题的提出	73
6.2 应用系统设计共性关键问题初探	73
6.2.1 现场环境设计与改造	73
6.2.2 摄像设备选择与安装	74

6.2.3 增强人脸检测功能.....	74
6.2.4 选择合适的注册原型图像.....	75
6.2.5 核心识别算法的选择.....	75
6.2.6 识别输出策略.....	76
6.2.7 学习集的构建.....	76
6.2.8 识别模型更新与在线学习.....	77
6.2.9 活体判别问题.....	77
6.3 基于面像的身份识别系统.....	77
6.3.1 人脸照片比对系统.....	78
6.3.2 模拟 Windows XP 用户自动登录系统.....	79
6.4 基于面像的身份验证系统.....	81
6.4.1 机场登机旅客身份验证系统.....	82
6.5 本章小结.....	83
第七章 结束语.....	85
7.1 研究的问题及其研究意义.....	85
7.2 本文研究内容及主要贡献.....	85
7.2.1 全面综述了自动人脸识别研究的发展状况.....	86
7.2.2 提出了局部纹理约束的主动表观模型 LTC-AAM.....	86
7.2.3 提出了基于 AdaGaDA 人脸识别新方法.....	86
7.2.4 提出并系统研究了人脸识别中的“误配准灾难”问题.....	87
7.2.5 初步探讨了人脸识别应用系统设计问题.....	87
7.3 后续工作.....	87
参考文献.....	i
致谢.....	xiv
作者简历.....	xvi

第一章 绪论

1.1 研究背景与项目情况

本论文的研究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)，863 计划项目“生物特征识别核心技术与关键问题研究”(合同号：2001AA114190) 的资助。同时还得到了上海银晨智能识别科技有限公司 (IS' VISION) 的大力资助。此外，还曾作为一个子课题同时得到了国家自然科学基金重点课题(项目编号 69789301)、863 项目“多功能感知技术”(合同号：863-306-ZD03-01-2) 以及中科院“百人计划”项目的资助。这些项目的目标是研究人体生物特征识别尤其是人脸识别的核心技术和关键问题，提出新的识别方法，解决一些影响识别率的关键问题。其中核心技术主要包括实时精确的人脸检测与跟踪技术，快速精确的面部特征提取算法，准确快速的人脸识别/确认算法；关键问题主要包括光照问题、姿态问题、表情问题、饰物问题、活体判别问题等。这些项目的共同的产品目标是开发出拥有完全自主知识产权的算法、产品和系统，实现中国人体生物特征识别领域尤其是人脸识别领域的跨越式发展。

1.2 自动人脸识别问题描述

人类似乎具有“与生俱来”的人脸识别能力，赋予计算机同样的能力是人类的梦想之一，这就是所谓的“自动人脸识别”(Automatic Face Recognition, 以下简称 AFR) 研究。如果我们把照相机、摄像机、照片扫描仪等看作计算机的“眼睛”，数字图像可以看作计算机观察到的“影像”，简单地说，AFR 研究试图赋予计算机根据其所“看到”的面部照片来判断人物身份的能力。广义的讲，这种识别能力包括以下不同功能的任务：

1. 人脸检测与跟踪

人脸检测任务要求判断计算机的“眼睛”观察到的“影像”中是否存在人脸。如果存在，需要同时给出其在图像中的坐标位置、人脸区域大小等信息。而人脸跟踪则需要进一步输出所检测到的人脸位置、大小等状态随时间的连续变化情况。

2. 面部特征检测与提取

这一任务既要求确定面部图像中眼睛、鼻子、嘴巴等器官的位置，还要求对这些器官及其面部轮廓的形状信息作出描述。

3. 人脸属性分类

根据面部特征检测与提取的结果，结合面部图像中的亮度分布信息，对检测到的人脸的性别、表情、种族、年龄等属性做出判断。

4. 基于人脸图像比对的身份识别

即通常意义上的人脸识别问题。通过将输入人脸图像与人脸数据库中的所有已知原型人脸图像计算相似度并对其排序来给出输入人脸的身份信息。这包括两类识别问题：一类是无拒识的闭集 (Close set) 人脸识别问题：即假定输入的人脸一定是人脸库中的某个个体；另一类是有拒识的开集 (Open set) 识别，即首先要对输入人脸是否在已知人脸库中做出判断，如果是，则给出其身份。

5. 基于人脸图像比对的身份验证

即通常意义上的人脸确认问题。系统在输入人脸图像的同时输入一个用户宣称的该人脸的身份信息,系统要对该输入人脸图像的身份与宣称的身份是否相符作出判断。

在这几种不同的功能中,尤其需要注意的是识别问题与验证问题的区别。简单地,可以认为识别是一个多类分类问题,而验证则通常是一个二类判别问题[JM00, KT00, TK01, JK02, SA03]。但其中涉及的关键问题是一致的,多数技术方法是通用的。

1.3 人脸识别的研究意义与典型应用

人脸识别的研究始于二十世纪六十年代中后期[CB65, GH71, Ka73, KB76, HH77, HK81],近40年来得到了长足的发展,尤其是近几年来,更成为一个热门的研究课题,国内外各知名大学、研究所、IT公司等都得到了大量的项目支持。人脸识别问题之所以得到重视,是因为其有重要的研究意义,突出表现在其对学科发展的贡献和巨大的潜在应用前景两个方面:

(1) 人脸识别研究可以极大地促进多门相关学科的发展

AFR 作为一个典型的图像模式分析、理解与分类计算问题,它为模式识别,图像处理、分析与理解,计算机视觉,人工智能,人机交互,计算机图形学,认知科学,神经计算,生理学,心理学等多学科提供了一个良好的具体问题,有利于构建这些学科领域的基础实验平台用于尝试新方法,验证新理论,解释新现象。人脸识别问题的深入研究和最终解决,可以极大的促进这些学科的成熟和发展。例如,作为一个模式识别问题,它被认为是最具有挑战性的问题之一:模式种类数目庞大,不同类别模式的差别非常微妙,这一点由于数据采集过程的噪声、成像设备的精度、外界条件的变化以及数据缺损等原因而显得更加突出。再如,作为一个计算机视觉问题,如何融合人脸的一般性先验形状信息来准确地恢复特定人脸的3D结构也是一个非常有价值的研究问题[Ve98, BR02, BT03]。从图形学的角度,研究3D人脸在光源照射下的成像过程,以指导人脸图像的3D形状分析、光照模式分析及其生成虚拟光照、不同视角视图问题也已经成为一个重要的研究课题[Sh97, GB00, BJ01, Ra02]。

人脸识别也是智能人机接口领域的重要研究内容之一。人脸识别本质上是要赋予计算机区分不同人类个体的能力,也就是“看”的能力,这恰是智能人机接口研究需要解决的问题之一,也是机器智能的重要表现。人脸识别问题的最终解决可以极大的改善目前呆板、不方便的人机交互环境,从而在一定程度上改变人们的生活方式。

(2) 作为生物特征识别技术的人脸识别具有巨大的潜在应用前景

身份识别与验证是人类社会日常生活中的基本活动之一。尽管也许是无意识的,我们每天都要对很多人的身份做出判别,同时,每个人也都要经常通过各种方式和手段证明自己的身份,比如在ATM上进行金融业务时需要输入密码,进入自己的计算机系统则需要输入用户名和密码等等。显然,目前我们大多数情况下仍然依赖于传统的身份验证手段来完成这些身份识别过程,这些手段包括身份证卡等各类标识物(如身份证、学生证等各类证件),钥匙,密钥(如密码、口令等),然而这些方式不方便、不安全、不可靠的缺点不言而喻:证件、钥匙携带不便;证件可以被伪造;钥匙可能会丢失;密码

可能会遗忘或被蓄意窃取。这些缺点使得他们已经越来越不能适应社会发展的需求，我们亟需更方便、更可靠、更安全的身份验证方式。同时，国家公共安全、信息安全等关系国计民生的领域更需要高可靠性、高安全性的全新身份验证技术，尤其是美国遭遇“911”恐怖袭击事件之后，安全问题更加成为人们关注的热点问题，而生物特征识别技术被认为是这些需求的终极解决方式。

所谓生物特征识别或生物测定学（**Biometrics**）即利用人类本身所拥有的、能够唯一标识其身份的生理特征或者行为特征进行身份验证的技术。广义的生物特征包括生理特征和行为特征两种。生理特征通常是指人类个体“与生俱来、独一无二、长期不变”（至少在一定程度上满足）的生理特性，即要满足普遍性、独特性和稳定性，如人脸、指纹、虹膜、视网膜、掌纹、手形、DNA、耳廓形状等；行为特征通常是指人类个体的习惯性行为特点，如笔迹、声纹、步态、敲键方式等。这些特征都在一定程度上是“人人拥有、人各不同、长期不变”的，都能反映个体特点，并与个体的身份一一对应，从而可以用来验证个体身份的真伪。

与传统的身份验证方法相比，生物特征识别突出的优势表现在生物特征可以从根本上杜绝伪造和窃取，而且是人类自身拥有的，从而具有更高的可靠性、安全性和可用性。因而，人们对生物特征识别技术寄予厚望，期望能够满足从国家公共安全，社会安全，到金融安全以及人机交互等方面的各类应用。

在国家公共安全方面，保障国家公共安全是社会发展的基础，社会安定更是国家持续发展的必备条件，而境内外违法犯罪人员，尤其是恐怖分子严重威胁人民的生命财产安全，是社会安定团结的极大威胁，对这些危险分子进行及时、有效的监控和抓捕至关重要，而生物特征识别技术，特别是人脸、指纹和声纹识别技术，为此提供了更加有力的技术保障。尤其在911恐怖袭击事件发生后，世界各国政府、安全部门、企业组织等都深刻地体会到了安全方便的身份识别和验证技术的重要性和必要性，对身份验证领域研究的投入迅速增长。我国当然也不例外，尤其是在人员流动日益频繁的今天，在我国这样一个人口众多的大国，利用生物特征识别技术来监控和抓捕危险分子显得尤为重要。而2008年北京奥运会，同样要求采用准确可靠的身份验证技术来防止各种可能的恐怖与破坏活动。在生物特征识别领域形成中国完全自主知识产权的用于保障国家公共安全的身份验证理论方法和技术手段势在必行。

在信息安全领域，信息网络日益普及，这在带给人民群众更多利益的同时，也带来了严峻的信息获取和访问的安全问题。信息安全已引起了全社会的广泛关注，通过信息网络窃取国家机密、商业资料等犯罪现象日益增多，严重影响了社会的公共安全和国家安全。另外，随着电子商务业务的日益普及，人们在享受网络带来的各种不同类型服务（如银行自动取款等）的同时，信息安全也开始与每个人的生活息息相关。生物特征识别技术则可以为此提供一种更加可靠、安全和方便的身份认证手段，在更加方便人们信息获取的同时，保障信息访问的安全合法。

而人脸识别作为一种典型生物特征识别技术，以其自然性、高可接受性等优点受到了人们的青睐，可应用于各行各业中。表1.1汇总了人脸识别的一些典型应用的情况。

1.4 人脸作为生物特征识别方法的优势与不足

如前所述，可供选用的生物特征识别手段有很多种，但他们在身份鉴别方面的安全性、可靠性等性能上是各有千秋的。其中，人脸识别是人类用来互相识别的最重要的方法。

表 1.1 人脸识别应用典型应用
Table1.1 Typical Applications of Face Recognition

应用模式	典型具体应用	特点说明	应用领域
身份识别	出入境管理	过滤敏感人物（间谍、恐怖分子等）	国家安全 公共安全
	嫌疑人照片比对	公安系统用于确定犯罪嫌疑人身份	
	敏感人物智能监控	监控敏感人物（间谍、恐怖分子等）	
	网上追逃	在 PDA 等移动终端上进行现场比对	
	会议代表身份识别	防止非法人员进入会场带来危险因素	
	关键场所视频监控	如银行大厅，预警可能的不安全因素	
	家政服务机器人	能够识别家庭成员的智能机器人	人机交互
	自动系统登陆	自动识别用户身份，提供个性化界面	
	智能 Agent	自动识别用户身份，提供个性化界面	
	真实感虚拟游戏	提供真实感的人物面像，增加交互性	
身份验证	护照、身份证、驾照等各类证件查验	海关、港口、机要部门等查验持证人的身份是否真实、合法	公共安全
	准考证查验	防止替考问题	教育
	机要部门物理门禁	避免钥匙和密码被窃取造成失窃	公共安全
	机要信息系统门禁	避免单纯的密码被窃取造成信息被窃	信息安全
	面像考勤系统	方便，快捷，杜绝代考勤问题	企业应用
	金融用户身份验证	避免单纯的密码被窃取造成财产损失	金融安全
	电子商务身份验证	安全可靠的身份验证手段	金融安全
	智能卡	安全可靠的授权	信息安全
	会议代表身份验证	防止非法人员进入会场带来危险因素	公共安全
	屏幕保护程序	方便快捷的允许合法用户打开屏保	人机交互

1.4.1 人脸识别的技术优势

与其它生物特征识别技术相比，人脸识别在可用性方面具有独到的技术优势，这主要体现在：

- 可以隐蔽操作，尤其适用于安全监控**
这一点特别适用于解决重要的安全问题、罪犯监控与网上抓逃等应用，这是指纹、虹膜、视网膜等其他人体生物特征识别技术不能比拟的；
- 非接触式采集，没有侵犯性，容易被接受**
因此，不会对用户造成生理上的伤害，另外也比较符合一般用户的习惯，容易被大多数的用户接受；
- 具有方便、快捷、强大的事后追踪能力**
基于面像的身份认证系统可以在事件发生的同时记录并保存当事人的面像，从而可以确保系统具有良好的事后追踪能力。例如，用于考勤系统时候，管理人员就可以方便的对代打卡进行事后监控和追踪；这是指纹、虹膜等生物特征所不具有的性质（一般人不具备指纹、虹膜鉴别能力）。
- 图像采集设备成本低**
目前，中低档的 USB CCD/CMOS 摄像头价格已经非常低廉，基本成为标准的

外设，极大的扩展了其应用空间；另外，数码相机、数码摄像机和照片扫描仪等摄像设备在普通家庭的日益普及进一步增加了其可用性；

5. 更符合人类的识别习惯，可交互性强

例如，对于指纹、虹膜等识别系统，一般用户对识别往往是无能为力的，而对人脸来说，授权用户的交互和配合可以大大提高系统的可靠性和可用性；

1.4.2 人脸识别的弱点

然而，人脸作为生物特征识别技术也有其固有的缺陷，这主要表现在：

1. 人脸特征稳定性较差

尽管面部通常不会发生根本性的变化（故意整容除外），但人脸是有极强可塑性的三维柔性皮肤表面，会随着表情、年龄等的变化而发生改变，皮肤的特性也会随着年龄、化妆乃至整容、意外伤害等发生很大变化；

2. 可靠性、安全性较低

尽管不同个体的人脸各不相同，但人类的面孔总体是相似的，而且地球上人口如此众多，以致很多人的面孔之间的差别是非常微妙的，技术上实现安全可靠的认证是有相当难度的；

3. 图像采集受各种外界条件影响很大，因此识别性能偏低

图像的摄制过程决定了人脸图像识别系统必须面对不同的光照条件、视角、距离变化等非常困难的视觉问题，这些成像因素都会极大影响人脸图像的表现，从而使得识别性能不够稳定；

这些缺点使得人脸识别成为一个非常困难的挑战性课题，尤其是在用户不配合、非理想采集条件下的人脸识别问题，更成为目前的热点问题。目前，世界上最好的人脸识别系统也只能在用户比较配合、采集条件比较理想的情况才可以基本满足一般应用的要求[PM00, PG03]。当然，随着技术的进步，相信这些问题也应该可以逐步解决，从而使得 AFR 技术能够更好的满足公众的期望。

1.5 人类视觉识别系统特性简介及其借鉴意义

让计算机具有人类视觉系统一样的自动、快速、准确地识别人脸的能力，是人脸识别研究人员梦寐以求的。人类视觉识别系统自然的成为 AFR 研究的参照系和仿生基础。本节简单的介绍一些文献中，尤其是文献[CW95, ZCR00]中总结的人类视觉识别系统的特性，以期能够对 AFR 的研究有借鉴意义。

1、人脸识别是否是一个特定的过程？[ZCR00]

人类的面像识别机制是否是完全不同于其他一般物体的识别？人类大脑皮层中是否存在专用的人脸识别功能区？这是很多研究人员长期争论不休的焦点问题之一。有些学者认为人脸识别与其他物体的识别没有区别的，使用同样的识别机制，人类大脑皮层中也不存在专门的人脸识别功能区，因此建议 AFR 研究人员应该更强调一般的识别框架的研究，而不应该过多依赖于人脸自身的特定知识。而更多的研究人员则认为人脸识别机制是一个特殊的过程，由专门对应的大脑皮层完成人脸这种特定对象的识别。该论点最有力的证据是“人脸识别能力缺失症（Prosopagnosia）”患者的存在，患有此症的人可以正常的识别其他物体，甚至可以正确的识别鼻子眼睛和嘴巴等面部器官，但是就是不能认出熟悉的人脸，因此有理由怀疑其人脸识别功能区遭到了破坏。因而建议 AFR

研究要更多的利用人脸对象的特定领域知识以便达到更加可靠、鲁棒的识别。实际上，目前的多数 AFR 算法都较多的利用了人脸模式的特定领域先验知识，形成特定的识别机制。

2、全局特征与局部特征孰轻孰重？

全局特征主要包括人脸的肤色特征（比如白皙、黝黑）总体轮廓（比如圆脸、鸭蛋脸、方脸、长脸等）以及面部五官的分布特征（比如，在绘画界就有“国田由用，目甲风申”8种脸形之说），中医也将人脸按照总体结构特征划分为“金木水火土”五行（侧重人脸 3D 结构和肌肉凸凹情况）。而局部特征则主要指面部五官的特点，比如浓眉毛、丹凤眼、鹰勾鼻、大豁嘴、八字胡须、尖下巴等，以及面部的一些奇异特征，比如黑痣、伤痕、酒窝等等。一种被广泛接受的观点认为：二者对识别都是必要的，但全局特征一般用来进行粗略的匹配，局部特征则提供更为精细的确认。但一个必须提及的现象是：如果存在独特的局部特征（比如刀疤、大的黑痣等），则其会首先被用来确定身份。而最支持全局特征比较重要的实验是所谓的“撒切尔夫人倒像实验（Thatcher Illusion）”，如图 1.1 所示。其中，a 和 b 是撒切尔夫人的两幅倒置图像，尽管图像是倒立的，但我们仍然可以比较容易地看出这都是撒切尔夫人的面像，他们看起来非常相似——但实际上，他们的差别是很大的：b 图像中，其眼睛和嘴巴都被倒置了。这可以从 c（a 的倒置）和 d（b 的倒置）中比较容易地看出来。

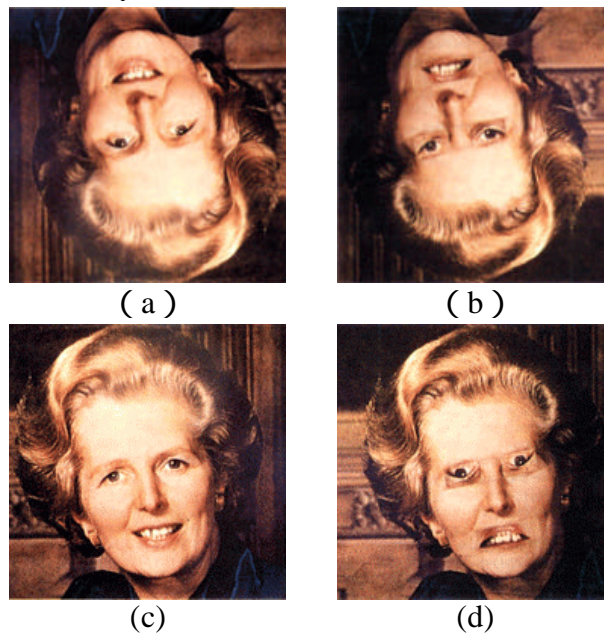


图 1.1 撒切尔夫人倒像实验
Figure 1.1 Thatcher Illusions

3、面部特征对识别的重要性分析[ZCR00]

不同的面部区域对人脸识别的重要性是不同的。一般认为面部轮廓、眼睛和嘴巴等特征对人脸识别是更重要的，人脸的上半区域对识别的意义明显比下半区域重要。发型的变化对人脸识别同样也是重要的，但值得注意的是发型是可变的，尤其是对于年轻人和女性，因此，一般来说发型是不能作为 AFR 的特征使用的。另外，鼻子在侧面人脸识别中的重要性要高于其他特征，这主要是因为侧面人像中鼻子区域包含了若干个关

键的特征点。这些现象对 AFR 的启示是：在分类器的设计中，不同的面部区域应该对最终的识别结果有不同的贡献率。

4、异族人脸识别困难现象

很多东方人都有过这样的体会，在见到足够多的西方人之前，区分不同的西方人往往是比较困难的，这揭示了人脸识别中经常存在的“异族人脸识别困难”现象。这说明了人类的人脸识别能力（至少其中一部分）很可能也是后天学习而来，人类成长过程中学习了足够多的人脸样本，因而具备了“超常”的识别能力，但即使这样，对与学习集中的样本差别较大的异族人脸，仍然不能做到完美的识别。这说明作为先验知识载体的学习集的特性在其中起到了重要的作用。对 AFR 而言，这涉及到识别算法的适应性和泛化能力问题，一方面可能需要尽可能大的学习集，另一方面也需要学习集必须具有较大的覆盖能力，以便最终得到的识别算法不至于只能用于特定类型的人脸的识别。

5、人物面部漫画的启示

人物漫画是我们经常能够看到的，图 1.2 给出了若干明星人物的漫画。不难看出：漫画刻意地夸大了人脸中最为突出的个性化特征（比如陈佩斯的光头和鹰勾鼻子）。这些特征进一步加深了我们对于人物面部特征的认识，从而使我们更容易记住这些人脸。这其实我们 AFR 系统应该强调提取“与众不同”的个性化特征，比如最直接的做法可能就是提取偏离平均脸的那些特性。而基于线性判别分析(Fisherface)的 AFR 方法则试图尽可能提取不同人脸之间的差别特征用于最终的识别。



图 1.2 一些明星人物的漫画示例

Figure 1.2 Example caricatures of some popular stars

6、性别和年龄阶段对于识别性能的影响

FRVT2002（Face Recognition Vendor Test 2002）的测试表明：女性的识别要比男性的识别更困难[PG03]，一般认为这与女性更多的喜欢化妆有关，另外，女性老化速度较快可能也是一个原因。而男性则很少化妆，而且老化的速度也相比女性要缓慢。FRVT2002 还表明年轻人的识别比老年人更难识别[PG03]，这也部分由于年轻人会较多的打扮、改变发型和生理心理变化较多有关，而老年人则相对比较稳定。当然，这些现象还需要认知领域专家的更为权威的解释。如何针对性地提高 AFR 技术对年轻人和女性的识别能力也是需要进一步研究的课题。

7、频域特性与人脸识别的关系[ZCR00]

研究表明：不同空间频段信息对于识别的贡献和作用是不同的。例如，要完成性别

的识别，低频信息往往已经足够，而如果要区分不同人之间的微妙差别，高频分量的作用就更大了。容易理解，低频分量其实更多的是对人脸图像表观总体分布特性的描述，而高频分量则对应局部的细节变化。例如，要想保留某人面部的一颗黑痣的信息，高频分量是无能为力的，必须保留足够的高频分量才可以。这些结论对于构建不同任务的人脸识别系统具有重要的借鉴意义。

8、特异度对人脸识别的影响

研究表明：最“漂亮的”、最“丑陋的”、最“奇异的”的人脸都是最容易被记住的，而大众化的人脸则不太容易被记住。需要指出的是，“大众脸”并不等于“平均脸”，大众脸是指经常可以见到的“脸”，而“平均脸”并不多见，图 1.3 (a) 和 (b) 分别给出平均德国男性脸和平均德国女性脸，可以看出平均脸是相当有魅力的。甚至有研究人员的实验表明：人们通常倾向于认为“平均脸最美”。

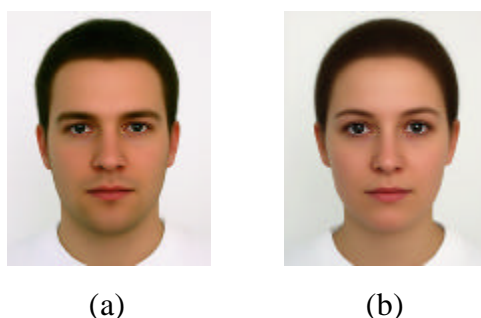


图 1.3 平均人脸示例 (a) 平均男性脸 (b) 平均女性脸

Figure 1.3 Average face (a) Average male face (b) Average female face

9、光照变化与人脸识别

光照变化会大大改变人脸图像的表观，因而影响人脸识别的性能[MA94, AM97]。人们很早就注意到负片人脸难以识别。而最近的研究则进一步表明：对于下方光源照明的人脸也存在识别困难[ZCR00]。这很可能与人类很难见到这类人脸模式因而对他们的学习不足有关。

人类视觉识别系统的上述特性或多或少都对 AFR 研究提供了一定的指导意义，甚至直接影响了很多人脸识别方法的原理和流程。但这方面的研究还需要进一步深入。

1.6 问题的提出及本文主要贡献

1.6.1 问题的提出

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,目前最好的识别系统在这样的情况下识别性能下降非常快,很多情况下识别系统正确识别率陡降至 75% 以下,验证系统等错误率攀升到 10% 以上——这样的性能显然是应用系统用户根本无法接受的!因此,现有的人脸识别系统尚未成熟,要开发出真正鲁棒、实用的 AFR 应用系统还需要解决大量的关键问题,尤其需要研究以下关键问题:

问题1. 作为识别必要前提条件的面部关键特征精确定位问题

面部关键特征精确定位(即面部特征精确配准)是鲁棒实用的人脸识别系统的基本前提。实用的 AFR 应用系统不应该要求用户有过多的交互,而应尽可能的完全自动,一个完全自动的人脸识别系统至少包括人脸检测、面部关键特征定位、人脸描述特征提取、判别分类几个步骤,面部关键特征定位是必不可少的一环。同时,面部关键特征定位的精度直接影响了后续的人脸描述特征提取和分类判别的精确度,可以断言,没有精确的特征配准,就没有高性能的识别。而大量的 AFR 研究文献往往都假定面部关键特征已经被精确定位,在实验中则采用手工标定的眼睛中心、眼角、嘴角等特征位置,这实际上只完成了半自动的人脸识别。因此,面部关键特征定位问题是一个尚未得到足够重视的课题,也是必须继续加大研究力度的一个研究方向。

问题2. 高效的人脸描述特征及其相应的高精度核心识别算法

AFR 算法的精度和鲁棒性不仅取决于采用什么样的分类器,而且在很大程度上取决于采用什么样的特征描述人脸,亦即人脸表示的问题。理论上讲,良好的人脸表示可以使得最简单的分类器具有良好的识别性能。人脸三维形状信息和表面反射属性应该比较好的人脸表示方法,但遗憾的是它们难以从 2D 图像数据中准确获取,因而并不实用。从本质上讲,目前的多数主流的人脸识别方法均直接利用了 2D 图像表现(像素的亮度值)作为人脸表示,其缺点是受成像条件和各种几何变换的影响比较大,难以取得很高的识别精度。Gabor 变换因具有优良的空间局部性和方向选择性,能够提取图像局部区域内多个方向的空间频率和局部结构特征,是一种良好的人脸描述方法。弹性图匹配方法(EGM)和 Gabor 特征 Fisher 判别分类方法(GFC)是应用 Gabor 特征的典型代表。实用系统中 EGM 需要解决关键特征点的定位问题,其速度和精度难以提高;而 GFC 对高维 Gabor 特征直接简单下采样的策略尽管避免了特征定位难题,但却遗漏了大量有效特征。因此,如何高效的利用 Gabor 特征用于人脸识别是一个重要的研究课题。

问题3. 如何提高 AFR 系统对不可避免的配准错误的鲁棒性问题

对实用的人脸识别系统而言,面部特征配准是不可或缺的步骤,现有的多数人脸识别系统都要依赖于面部特征(比如眼睛位置)的严格配准来归一化人脸以便提取人脸描述特征,但面部特征配准的精度如何影响人脸识别算法的性能?在面部特征配准出现错误的时候如何尽可能的保障识别算法的精度不至于下降得太快?如何量化的评估和比较不同算法对配准误差的鲁棒性能?这些关键的问题却没有得到足够的重视,更没有专门的文献研究者方面的内容。本文首次对此进行了系统的专门研究。

此外,对开发鲁棒实用的 AFR 系统而言,研究应用系统设计层面的诸多工程技术问题同样至关重要的。核心识别算法的性能当然是实用 AFR 应用系统最终识别性能的

决定性因素，但系统设计是否合适同样在很大程度上影响着应用系统的成败，尤其在人脸识别技术尚未完全成熟的今天，系统设计的作用更加突出，良好的系统设计可能起到事半功倍的效果。

1.6.2 本文主要贡献

本文重点探讨了实现鲁棒实用的 AFR 系统必须的上述核心技术和关键问题，对面部特征精确配准、基于 Gabor 特征的人脸表示和识别方法、识别算法的误配准鲁棒性问题进行了专门研究。本文的主要贡献总结如下：

1、全面综述了人脸识别研究的历史和现状

人脸识别研究国内外综述文章最晚也是 2000 年发表的，实际综述内容大多是 1999 年前的文献情况。本文给出的 AFR 综述首先探讨了 AFR 的一般计算模型，并按照其发展历史，将 AFR 研究划分为三个阶段，对每个阶段研究的特点以及代表性方法进行了较为详尽的总结，并从面部特征自动定位、人脸表示模型、分类和识别算法三个角度对人脸识别领域的主要方法进行了分类整理。在此基础上，结合近年来主要的人脸识别评测，对人脸识别研究的现状进行了阐述。此外，还总结了现有的主要商业人脸识别系统以及主要的公共人脸图像库的情况。最后讨论了人脸识别领域目前仍然面临的主要技术难题和开放问题。

2、研究了特征精确配准问题，提出了局部纹理约束的主动表观模型 LTC-AAM

面部特征精确配准是鲁棒实用的人脸识别系统的基本前提。主动形状模型 (ASM) 和主动表观模型 (AAM) 是目前解决该问题的主流方法。在对它们的优缺点进行综合分析的基础上，本文对二者进行集成，提出了一种局部纹理约束的主动表观模型 LTC-AAM。该模型的要点包括：(1) 继承了 ASM 的局部纹理匹配模型，并针对面部轮廓点采用了边缘增强的局部纹理匹配模型；(2) 保留了 AAM 中度量当前纹理与模型纹理匹配度的目标函数，从而将 ASM 形状参数调整规范到了与 AAM 统一的参数优化框架中；(3) 采用了 AAM 利用当前纹理预测表观模型参数变化的启发式参数优化过程。这些特点使得 LTC-AAM 融合了 ASM 局部搜索模型、AAM 纹理约束以及启发式的参数优化过程，从而在较大程度上继承了二者的优势。实验表明，LTC-AAM 能够更精确的定位面部关键特征，为后续的人脸识别打下了良好的配准基础。

3、研究了 Gabor 特征人脸描述方法，提出了一种 AdaGaDA 人脸识别新方法

良好的人脸表示是鲁棒高效的人脸识别算法的关键因素。Gabor 具有优良的空间局部性和方向选择性，能够提取图像局部区域内多个方向的空间频率和局部结构特征，是目前主流的人脸描述方法之一。弹性图匹配方法 (EGM) 和 Gabor 特征 Fisher 判别分类方法 (GFC) 是应用 Gabor 特征的典型代表。实用系统中 EGM 需要解决关键特征点的定位问题，其速度和精度难以提高；而 GFC 对高维 Gabor 特征直接简单下采样的策略尽管避免了特征定位难题，但却遗漏了大量判别特征。针对这一问题，本文考虑了如何对 Gabor 特征进行有效降维的问题，将 Boosting 策略创新性的应用于 Gabor 特征的优化选择以产生维数较低的 AdaGabor 特征，并最终通过对 AdaGabor 特征的判别分析实现人脸识别——即所谓的 AdaGaDA 人脸识别新方法。在 CAS-PEAL 和 FERET 两个大型人脸图像库上的对比实验表明 AdaGaDA 方法不但可以有效降低 Gabor 特征维数，而且

识别精度也比传统方法有了较大提高。

- 4、首次提出并系统研究了 AFR 中的“误配准灾难”问题，给出了 AFR 算法误配准鲁棒性能的量化评价方法，提出了一种对误配准进行强化训练的 E-Fisherface 方法，同时还分析了 AdaGaDA 等基于 Gabor 特征的 AFR 方法的误配准鲁棒性

本文首次系统研究了面部特征配准精度对人脸识别算法性能的影响问题，发现 Fisherface 的识别率会随着误配准的增大而急剧下降——本文称这一现象为“误配准灾难”(COMA)问题。在分析产生原因的基础上，探讨了 COMA 问题的四类解决方案：精确再配准方法、不变特征方法、误配准学习方法和误配准不变距离度量方法。为了量化比较不同识别算法对误配准的鲁棒性，本文提出了一套识别算法误配准鲁棒性能量化评估方法。针对 Fisherface 的“误配准灾难”问题，我们提出了一种对误配准进行强化训练的 E-Fisherface 方法。作为补充还进一步考察所提出的 AdaGaDA 等基于 Gabor 特征的 AFR 方法对误配准的鲁棒性。这些方法在 FERET 和 CAS-PEAL 人脸图像数据库上的对比实验表明：Gabor 特征比单纯的灰度特征具有更强的鲁棒性；对误配准进行强化训练的 E-Fisherface 方法可以大大提高识别系统的鲁棒性。

- 5、初步探讨了 AFR 应用系统设计中的共性关键问题，实现了三种应用原型系统

核心识别算法的性能当然是实用人脸识别应用系统最终识别性能的决定性因素，但系统设计是否合适同样在很大程度上影响着应用系统的成败，尤其在人脸识别技术尚未完全成熟的今天，系统设计的作用更加突出，良好的系统设计可能起到事半功倍的效果，而拙劣的系统设计可能会葬送最好的核心识别算法的前途。鉴于此，本文初步探讨了人脸识别应用系统设计中需要考虑的一些共性关键问题，如现场环境设计、摄像设备选择与安装、注册原型图像的挑选、核心算法选择标准、学习集选择、模型在线更新等。最后，基于本文提出的核心技术和关键问题解决方案，给出了人脸照片比对系统、模拟的 XP 用户自动登录系统和机场登机旅客身份验证系统三个典型 AFR 应用原型系统的设计方案。

1.7 论文的组织结构

本论文各章的组织结构如下：

第一章为绪论部分。主要介绍了本论文的研究背景，介绍了人脸识别的问题描述、研究意义，还对人类视觉识别系统的一些特性进行了简单介绍和分析。最后给出了本文研究的主要问题及本文的主要贡献。

第二章为专门的综述部分。从计算模型、发展历史、方法归类、商业系统等不同的层面对国内外的人脸识别研究状况进行了总结，并指出了人脸识别研究中的开放问题。

第三章介绍了本文在面部特征配准方面提出的一种局部纹理约束的主动表观模型。

第四章探讨了基于 Gabor 特征的人脸识别方法，介绍了本文提出的 AdaGaDA 人脸识别方法，与传统方法进行了对比实验分析。

第五章提出并讨论了人脸识别中的“误配准灾难”问题，分析了其产生原因，对其解决方案进行了归类，并提出了本文的解决方案。

第六章对 AFR 应用系统设计的共性关键问题进行了初步探讨，介绍了本文实现的三个典型应用原型系统的设计和实现情况

第七章在对全文进行总结的基础上，讨论了本论文工作可能的后续扩展。

第二章 人脸识别研究综述

人脸识别研究国内外综述文章最晚也是 2000 年发表的，考虑到期刊论文发表的滞后期，所以它们实际综述的内容都是 1999 年前的文献情况，本章的综述则囊括了近年来的最新研究结果。本章首先初步探讨了自动人脸识别的一般计算模型，并按照人脸识别算法的发展历程，将人脸识别研究划分为三个阶段，对每个阶段的研究特点以及代表性方法进行了较为详尽的介绍，在此基础上从面部特征自动定位、人脸表示模型、分类和识别算法三个角度对人脸识别领域的主要技术方法进行了分类整理。然后结合近年来的主要人脸识别评测情况，对人脸识别的研究现状进行了阐述。此外，本文还总结了现有的主要人脸识别商业系统的情况，概述了其特点和技术基础。最后，本文讨论了人脸识别领域目前仍然面临的主要技术难题和开放问题。

2.1 人脸识别的一般计算模型探讨

通常我们所说的人脸识别是基于光学人脸图像的身份识别与验证的简称。光学人脸图像（以下简称人脸图像）是外界光源（包括太阳光、室内人造光源和其他物体表面反射而来的光线）发出的光线照射在人脸上，经人脸表面反射后传播到摄像机传感器的光线强度的度量。不难理解，这一成像过程实际上涉及到三大类关键要素：

1. 人脸内部属性
包括人脸表面的反射属性（如反射系数等，通常简称为纹理[Texture]）、人脸 3D 形状（表面法向量方向）以及人脸表情、胡须等属性的变化；
2. 外部成像条件
包括光源（位置和强度等）、其他物体（比如眼镜、帽子）或者人体其他部件（比如头发）对人脸的遮挡等；
3. 摄像机成像参数
包括摄像机位置（视点）、摄像机的焦距、光圈、快门速度等内外部参数

因此，光学人脸图像的成像过程可以简单地形式化为：

$$I = f(F; L; C), \quad (2-1)$$

其中，函数 f 表示成像函数， F 、 L 和 C 分别表示人脸内部属性、外部成像条件和摄像机成像参数这三类要素， I 为生成的人脸图像。如果我们进一步假设：人脸皮肤的反射属性满足朗博（Lambertian）模型，人脸为凸表面结构，光源为无穷远处的单色点光源，上述成像公式可以进一步改写为：

$$I(x, y) = f_c(\mathbf{r}_{(x,y,z)} \cdot \vec{n}_{(x,y,z)} \cdot \vec{s}_{(x,y,z)}) , \quad (2-2)$$

其中， (x, y, z) 表示人脸表面的一点 P 的三维坐标； $\mathbf{r}_{(x,y,z)}$ 表示 P 点的表面反射率； $\vec{n}_{(x,y,z)}$

表示 P 点的表面法向量方向； $\vec{s}_{(x,y,z)}$ 表示光源的方向和强度； f_c 则表示摄像机的成像函数；而 $I(x, y)$ 则为摄像机最终输出的对应 P 点的图像像素的强度。

不难理解,在上述成像过程中,人脸表面 3D 结构及其反射属性才是人脸相对稳定的本质属性,是人脸识别算法应该赖以区分不同人脸的主要特征;而人脸表情变化、有无胡须等尽管也属于人脸内部属性,但因为具有可变性而不能作为人脸的区分特征。光源等外部成像条件以及摄像机参数等外部因素就更不能作为人脸识别依赖的特征属性。

综上所述,理想情况下,要根据人脸图像区分不同的人脸,根本上似乎需要从人脸图像表观中分离出人脸稳定不变的本质属性(3D 形状与表面反射率)与外界条件和摄像参数。然后,从 3D 形状与表面反射率属性中提取不同人脸的差异信息,馈入到后端的判别分类器中进行识别。设输入图像为 I ,这一过程可以形式化为以下三个步骤:

- 1、属性分离。分离人脸本质属性要素与光源条件 s^* 、摄像参数 c^* 等外部参数要素:

$$(\mathbf{r}_{(x,y,z)}^*, \vec{n}_{(x,y,z)}^*, s^*, c^*) = f_c^{-1}(I(x, y)); \quad (2-3)$$

- 2、特征提取。从人脸属性要素中提取能够体现特定人脸身份的特征:

$$F^* = T(\mathbf{r}_{(x,y,z)}^*, \vec{n}_{(x,y,z)}^*) , \quad (2-4)$$

其中, T 表示特征提取过程。

- 3、分类判别。将提取的特征与数据库中存储的已知人脸特征进行对比,选择相似度最大的人脸作为输入人脸的身份信息:

$$ID^* = \arg \max_{i \in \Omega} (Sim(F^*, F_i)) , \quad (2-5)$$

其中, $Sim(.)$ 表示计算特征之间的相似度, Ω 表示已知人脸集合。

然而,这一过程并不那么简单直接:从单一未知光源条件的图像中恢复 3D 形状信息和表面反射率是经典的视觉难题,本质上是一个病态的问题[ZT99],尽管近年来研究人员通过利用各种约束条件和先验知识可以在一定程度上实现该功能[AG96, HJ01, WH01, ZC00, BV99, BR02, BT03],或者要求多幅不同光照条件下的图像[YS97, GK98, GB01, Sh97]以便得到更精确的估计结果,但直到现在仍没有取得本质的突破。除了少数采用特殊设备获取人脸 3D 结构外,多数系统不得不退而求其次,采用的人脸建模方法仍然停留在图像表观层面上,并没有对 3D 形状和纹理进行显式分离的步骤,而是直接从图像表观中提取判别特征并进行分类来完成识别,此时公式(2-4)相应地改写为:

$$F^* = T(I) . \quad (2-6)$$

例如,在基于面部结构几何特征的人脸识别方法中, T 通常是一个计算面部主要器官(眉毛、眼睛、鼻子、嘴巴和下巴)的大小、形状、位置和角度关系等几何度量参数的过程,最终形成的特征 F^* 是一个反映这些几何度量的特征向量。对于模板匹配方法而言, T 则是对人脸图像进行几何归一化和亮度校正的过程,得到的 F^* 则是一个反映图像各像素点亮度的二维数据矩阵[BP92, BP93]。再比如对“特征脸”方法[TP91]而言, T 则是一个将高维空间中的人脸图像降维到一个低维子空间中的变换(主成分分析),形成的特征 F^* 是降维后的主成分特征分量。

需要指出的是,近年来 Blanz&Vetter 等人提出的 3D 变形模型方法[BV99, BR02, BT03]是上述理论计算模型的重要尝试,已经引起了研究人员的高度重视,尽管还存在很多问题,比如速度和恢复精度问题,但相信此方面的工作近期会有更大的突破。

2.2 人脸识别研究历史、现状概述

人脸识别的研究历史是比较悠久的, Galton 早在 1888 年和 1910 年就分别在《Nature》

杂志发表了两篇关于利用人脸进行身份识别的文章，对人类自身的人脸识别能力进行了分析，但当时还不可能涉及到人脸的自动识别问题。最早的 AFR 的研究论文见于 1965 年 Chan & Bledsoe 在 Panoramic Research Inc. 发表的技术报告[CB65]，到现在已有三十多年的历史。近年来，人脸识别研究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。1999 年前的人脸识别研究内容可参见下面几篇相关综述文章：

- [SI92] Samal & Iyengar, “Automatic Recognition and Analysis of Human Faces and Facial Expressions,” Pattern Recognition, vol. 25, 1992
- [VA94] Valentin, Abdi, O’Toole & Cottrell, “Connectionist Models of Face Processing: A Survey,” Pattern Recognition, vol. 27, 1994
- [CW95] Chellappa, Wilson & Sirohey, “Human and Machine Recognition of Faces: A Survey,” Proc. IEEE, vol. 83, 1995.
- [Gr00] Grudin, “On Internal Representations in Face Recognition Systems,” Pattern Recognition, vol. 33, 2000
- [ZCR00] Zhao, Chellappa, Rosenfeld & Phillips, “Face Recognition: A Literature Survey”, UMD CS-TR-4167, 2000
- [周 00] 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
- [张 00] 张翠平，苏光大，人脸识别技术综述，中国图形图像学报，Vol.5, 2000

需要指出的是，人脸识别是一个被广泛研究着的热门问题，大量的研究论文层出不穷，在一定程度上有泛滥成“灾”之嫌。为了更好的对人脸识别研究的历史和现状进行介绍，本文将 AFR 的研究历史按照研究内容、技术方法等方面的特点大体划分为三个时间阶段，如表 2.1 所示，该表格概括了人脸识别研究的发展简史及其每个历史阶段代表性的研究工作及其技术特点。下面对三个阶段的研究进展情况作简单介绍：

第一阶段（1964 年~1990 年）

这一阶段人脸识别通常只是作为一个一般性的模式识别问题被研究，所采用的主要技术方案是基于人脸几何结构特征（Geometric feature based）的方法[CB65, GH71, Ka73]，这集中体现在人们对于剪影（Profile）的研究上，人们对面部剪影曲线的结构特征提取与分析方面进行了大量研究[KB76, HH77, HK81]。人工神经网络也一度曾经被研究人员用于人脸识别问题中。较早从事 AFR 研究的研究人员除了 Bledsoe 外还有 Goldstein[GH71]，Harmon[HH77, HK81]以及 Kanade[Ka73]等。Kanade 博士于 1973 年在京都大学完成了第一篇 AFR 方面的博士论文[Ka73]，直到现在，作为卡内基-梅隆大学（CMU）机器人研究院的一名教授，Kanade 博士仍然是人脸识别领域的活跃人物之一，他所在的研究组也是人脸识别领域的一支重要力量[RB98, GM02]。总体而言，这一阶段是人脸识别研究的初级阶段，非常重要的成果不是很多，也基本没有获得实际应用。

第二阶段（1991 年~1997 年）

这一阶段尽管时间相对短暂，但却是人脸识别研究的高潮期，可谓硕果累累：不但诞生了若干代表性的人脸识别算法，美国军方还组织了著名的 FERET 人脸识别算法测试[PM00]，并出现了若干商业化运作的人脸识别系统，比如最为著名的 Visionics（现为

Identix) 的 FaceIt 系统[PA96]。

表 2.1 人脸识别发展历史简表

Table 2.1 A Brief History of Face Recognition Research

发展阶段	1964-1990	1991-1997	1998-Current
主要特征	作为一般识别问题被研究 ;基于特征的方法是主流	重点解决的是较理想条件下、用户配合、中小规模人脸数据库上的人脸识别问题 ;基于 Appearance 的 2D 人脸图像线性子空间分析和统计模式识别方法是主流 ;	重点研究非理想条件下、用户不配合、大规模人脸数据库上的人脸识别问题 ;基于 3D 模型和非线性建模的方法等可能是发展趋势
代表性的人脸识别技术与方法及其关键性事件和作品	已知的最早的 AFR 研究论文 [CB65]	Eigenface[TP91]	光照锥技术 [GK98, GB01]
	第一个半自动人脸识别系统 [Ka73]	基于特征的方法与基于模板的方法的对比 [BP93]	支持向量机 SVM 用于人脸识别中 [Ph98, GL00, JM00, DC01, JK02]
		美国 DARPA 启动 FERET 测试项目 [PM00]	3D 可变形模型 [BV99, RB02, BR02, BT03]
	第一篇 AFR 方面的博士论文 [Ka73]	局部特征分析 (LFA) 人脸识别方法发展成为 Visionics 公司 FaceIt 商业系统 [PA96]	基于 AdaBoost 的人脸检测技术 [VJ01, VJo01]
	基于剪影分析的 (Profile) 人脸识别 [KB76, HH77, HK81]	基于双子空间的贝叶斯概率学习 [MP97]	流形学习 ISOMAP [TS00] 以及 LLE [RL00] 等
		人脸识别研究综述 [CW1995]	朗博反射与线性子空间分析 [BJ01, BH01, LH01, Ra02]
	人脸的低维表示 [SK87, KS90]	Fisherface [BH96, BH97]	基于商图像的人脸识别方法 [SR01]
		弹性图匹配技术 [BL90, LV93, WF97]	人脸检测综述 [YK02]
		ASM/AAM [LT94, LT95, LT97, CE98, EC99]	Face Recognition Vendor Test 2000, FRVT 2002 测试 [PG03]
		FERET'1996 测试 [PM00]	
技术特点	基于特征的方法		基于模板的方法
	---	基于神经网络的识别方法	---
	---	统计学习理论	
	---	基于 Appearance 的 2D 人脸子空间分析与建模	
	---	非线性流形分析技术	
	基于 2D 图像模型的人脸识别		基于 3D 模型的人脸识别

麻省理工学院媒体实验室的 Turk 和 Pentland 提出的 “Eigenface” 方法 [TP91, PM94] 无疑是这一时期内最负盛名的人脸识别方法。其后的很多人脸识别技术都或多或少与

Eigenface 有关系[BH97, MP97, MP00, ZC98, Liu99], 现在 Eigenface 已经与归一化的协相关量(Normalized Correlation)方法[BP93]一道成为人脸识别的性能测试基准算法[PM00]。

这一时期的另一个重要工作是麻省理工学院 AI 实验室的 Brunelli 和 Poggio 于 1992 年左右做的一个对比实验, 他们对比了基于结构特征的方法与基于模板匹配的方法的识别性能, 并给出了一个比较确定的结论: 模板匹配的方法优于基于特征的方法[BP93]! 这一导向性的结论与 Eigenface 共同作用, 基本中止了纯粹的基于结构特征的人脸识别方法研究, 并在很大程度上促进了基于表观 (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[BH96, BH97]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)对图像表观特征进行降维, 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[CL00], 子空间判别模型[ZC98, ZC00], 增强判别模型[LW01, LW02], 直接的 LDA 判别方法[YY01], 以及近期的一些基于核学习的改进策略[YA00, Ya02, LH02, 刘 03]。

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[MP95, MP97, MP00, Mo02, Liu99, Liu03, WT04]。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。

人脸识别中的另一种重要方法——弹性图匹配技术(Elastic Graph Matching, EGM)也是在这一阶段提出的[BL90, LV93, WF97]。其基本思想是用一个属性图来描述人脸: 属性图的顶点代表面部关键特征点, 其属性为相应特征点处的多分辨率、多方向局部特征——Gabor 变换特征, 称为 *Jet*; 边的属性则为不同特征点之间的几何关系。对任意输入人脸图像, 弹性图匹配通过一种优化搜索策略来定位预先定义的若干面部关键特征点, 同时提取它们的 *Jet* 特征, 得到输入图像的属性图。最后通过计算其与已知人脸属性图的相似度来完成识别过程。该方法的优点是既保留了面部的全局结构特征, 也对人脸的关键局部特征进行了建模。近来还出现了一些对该方法的扩展[KT00, TK01, 丁 02]。

局部特征分析技术(Local Feature Analysis, LFA)是由洛克菲勒大学(Rockefeller University)的 Atick 等人提出的[PA96]。LFA 在本质上是一种基于统计的低维对象描述方法, 与只能提取全局特征而且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, 从而具有更佳描述和判别能力。LFA 技术已商业化为著名的 FaceIt 系统, 因此后期没有发表新的学术进展。

由美国国防部反毒品技术发展计划办公室资助的 FERET(FacE REcognition technology Test)项目无疑是该阶段内的一个至关重要的事件[PM00]。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 AFR 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(有关测试的细节参见本章 2.5.2 小节)。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。

柔性模型(Flexible Models)[LT94, LT95, LT97]——包括主动形状模型(Active Shape

Models, ASMs)[CT95]和主动表观模型(Active Appearance Models, AAMs)[CE98, EC99, CW00, St00]是这一时期内在人脸建模方面的一个重要贡献。ASMs/AAMs 将人脸描述为 2D 形状和纹理两个分离的部分,分别用统计的方法进行建模(PCA),然后再进一步通过 PCA 将二者融合起来对人脸进行统计建模。柔性模型具有良好的人脸合成能力,可以采用基于合成的图像分析技术来对人脸图像进行特征提取与建模。柔性模型目前已被广泛用于人脸特征对准(Face Alignment)和识别中[LW01],并出现了很多的改进模型[HL01, LY02, YL03, WA03, ZG03, SGW03]。

总体而言,这一阶段的人脸识别技术发展非常迅速,所提出的算法在较理想图像采集条件、用户配合、中小规模正面人脸数据库上达到了非常好的性能,也因此出现了若干知名的人脸识别商业公司。从技术方案上看,2D 人脸图像线性子空间判别分析、统计表观模型、统计模式识别方法是这一阶段内的主流技术。

第三阶段(1998 年~现在)

FERET'96 人脸识别算法评估表明:主流的人脸识别技术对光照、姿态等由于非理想采集条件或者用户不配合造成的变化鲁棒性比较差[PM00]。因此,光照、姿态问题逐渐成为研究热点。与此同时,人脸识别的商业系统进一步发展,为此,美国军方在 FERET 测试的基础上分别于 2000 年和 2002 年组织了两次商业系统评测[PG03]。

Georghiades 等人提出的基于光照锥(Illumination Cones)模型的多姿态、多光照条件人脸识别方法是这一时期的重要成果之一[GK98, GB01]。Georghiades 等人证明了一个重要结论:同一人脸在同一视角、不同光照条件下的所有图像在图像空间中形成一个凸锥——即光照锥。为了能够从少量未知光照条件的人脸图像中计算光照锥,他们还对传统的光度立体视觉方法进行了扩展,能够在 Lambertian 模型、凸表面和远点光源假设条件下,根据未知光照条件的 7 幅同一视点图像恢复物体的 3D 形状和表面点的表面反射系数(传统光度立体视觉能够根据给定的 3 幅已知光照条件的图像恢复物体表面的法向量方向),从而可以容易的合成该视角下任意光照条件的图像,完成光照锥的计算。识别则通过计算输入图像到每个光照锥的距离来完成。

以支持向量机(Support Vector Machines, SVMs)为代表的统计学习理论也在这一时期内被应用到了人脸识别与确认中来[Ph98, GL00, JM00]。SVMs 是一个两类分类器,而人脸识别则是一个多类问题,通常有三种策略解决这个问题,即:类内差/类间差法[Ph98]、一对多法(one-to-rest)[JM00]和一对一法(one-to-one)[GL00]。

Blanz 和 Vetter 等提出的基于 3D 变形(3D Morphable Model)模型的多姿态、多光照条件人脸图像分析与识别方法是这一阶段内一项开创性的工作[BV99, RB02, BR02, BT03]。该方法在本质上属于基于合成的分析技术,其主要贡献在于它在 3D 形状和纹理统计变形模型(类似于 2D 时候的 AAMs)的基础上,同时还采用图形学模拟的方法对图像采集过程的透视投影和光照模型参数进行了建模,从而可以使得人脸形状和纹理等人脸内部属性与摄像机配置、光照情况等外部参数完全分开,更加有利于人脸图像的分析与识别。Blanz 的实验表明,该方法在 CMU-PIE(多姿态、光照和表情)人脸库和 FERET 多姿态人脸库上都达到了相当高的识别率,证明了该方法的有效性。

2001 年的国际计算机视觉大会(ICCv)上,康柏研究院的研究员 Viola 和 Jones 展示了他们一个基于简单矩形特征和 AdaBoost 的实时人脸检测系统,在 CIF 格式上检测准正面人脸的速度达到了每秒 15 帧以上[VJ01, VJo01]。该方法的主要贡献包括:1、用可以快速计算的简单矩形特征作为人脸图像特征;2、基于 AdaBoost 将大量弱分类器进行组合形成强分类器的学习方法;3、采用了级联(Cascade)技术提高检测速度。目

前,基于这种人脸/非人脸学习的策略已经能够实现准实时的多姿态人脸检测与跟踪[ZZ02,LZ04]。这为后端的人脸识别提供了良好的基础。

Shashua 等于 2001 年提出了一种基于商图像的人脸图像识别与绘制技术[SR01]。该技术是一种基于特定对象类图像集合学习的绘制技术,能够根据训练集合中的少量不同光照的图像,合成任意输入人脸图像在各种光照条件下的合成图像,基于此,Shashua 等还给出了对各种光照条件不变的人脸签名(Signature)图像的定义,可以用于光照不变的人脸识别中去,并通过实验表明了该方法的有效性。

Basri 和 Jacobs 则利用球面谐波(Spherical Harmonics)表示光照、用卷积过程描述朗博反射的方法解析地证明了一个重要的结论:由任意远点光源获得的所有朗博反射函数的集合形成一个线性子空间[BJ01]。这意味着一个凸的朗博表面物体在各种光照条件下的图像集合可以用一个低维的线性子空间来近似,这不仅与先前的光照统计建模方法的经验实验结果相吻合,更进一步从理论上促进了线性子空间对象识别方法的发展。而且,这使得用凸优化方法来强制光照函数非负成为可能,为光照问题的解决提供了重要思路[BH01,LH01,Ra02]。

FERET 项目之后,涌现了若干人脸识别商业系统,美国国防部有关部门进一步组织了针对人脸识别商业系统的评测——Face Recognition Vendor Test(FRVT),至今已经举办了两次:FRVT2000 以及 FRVT2002 测试(有关测试的细节参见本章 11 小节)[PG03]。这两次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 三个商业产品遥遥领先于其他系统,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还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[PG03]。

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。

2.3 人脸识别的国内研究状况

在国内,九十年代中后期以来,在 NSFC、863 等资助下,国内众多研究机构的研究组开始对人脸识别进行研究,主要研究组包括*:清华大学计算机系徐光佑教授[游 96]和艾海舟教授[梁 01, 梁 02, SA03]、自动化系边肇祺教授、张长水教授[彭 97, 卢 99, 卢 98, 彭 98]、周杰教授[周 00]、电子系苏光大教授[丁 02, 张 00]、丁晓青教授[杨 03]等;哈工大计算机系和中科院计算所高文教授、陈熙霖教授研究组[刘 97, 陈 99, 山 01, SG03, SC04];中科院自动化所模式识别国家重点实验室的谭铁牛博士、王蕴红博士[王 00, TT03]、马颂德研究员、卢汉青研究员[刘 97, 刘 03]、王阳生博士[WL04];北京交通大学袁保宗教授[WY01];南京理工大学杨静宇教授[Ho91, 洪 92, 洪 94, 黄 98, 荆 98, 李 00];中山大学冯国灿教授、赖剑煌教授[赖 99, 赖 99];南京大学周志华博士[HZ00, 周 01];北京工业大学沈兰荪教授[刑 00]、尹宝才教授[MY99];中科院声学所

*考虑到篇幅的关系,这些研究组的著作列表没有完全列出

杜利民研究员[张徐 00]；上海交大施鹏飞教授；北京大学查红彬教授、封举富教授、吴玺宏教授；浙江大学吴朝晖教授；西北大学周明全教授；西安交通大学谈正教授；复旦大学张立明教授等知名专家学者领导的研究组。这些研究组都在人脸检测、特征提取与识别方面进行了许多有意义的尝试，积累了宝贵的经验。

另外，微软亚洲研究院的李子青博士领导的研究组[GL00, GZ01, HL01, LL99, LY02, LZ04, WL04, YL03, ZZ02]也开展了大量人脸识别方面的工作，取得了很大的进展。尤其值得关注的是李子青博士近期还联合 Anil K.Jain 教授共同组织该领域内的著名专家联合编著了一本人脸识别手册[LJ04]，非常全面的介绍了人脸识别的最新研究进展，必将极大推动该领域的研究进展。

2.4 人脸识别主要技术方法归类

一个实用的 AFR 系统至少包括人脸检测、面部关键特征点自动标定、人脸特征提取（人脸表示）和分类判别几个步骤。对一个模式识别系统而言，采用什么样的特征来表示模式，如何提取这种表示特征，往往是系统成功的关键。对 AFR 而言，面部关键特征点的自动标定又是人脸表达特征的必须环节。因此，本节用表格的形式，从面部关键特征点的自动定位方法（表 2.2）、人脸表示方法（表 2.3）、判别与分类方法（表 2.4）三个角度分别枚举了人脸识别领域的主要技术方法，并给出了相应的参考文献和简要说明。需要说明的是，人脸检测已经逐渐成长为一个相对独立的研究课题，已有很多专门的研究论文和综述[YK02]，因此本文没有过多涉及。

表 2.2 面部关键特征点的自动标定主要方法
Table 2.2 Main approaches for face alignment

方法	简要说明及主要参考文献
模板匹配	通过计算与预定五官模板的相关量来定位五官的位置[BP92]
峰谷分析、积分投影分析技术	用于定位眼睛中心、嘴角和鼻孔等[Ka73, BP92]
Snake	用于提取人脸轮廓特征[KW88, XS94]
特征脸/特征子脸	用于定位人脸及其面部五官的位置[TP91]
可变形模板匹配	调整抛物线、双曲线等数学模型的参数使之最佳匹配眼睛、嘴巴、下巴等处的轮廓特征[Yu91, YH92, XS94, CL93]
光流, <i>Vectorize</i>	采用光流迭代的方法建立输入人脸与参考人脸之间的稠密的像素级对应关系[Be95]
弹性图匹配技术	通过属性图的形变，一方面匹配顶点处的 Gabor 局部特征，另一方面匹配全局几何结构特征[LV93, WF97, KT00, TK01, 丁 02]
主动形状模型 ASM	全局形状统计模型约束局部纹理匹配的结果[CT95, LW01, HL01, LY02, YL03, ZG03, SGW03, WA03]
主动表观模型 AAM	对形状和纹理联合统计(表观)模型参数进行优化使得模型与输入最佳匹配，得到形状和纹理[CE98, EC99, CW00, St00]
3D 形变模型	对 3D 形状、纹理、成像参数等形成的形变模型参数进行优化，使得合成的模型图像最佳匹配输入图像，从而得到人脸 3D 形状和纹理[BV99, RB02, BR02, BT03]

表 2.3 主要的人脸表示方法

Table 2.3 Main approaches for face representation

人脸描述模型	简要说明及主要参考文献
面部几何结构特征	距离、角度、面积等参数[Ka73, KB76, HK81, BP92, BP93, MC92]
灰度模板	所有像素亮度值串接形成的向量或矩阵[BP92]
主成分特征	对图像进行 PCA 降维(去无关)后的特征[SK87, KS90, TP91, PM94, BH97, MP97, ZC98]
奇异值等代数特征	对图像进行 SVD 分解得到的奇异值[Ho91, LC93, 黄 98, 洪 94, TT03, 王 00, 洪 92]
可变形模板	抛物线、双曲线等数学模型的参数[Yu91, YH92, XS94, CL93]
2D 形状模型	稀疏特征点统计分布模型(PDM)[CT95, LT94, , LW01]
光流形状模型	与参考图像的稠密的、像素级的对应关系[Be95]
形状无关纹理模型	将稀疏特征点与平均形状对齐后的图像模型[LT97, LW01]
表观模型	形状与纹理通过 PCA 进一步建模[LT97, LW01]
3D 形变模型	3D 形状 PCA、纹理 PCA 二者的结合[BV99, RB02, BR02, BT03]
局部特征	能够保留全局拓扑关系的局部结构信息[PA96]
弹性属性图	关键特征点的局部结构信息及其几何关系[LV93, WF97, KT00, TK01]
频域特征, 如 Gabor, , FFT DCT	不同频段、不同类型的滤波器在图像上进行卷积得到的频域特征[LY99, LY01, CW02, 赖阮 99, 赖 99, LW02, LW03, Lee96, Kr00]
独立元特征	各特征之间具有统计独立性[BL98, DC01, LW99, LW03]
光照锥	凸朗博表面在任何远点光源照射下同一视角光照图像分布的几何表示[GK98, GB01]
商图像	表面反射率之比[SR01, WL04]

表 2.4 判别与分类方法

Table 2.4 Main approaches for discrimination and classification

判别与分类方法	简要说明及主要参考文献
最近邻/k-近邻	最常用的方法, 文献不能枚举
人工神经网络	包括自组织映射, BP, DLA 等 [LO95, PL92, RB98, LG97, LV93]
最近特征线方法	新的距离度量[LL99]
线性子空间法	每个类别分别建立子空间[BH97, GB01, BJ01, BH01, LH01, SG03]
线性判别分析	例如: [BH96, BH97, CL00, ZC98, ZC00, LW01, LW02, YY01]
核判别分析技术	例如: [YA00, Ya02, LH02, 刘 03]
支持向量机	例如: [Ph98, GL00, JM00]
贝叶斯决策	类内差、类间差双因子空间, 通过贝叶斯分类[MP95, MP97, MP00, Mo02, Liu99, Liu03, WT04]
弹性图匹配技术	例如: [BL90, LV93, WF97, KT00, TK01, 丁 02]
隐马尔可夫模型方法	将结构信息看作时序信号变化[Sa93, SY94, Sa94, Ne99]
AdaBoost	弱分类器加权组合为强分类器[VJ01, VJo01, ZZ02, LZ04, PS04]
流形学习技术	例如: ISOMAP, LLE, LPP 等[TS00, RL00, 张 03]

2.5 国内外公用人脸图像数据库及其性能评测情况

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。而性能评测则是不同算法性能比较、特定算法改进和模型调整、发现问题进而提出新算法的重要环节。本节主要介绍国内外主要公共人脸图像数据库的情况，以及 FERET、FRVT 等人脸识别性能评测的情况[PM00, PG03]，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。

2.5.1 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、算法测试比较是不可或缺的。尤其是在统计学习占主导地位的今天，模型训练所采用的人脸库的人数规模、覆盖的成像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的人脸库的规模和属性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。

✧ FERET 人脸数据库[PM00]

由 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，并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.html

✧ PIE 人脸数据库 [SB03]

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

✧ BANCA 人脸库 [BB03]

该人脸库是欧洲 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：<http://www.ee.surrey.uk/research/vssp/banca/>

✧ CAS-PEAL 人脸数据库 [GC04]

该人脸库包含了 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。该人脸库的介绍网址为：<http://www.jdl.ac.cn/peal/index.html>

✧ AR 人脸库 [MB98]

该人脸库是由西班牙巴塞罗那计算机视觉中心于 1998 年创建的，包括 116 人不同光照、表情、遮挡和老化的人脸图像共 3288 幅。该人脸库的介绍网址为：http://rvll.ecn.purdue.edu/~aleix/aleix_face_DB.html

✧ ORL 人脸数据库[SH94]

由剑桥大学 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。

✧ XM2VTS 多模态生物特征数据库 [MM99]

该数据库包括 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。

✧ MIT 人脸数据库[TP91]

由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。

✧ Yale 人脸数据库[BH97]

由耶鲁大学计算视觉与控制中心创建，包含 15 位志愿者的 165 张图片，包含光照、表情和姿态的变化。

✧ Yale 人脸数据库 B[GB01]

包含了 10 个人的 5,850 幅多姿态、多光照的图像。其中的姿态和光照变化的图像都是在严格控制条件下采集的，主要用于光照和姿态问题的建模与分析。

✧ PF01 人脸数据库[DG01]

由韩国浦项科技大学创建，包含 103 人的 1,751 张不同光照、姿态、表情面部图像，志愿者以韩国人为主。

✧ KFDB 人脸图像库 [HB03]

该人脸库包含 1000 名韩国人的共 52,000 幅图像，涵盖 7 种不同姿态、16 种不同光照条件和 5 种表情变化。

2.5.2 FERET 测试

基于对人脸识别技术在军事、安全和法律等领域重要性的考虑，美国国防部的反毒品技术开发计划通过 DARPA 资助了一个对现有人脸识别技术进行评测的项目，从 1993 年到 1997 年，共组织了三次人脸识别/确认的性能评测(分别在 1994,1995,1996 年进行，1997 年有一次补充测试)[PM00]，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。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%。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。

2.5.3 FRVT2000 和 FRVT2002 测试

FERET 测试之后，很快出现了很多人脸识别商业系统提供商，另一方面，美国安全、军事和法律等部门也开始更加关注人脸识别技术的应用。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 年和 2002 年联合资助了两次专门对商业人脸识别产品进行评测

的项目。

共有 8 个商业人脸识别系统参加了 FRVT2000，但只有 5 家在规定时间内完成了全部测试，它们是 Visionics、Lau、Miros、C-VIS 和 Banque-Tec。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，其性能下降的就越大。

而 FRVT2002[PG03]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[PG03]，分别是：

- 1、身份认证。这类任务的评测指标有两个：错误接受率和错误拒绝率。FRVT2002 并没有绘制 ROC 曲线，而是采取了限定一个错误率，比较另外一个错误率的方法。
- 2、闭集识别。闭集识别不要求算法判断输入人脸是否在原型图像库（Gallery）中，而只需要给出识别结果。FRVT2002 延续了 FERET 的做法，统计前 n 候选累计正确识别率来比较不同算法的性能。
- 3、开集识别。FRVT2002 称之为 Watch List 测试，要求算法首先判断输入人脸是否在 Watch list 集合中，如果是，给出识别结果。FRVT2002 的测试指标是给定一个虚警率，比较不同算法的检测识别率。

FRVT2002 的主要测试结果包括[PG03]：

- ✧ 对目前最好的人脸识别系统，在 37,437 人签证照测试库上，闭集识别任务的最高首选识别率为 73%；前 10 候选累计识别率 82%；前 50 候选识别率 87%；
- ✧ 对目前最好的人脸识别系统，对签证照质量的人脸图像，人脸验证错误接收率为 0.01% 时，最低错误拒绝率 30% 左右；错误接受率为 0.1% 时，最低错误拒绝率 18% 左右；错误接受率为 1% 时，最低错误拒绝率 10% 左右；
- ✧ 在虚警率控制在 1% 时，最好的识别系统在 25 人规模的 Watch List 上的正确检测识别率为 77%，而在 3000 人的 Watch list 上，正确率则下降为 56%；因此，对 Watch list 应用，如果可能，要尽量减少 Watch list 中人脸的数量；
- ✧ 非理想条件测试，虚警率为 1% 时，在 787 人的数据库上，Gallery 中图像均为正面中性表情，室内白炽灯光源条件下采集，对室外同一天测试图像最高首选识别率 54%；室外 152-505 天之间的测试图像最高首选识别率 46%；因此，室外光照条件下的人脸识别问题是一个亟待解决的研究课题；
- ✧ 对姿态问题，在 87 人的人脸库上，Gallery 和测试图像光照条件一致时，左右深度旋转 45 度时最高首选识别率 42%；低头/抬头 30 度时最高首选识别率 53%。而经过 3D 形变模型校正到正面后，分别可以提高到 84% 和 85%。说明 3D 形变模型校正有非常好的效果；
- ✧ 对大约 3 年后的照片，在错误接受率为 1% 时，最高首选识别率 60% 左右，而最低

的错误拒绝率则为 15% 左右。总体而言，时间跨度增加一年，最好识别系统的识别性能大约下降 5 个百分点；

- ✧ 数据库的规模每增加一倍，最好的商业识别系统的首选识别率大约会下降 2 到 3 个百分点；
- ✧ 对视频序列图像的识别效果并不比对静态图像的识别效果好，这一点需要进一步的实验验证；
- ✧ 不同属性的人群识别性能会有差别，例如男性比女性更易于识别（大约有 6 到 9 个百分点的差别）；年轻人比老年人难识别。

2.6 国内外主要人脸识别商业系统

人脸识别经过 40 年左右的发展，技术上已经达到了一定的成熟度，因此从二十世纪九十年代中期开始，陆续出现了一些人脸识别商业系统，这些商业系统基本都依托了知名研究机构多年的研究成果。表 2.5 给出了典型的人脸识别商业系统的情况，还对它们所采用的技术作了简要说明。

表 2.5 主要商业人脸识别系统简表

Table 2.5 Brief introductions of the main commercial face recognition systems

系统名称	提供商	简要说明
FaceVACS-SDK	Cognitec	其算法采用了多阶高斯导数滤波特征，并对这些特征进行判别分析用于最终的匹配识别。其 CEO Alfredo Herrera 最初供职于 Siemens Nixdorf Advanced Technologies GmbH，后转到 Plettac Electronics Security GmbH，2002 年创建 Cognitec。
FaceIt	Identix (其前身为 Visionics)	基于 Rockefeller 大学开发的 LFA 算法，Identix 的现任 CEO 和 CTO 分别为 LFA 的提出者 Atick 博士和 Gruffin 博士[AG96]
PersonSpotter 等	Neven Vision (前身为 Eyematic)	基本算法是基于 Gabor 特征的人脸识别方法。其前身为 Eyematic，2003.7 被 Neven Vision 收购，其现任 CEO Dr. H.Neven 为 USC-ISI 职员。
ZN face II 等*	ZN Bochum GmbH*	基本算法来自美国 USC 和德国 Bochum 大学联合开发的基于 Gabor 特征的弹性图匹配人脸识别方法
Viisage Face TOOLS	Viisage*	基于 MIT 的特征脸技术，后来开发了 ICA 技术，2004 年 1 月开始又融合了 ZN 公司的弹性图匹配技术
Hunter	LAU	基于 MIT 的特征脸技术
FaceSnap	C-VIS	基于弹性图匹配技术
Unmask(+)	VisionSphere	基于 McGill 大学的基于几何结构特征的识别方法
TrueFace	eTrue	基于 Multiple Neural Networks
BioID	BioID	基于人脸、语音和唇动三者的融合进行身份认证
GodEye	IS' Vision	基于中科院计算所高文研究组研发的自适应多特征融合判别分析方法
未知	海青华远	很可能为弹性图匹配技术

*注：2004 年 1 月，ZN 已经与 Viisage 合并为一，公司名仍为 Viisage

2.7 人脸识别领域的若干开放问题和技术发展趋势

2.7.1 若干开放问题

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，目前最好的识别系统在这样的情况下识别性能下降非常快，很多情况下识别系统正确识别率陡降至 75% 以下，验证系统等错误率攀升到 10% 以上——这样的性能显然是应用系统用户根本无法接受的！因此，现有的人脸识别系统尚未成熟，尤其需要针对性的解决在非理想摄像条件下（光照变化、背景变化、摄像设备差异）和用户不配合（视角变化、表情变化、饰物乃至化妆）时必然遇到识别性能下降问题。图 2.1 给出了影响人脸识别性能的内、外部变化情况。

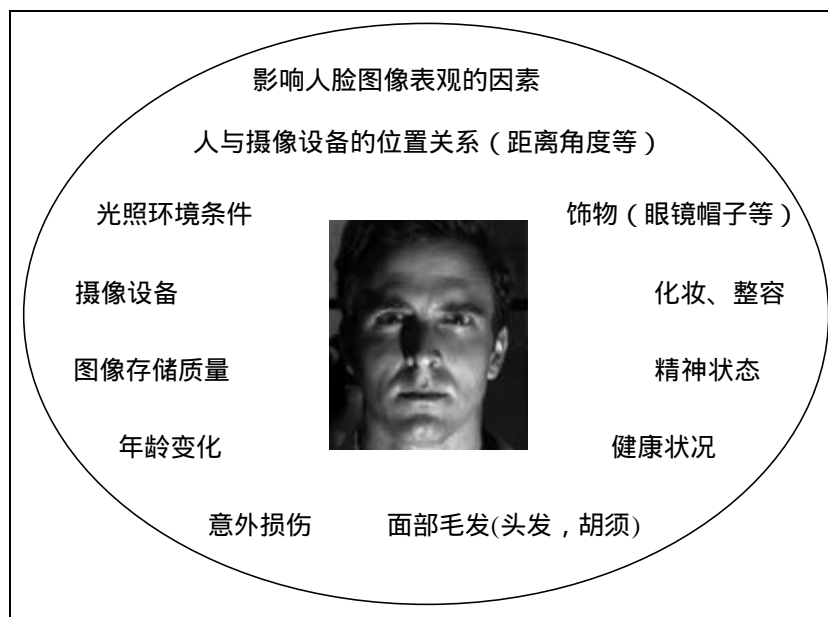


图 2.1 影响人脸图像表观的内外因素

Fig.2.1 Factors affect the appearance of face images

图 2.1 所示的这些变化因素在不同的应用系统中均会有不同程度的出现，因而会极大的影响实用识别系统的性能，导致识别系统性能的下降。概括而言，目前人脸识别领域的主要开放问题包括：

（1）鲁棒、准确的特征定位问题

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。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。

（2）对各种变换鲁棒的核心识别算法

外部条件变化、特征定位的精度都会极大的改变图像的表现，进而影响特征提取和识别的精度。寻求对不改变输入模式类别属性的变化（特征配准误差造成的相似变换）足够鲁棒的特征或者识别算法也是至关重要的问题。

(3) 识别算法的泛化能力和自适应学习问题

统计和学习是目前最成功的人脸识别算法的主流方法，因此数据量和数据的覆盖情况往往是成功的关键，遗憾的是已知的数据往往不能覆盖不同应用中的实际数据的分布情况，因而出现严重的有偏性，亦即泛化能力差的问题。可以考虑两种不同的解决思路，其一是研究如何提高学习算法的泛化能力，其二是研究对新的环境或者数据在线自适应的方法。

(4) 光照变化问题，尤其是室外光照变化。

大量的实用系统是需要室外 24 小时全天候工作的，摄像时的光照条件必然有剧烈的变化。经典的评测也表明：即使是最好的商业系统，其性能也会在室外光照环境下急剧下降，尽管研究人员已经提出了部分解决方案，但多数基于实验室条件下的点光源环境和简单的 Lambert 模型假设，因而并不实用。因此，实用的光照问题解决方案是亟需的。

(5) 姿态不变的人脸识别算法

与光照问题类似，姿态问题也是目前的人脸识别系统实用化的一个巨大障碍。现有的多数系统对正面人脸识别性能尚可接受，但实用的系统并不能保证输入的人脸是准正面的。如何快速准确地估计姿态并进行相应的矫正以便与数据库中的原型人脸准确匹配是一个极具挑战性的课题。

(6) 人脸信息采集设备带来的问题

摄像头是目前主流的人脸信息采集设备，它们的种类、型号、传感方式（比如 CCD 或者 CMOS）、采样和量化精度等各不相同，采集时刻的焦距、光圈、增益、白平衡等参数设置区别也很大。这些变化因素使得获取的人脸图像数据的变化更大。如何解决这些问题也是值得关注的。

(7) 低质量照片的识别问题

目前的多数人脸识别系统在算法设计和模型训练方面都往往只针对图像质量很好的情况，但对智能监控、公安系统犯罪嫌疑人照片比对等应用而言，由于其来源不一，照片的质量非常差，比如模糊、高噪声、分辨率极低等。如何提高系统对这些低质量照片的识别能力也是人脸识别亟待解决的关键问题之一。

(8) 照片老化问题

人的面部会随着年龄的增长而发生变化，但照片不会，因而导致识别困难，这就是所谓的老化问题。老化分为短期老化和长期老化，短期老化通常与人的生活和精神状态导致的肌肉胖瘦和皮肤属性有关，而长期老化则与头部骨骼的老化密切相关。对很多实用系统来说，老化问题是必须解决的。

(9) 墨镜、帽子等饰物遮挡问题

对不能要求用户配合的场景，例如全自动的安全监控，被识别的对象往往带有墨镜、帽子等饰物，使得面部被遮挡，从而造成信号数据的缺失，如何解决很好的该问题也是值得关注的方向之一。

(10) 化妆、整容带来的问题

现代社会中，面部化妆和整容技术日臻成熟和普遍，因此带来人们面孔表现的更大可变性，这对目前的识别系统而言还是难以解决得，需要从信号传感方式、特征提取等多个环节共同考虑解决该问题。

另外需要注意的是：上述开放问题并不是单独作用的，很多都是组合性的，例如姿

态和光照问题同时出现，会更进一步的增加问题的难度。而目前出现的主要研究论文，往往都是试图孤立的解决其中一个问题，而没有考虑它们同时出现情况下的解决方案，这也是值得关注的一个问题。

2.7.2 可能的技术趋势浅析

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，而这一问题的产生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。因此，本文认为要解决这些问题就需要从信号、特征和符号等不同层面探讨人脸识别的本质计算模型。图 2.2 中列出的从这三个层次的角度列出了十几个技术趋势供读者参考。其中，多摄像机融合识别技术、序列图像分析、基于非线性流形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、泛化能力强对各种变换鲁棒的分类算法等研究课题尤其值得关注。

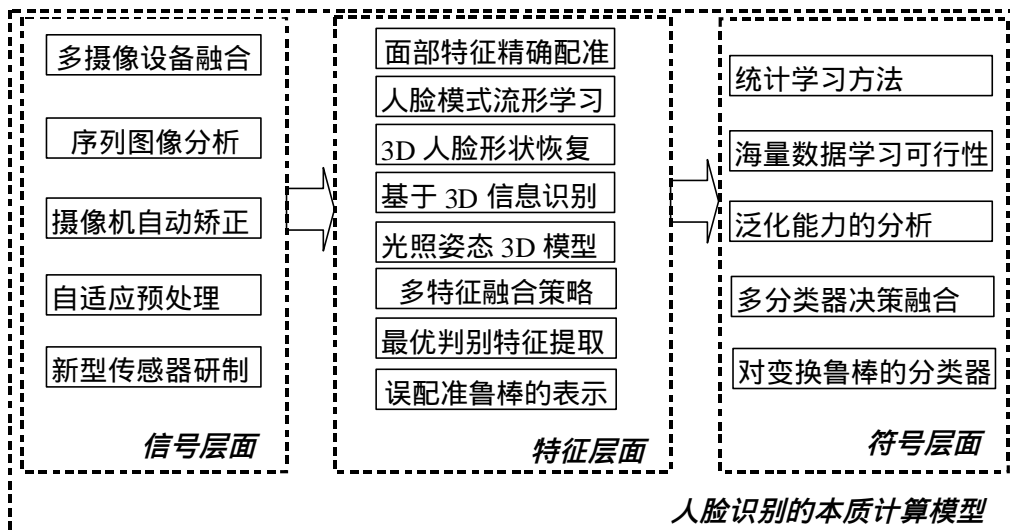


图 2.2 人脸识别技术趋势

Fig.2.2 Some new trends in AFR

2.8 人脸识别研究发展现状总结

自动人脸识别的研究方兴未艾，本章从多个不同的角度探讨了人脸识别研究的发展情况。自动人脸识别研究经过近四十年的研究发展，近年来取得了长足的进步。对于已知人脸数据库规模在 1000 人左右、成像条件理想、用户比较配合的场合，目前最好的人脸识别系统的性能已经达到了 90% 以上，并已经在限定范围内得到了一些应用。但这并不意味着人脸识别已经是一个被解决了的成熟技术问题，相反，人脸识别研究目前仍旧面临诸多挑战性的问题，客观地说，在很多应用条件下，最好的识别系统的识别率也往往达不到 50%，这也是为什么直到目前人脸识别几乎没有典型的成功案例、难以推广的主要原因。因此，研究更加鲁棒、实用的人脸识别核心算法，提供更加有效的影响识别性能的关键问题解决方案，仍将是相关领域内的重点研究内容。

第三章 局部纹理约束的主动外观模型

主动形状模型 (ASM) 和主动外观模型 (AAM) 是图像分析与理解领域著名的算法。在对他们的优缺点进行综合分析的基础上, 本章考虑了二者的集成, 提出了一种局部纹理约束的主动外观模型 (Local Texture Constraint AAM, LTC-AAM), LTC-AAM 的主要特点包括: (1) 针对 ASM 对面部轮廓点定位不够理想的问题, 本文提出了将候选点边缘强度加入面部轮廓点局部纹理匹配模型中的策略, 有效地提高了轮廓点的定位精度; (2) 在利用局部纹理匹配更新特征点位置的基础上, 不但采用了 ASM 的统计形状约束, 还进一步保留了 AAM 中衡量当前纹理与模型纹理匹配度的目标函数, 从而将 ASM 搜索规范到了与 AAM 统一的参数优化框架中; (3) 在 ASM 单步搜索和调整, 仍然借鉴 AAM 的思想, 利用当前纹理来预测外观模型参数的可能变化, 从而实现了启发式的参数优化过程。这些特点使得 LTC-AAM 融合了 ASM 局部模型、AAM 的纹理约束以及基于线性回归的启发式参数优化过程, 从而在较大程度上继承了二者的优势。实验表明, LTC-AAM 能够比较精确的定位预定义的面部关键特征, 为后续的人脸识别与确认打下了良好的特征基础。

3.1 问题的提出

一个实用的完整人脸识别系统通常至少应包括人脸检测与跟踪、面部关键特征点定位、人脸特征提取与比对等不可缺少的步骤, 其中第一个、第三个环节已经被广泛研究, 而面部关键特征点定位却显得相对薄弱, 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, 而大多数分类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, 例如, 我们的研究就表明, 不精确的眼睛特征位置会极大的降低后续的分类识别算法的性能 (参见[SC04]及本论文第五章内容), 而文献研究表明, 特征定位远不是一个已经解决的问题, 尤其需要注意非理想成像条件下的特征定位问题。因此, 为了开发出真正能够实用的人脸识别系统, 必须对面部关键特征的定位问题进行进一步的研究。

在人脸识别乃至模式识别和计算机视觉领域, 目标对象的精确定位和特征配准都是至关重要的。为此, 人们提出了很多方法, 包括蛇模型 (Snake) 方法[KW88]、可变形模板方法 (Deformable Template) [Yu91, YH92]、弹性图匹配技术 (Elastic Graph Matching) [LV93, WF97]、基于光流的向量化技术 (Vectorize) [Be95]、主动形状模型 (Active Shape Model, ASM) [CT95]、主动外观模型 (Active Appearance Model, AAM) [CE98, EC99] 以及最近的 3D 形变模型[BV99, RB02]等。其中, ASM 和 AAM 是最引人注目的, 并出现了他们的很多改进模型, 包括直接外观模型 (Direct Appearance Model, DAM) [HL01], 多视角 AAM [CW00], 纹理约束的 ASM (Texture Constraint ASM, TC-ASM) [LY02, YL03] 和贝叶斯切空间形状模型 (Bayesian Tangent Shape Model, BTSM) [ZG03]等。

可变形模板方法 [Yu91, YH92] 把眼睛特征定义为有两条抛物线 (上下眼睑) 和一个圆 (虹膜) 构成的几何图形, 通过优化的方法来调整其参数以达到最佳的匹配。嘴巴、下巴等也可以采用类似的几何图形建模。但是, 固定的几何图形并不能很好的建模实际上会千差万别的眼睛形状, 而光照、姿态和表情的变化更使得采用这种固定的简单数学

模型难以适应这些复杂的变化，因此很难实现鲁棒的特征提取。优化速度慢、容易陷入局部极小也阻止了该算法的进一步发展。

弹性图匹配[LV93, WF97]是面部关键特征定位的另一个重要算法，该方法将人脸面部关键特征点的属性及其它们之间的位置关系通过一个属性图进行描述，图的顶点对关键特征点的局部纹理建模（通过 Gabor 特征），图的边则反映了特征点之间的距离等位置关系。对新输入的图像，其特征点则通过基于相位预测的位移估计结合图匹配技术来定位。尽管弹性图匹配可以达到较高的定位精度，但速度较慢。

ASM 和 AAM 都是基于点分布模型（Point Distribution Model, PDM）的算法[CT95, CE98, EC99, CW00, St00]。在 PDM 中，外形相似的特定类别物体（比如人脸、人手）的形状通过若干关键的特征点（Landmarks）的坐标串接成原始形状向量。对训练集中的所有形状向量进行对齐操作后，对他们进行 PCA 分析建模，保留的主成分形成最终的形状模型，形状模型的参数反映了形状的主要可变化模式[SH02]。ASM 搜索则首先通过局部纹理模型匹配得到各个特征点的更佳的位置，经过相似变换对齐后，通过统计形状模型对其进行约束，而后再进行局部纹理模型的匹配，形成一个迭代过程，以期形状模型最终匹配到输入的形状模式上去。在 ASM 中，仅使用了特征点局部纹理特征作为启发式信息，没有使用全局的纹理约束，实践中发现 ASM 很容易陷入局部极小。

而在 AAM 中，则采用了形状和纹理二者融合统计约束，即所谓的统计表观模型。AAM 搜索借鉴了基于合成的分析技术（Analysis-By-Synthesis, ABS）的思想，通过模型参数的优化调整使得模型能够不断逼近实际输入模式。模型参数的更新则放弃了 ASM 中的局部纹理搜索过程，仅使用一个线性预测模型根据当前模型和输入模式之间的差别来预测和更新模型参数。AAM 尽管利用了全局纹理，但却抛弃了局部纹理匹配过程，因此会在一定程度上降低关键特征点配准的精度，而且其线性预测模型也有较大的局限性，在初始位置偏离目标位置过大时，则很难收敛到正确位置。

基于对 ASM 和 AAM 的分析，本文提出了一种局部纹理约束的主动表观模型——LTC-AAM，在 AAM 全局纹理约束能量函数以及基于纹理预测形状的启发式参数优化过程之前，引入了 ASM 的局部纹理匹配和统计形状约束[SGW03]。另外，在局部模型匹配方面，针对 ASM 对面部轮廓点定位不够理想的问题，提出在 ASM 的面部轮廓点局部匹配模型中增加边缘约束的策略[WS02]，有效提高了轮廓点的定位精度。这些改进使得 LTC-AAM 融合了 ASM 局部模型、AAM 的纹理约束和启发式的参数优化过程，因而具有更佳的特征配准精度。

3.2 前人工作简介

因本文的算法是 ASM 与 AAM 的融合改进，为了更好的阐述本文的算法，本节首先简单介绍 ASM 以及 AAM 的基本概念和流程。

3.2.1 主动形状模型

基于关键特征点的点分布模型（PDM），Cootes[CT95]等人提出了主动形状模型用于图像分析和理解。为了建立 PDM 模型，需要给定一个标定了面部关键特征点的学习集合 $L = \{(I_i, s_i) | i = 1, \dots, m; s_i = (x_i^1, y_i^1, \dots, x_i^N, y_i^N)^T\}$ ，其中 m 为训练样本数目， N 为预先定义的关键特征点的数目， L 中的每个形状向量 s_i 是由训练图像 I_i 上预先定义并手工标

定的 N 个关键特征点的横纵坐标串接而成。然后，通过形状对齐算法将学习集中的所有形状对齐到同一坐标体系下，得到对齐后的特征集合 $L' = \{s'_i | i = 1, \dots, m\}$ ，对这些对齐后的形状作主成分分析，得到统计形状模型如下：

$$s \approx \bar{s} + P_s b_s \quad (3-1)$$

其中， \bar{s} 为平均形状， b_s 为形状参数， P_s 为主成分特征向量构成的变换矩阵，通过对训练集的协方差矩阵的特征值分解得到，主成分特征向量反映了形状变化的主要模式。公式 3-1 可以直观解释为：任意形状均可以近似表示为对平均形状的“变形”，而且这种“变形”是通过形状参数 b_s 对若干形状变化模式的加权和来建模的。

ASM 的另一个重要特点是它针对每个 Landmarks 建立局部纹理模型，以便在形状搜索时作为启发式信息更新每个特征点的位置。简单的说，局部纹理模型就是每个特征点在其某种邻域内的图像亮度分布模型，Cootes 等人采用的是 Profile 邻域，定义为形状上特征点处法线方向上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，用高斯分布建模作为局部纹理模型，如图 3.1 所示。

局部纹理模型在训练阶段从学习集中统计得到，在模型适应阶段则用来进行 Landmarks 位置的自动更新，也就是说，对 p 点 Profile 邻域内的每个候选点 q ，都计算其局部纹理采样 l_q 与训练集中 p 点局部纹理之间的马氏距离：

$$f(l_q) = (l_q - \bar{l}_p)^T \Sigma_p^{-1} (l_q - \bar{l}_p) \quad (3-2)$$

其中， \bar{l}_p, Σ_p 分别为训练集中图像在 p 点处局部纹理的平均和协方差矩阵。该距离最小的候选点 p^* 作为新的特征点。

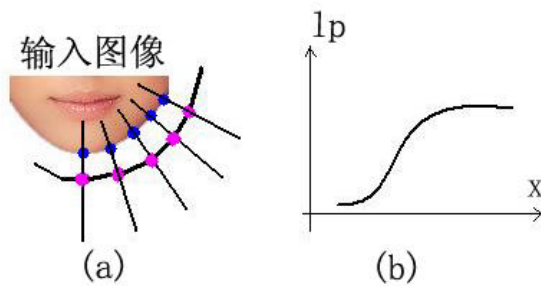


图 3.1 局部纹理模型图示

Fig. 3.1 Profile model for each landmark

基于上述统计形状模型和局部纹理模型，给定一幅新的输入人脸图像 I ，ASM 提取其中人脸形状的基本过程如下，图 3-2 形象的图示了 ASM 的搜索过程。

算法 3.1：ASM 搜索过程

1. 令 $t=0$, 用平均形状作为初始化形状 s_t , 参见图 3.2 (a);
2. 对每个 landmark, 对当前形状在该点处的 Profile 上的每个候选点, 计算公式 3-2 给出的距离度量, 选择具有最小距离的那个点作为该 Landmark 的新位置, 从而得到 s'_t , 参见图 3.2 (b);
3. 调整平移、旋转和尺度等相似变换参数以及形状参数 (x_0, y_0, t, q, b_s) , 使得 s'_t 与模型形状 $s'' = T^{-1}(\bar{s} + P_s b_s)$ 最佳匹配; 令 $s_{t+1} = T^{-1}(\bar{s} + P_s b_s)$, 参见图 3.2 (c);
4. 如果 s_t 和 s_{t+1} 之间差别足够小, 则宣布算法收敛; 否则转 2;

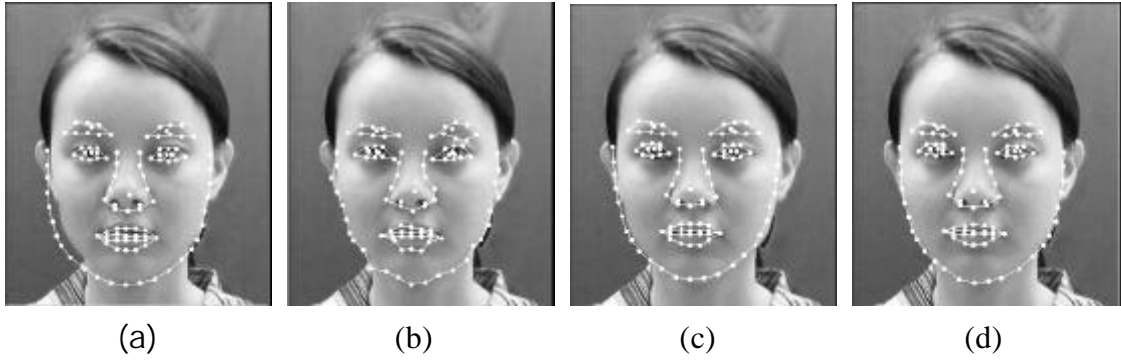


图 3.2 ASM 搜索过程示例

Fig. 3.2 Illustration of ASM searching

3.2.2 主动表观模型 (AAM)

在 ASM 的基础上, Cootes 等人进一步提出了 AAM 模型[CE98, EC99], 在统计形状模型的基础上, 进一步对纹理 (将人脸图像变形到平均形状而得到的形状无关图像 g) 进行统计建模, 并将形状和纹理两个统计模型进一步融合为表观模型。即给定如 3.2.1 节所述的学习集 $L = \{(I_i, s_i) | i = 1, \dots, m; s_i = (x_i^1, y_i^1, \dots, x_i^N, y_i^N)^T\}$, 所有形状对齐后计算平均形状 \bar{s} , 将所有训练图像变形到平均形状, 从而得到 $L' = \{(I_i, s'_i, g_i) | i = 1, \dots, m\}$, 对所有的纹理进行主成分分析, 从而得到如下统计模型:

$$g = \bar{g} + P_g b_g \quad (3-3)$$

其中, \bar{g} 为平均纹理, P_g 为 PCA 计算得到的纹理主成分特征向量形成的变换矩阵, b_g 为控制纹理变化的统计纹理参数。

统计表观模型则进一步将公式 3-1 和公式 3-3 定义的形状和纹理模型融合起来。即: 将 b_s 和 b_g 串接起来得到新的表观特征向量:

$$b = \begin{pmatrix} w_s b_s \\ b_g \end{pmatrix} \quad (3-4)$$

其中，对角阵 w_s 用来调整 b_s 和 b_g 二者之间量纲的不同。对得到的 b 进行 PCA，进一步消除形状和纹理之间的相关性，从而得到表观模型：

$$b = \bar{b} + Qc \quad (3-5)$$

其中， \bar{b} 为平均表观向量， Q 为表观主成分特征向量形成的变换矩阵， c 为控制表观变化的统计表观参数。这样，给定表观模型参数 c 以及相应的相似变换参数 $J = (x_0, y_0, t, q)$ ，就可以合成一幅模型人脸图像：

$$I_m = T(\text{warp}^{-1}(s_m, g_m); J) \quad (3-6)$$

其中： warp^{-1} 表示将平均形状下的模型纹理 g_m 变形为模型形状 s_m 的操作， T 则表示对其进一步进行参数为 J 的相似变换，从而得到模型图像 I_m 。而 g_m 和 s_m 则分别通过公式 3-7 计算：

$$\begin{aligned} s_m &= \bar{s} + P_s b_s \\ g_m &= \bar{g} + P_g b_g \end{aligned} \quad (3-7)$$

其中 b_s, b_g 通过下式计算：

$$b = (W_s b_s, b_g)^T = \bar{b} + Qc \quad (3-8)$$

基于上述模型，输入一幅新的人脸图像 I_n ，AAM 对其进行特征配准的基本思路是一个表观模型参数 c 的优化过程，以期最终达到模型纹理与输入纹理的最佳匹配。具体的是要优化公式 3-9 给出的目标函数：

$$\Delta = \|dg\|^2 = \|g_s - g_m\|^2 \quad (3-9)$$

其中 g_m 通过公式 3-6 计算得到，而 g_s 则是输入图像根据当前形状 s 形变到平均形状得到的纹理。

尽管上述优化过程比较清晰明了，但通常来说，表观模型参数可能接近有数十个之多（例如本文实验中形状参数 20 个，纹理参数 110 个，表观模型参数为 80 个），因此，采用一般的优化算法速度会难以接受。Cootes 等人进一步提出了一种基于纹理预测参数变化的启发式参数搜索策略，这种策略利用了输入和模型纹理之差别 dg 与模型参数变化 dc 以及相似变换参数变化 dJ 之间存在一定程度的线性关系，因此，可以通过线性回归进行参数预测，即：

$$dc = R_c dg \quad (3-10)$$

$$dJ = R_J dg \quad (3-11)$$

其中线性回归变换矩阵 R_c 和 R_t 可以在训练阶段通过人为加扰的方法得到。

最终的 AAM 搜索过程如算法 3.2 所示。图 3.3 给出了 AAM 搜索过程的一个示例。

算法 3.2：AAM 搜索流程

- 1、初始化： $t=0$ ，模型参数初始化为 $c_t = 0$ ；参见图 3.3 (b)。
- 2、计算模型纹理和当前纹理之差： $dg = g_s - g_m$ ；
- 3、利用线性回归模型预测参数变化： $dc = R_c dg$ ；
- 4、尝试新的模型参数： $c'_t = c_t - kdc$ ， $k=1$ ；
- 5、重新计算误差函数： dg'
- 6、比较 $\|dg\|, \|dg'\|$
 如果 $\|dg'\| < \|dg\|$ ，则接受 c'_t 作为新的参数；参见图 3.3 (c, d)。
 否则转第 4，尝试 $k=1.5; 0.5; 0.25$ 等
- 7、 $t=t+1$ ，如果 $\|dg'\| - \|dg\| < \epsilon$ 或者迭代次数 t 超过预定次数，则退出，否则转 2

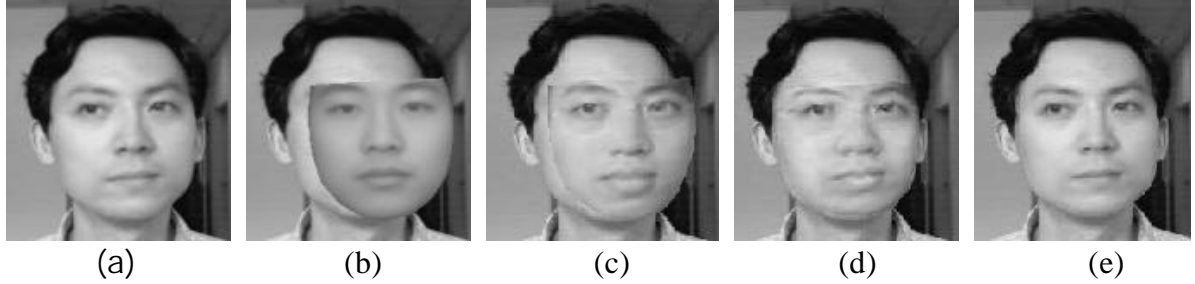


图 3.3 AAM 优化过程示例 (a)输入图像 (b)初始模型 (c)一次迭代后的模型图像 (d)五次迭代后的模型图像 (e)最终模型图像

Fig. 3.3 Illustration of AAM Searching (a) Input Image; (b) Initial model; (c) Model image after one iteration; (d) Model image after 5 iterations; (e) Final model.

3.3 局部纹理约束的主动表观模型 LTC-AAM

在 ASM 中，仅使用了特征点的局部纹理特征作为启发式信息，没有使用全局的纹理约束。而在 AAM 中，模型参数的更新则放弃了 ASM 中的局部纹理搜索过程，因此会在一定程度上降低关键特征点配准的精度。基于此，本文提出了一种局部纹理约束的主动表观模型——LTC-AAM，除保留了 AAM 中的全局纹理约束和基于纹理预测形状的启发式的参数优化过程之外，还引入了 ASM 局部纹理匹配和全局形状约束过程。另外，在局部模型匹配方面，针对 ASM 对面部轮廓点定位不够理想的问题，本文提出在 ASM 的面部轮廓点局部匹配模型中增加边缘约束的策略，有效提高了轮廓点的定位精度。这些改进使得 LTC-AAM 融合了 ASM 局部纹理模型、AAM 的纹理约束和启发式的参数优化过程，因而具有更佳的特征配准精度。

3.3.1 边缘约束局部纹理模型

在 ASM 中,基于局部纹理模型匹配的关键特征点更新是形状搜索成功的关键环节。在实验中我们发现,采用单纯的 Profile 邻域一阶导数建模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。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[SGW02]。具体的,本文将局部纹理模型距离度量(公式 3-2)修正为:

$$f(l_q) = (C - E_q) \cdot (l_q - \bar{l}_p)^T \Sigma_p^{-1} (l_q - \bar{l}_p) \quad (3-12)$$

其中, E_q 为候选点 q 处的边缘强度, C 为一个经验常数。公式 3-12 的意义在于它可以更容易的将特征点“吸引”到边缘强度大的位置上,这对于面部轮廓点来说是非常有利的。本文实验中,采用了 Sobel 边缘算子,常数 C 取值为 300。图 3.4 是一个采用这种改进的效果图示。从图 3.4 可以看出,由于引入了边缘强度约束项,大多数面部轮廓上的特征点局部匹配变得更加鲁棒精确,因而更加有利于最终人脸形状的精确获取。

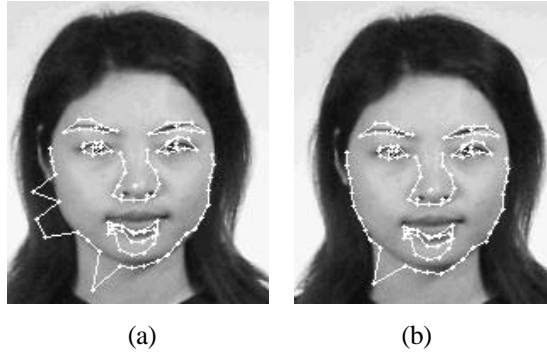


图 3.4 加入边缘强度约束的局部纹理匹配效果 (a)ASM 一次迭代局部匹配结果;(b)边缘强度约束后的局部纹理匹配结果

Fig. 3.4. Comparison of the local texture matching w/o edge constraints. (a) Matching result of the standard ASM. (b) Matching result with edge constraint

3.3.2 LTC-AAM 基本思想

本质上,ASM 和 AAM 都是基于参数优化策略来完成特征定位的,不同的是 ASM 仅使用了特征点的局部纹理匹配结果作为参数搜索时的启发式信息,而没有使用全局的纹理约束,也没有像 AAM 那样明确定义需要最小化的目标函数。这使得 ASM 具有较快的搜索速度,但却难以保证最终总体定位结果的正确性,实践中我们经常发现 ASM 会出现从初始形状搜索到一个正确结果,而后却再此偏离正确结果的奇怪现象。

在 AAM 中,不但明确定义了需要优化的目标函数——即当前纹理和模型纹理的差别,而且模型参数的更新则放弃了 ASM 中的局部纹理搜索过程,使用了纹理差与模型参数之间的近似线性关系作为参数搜索时的启发式信息。其优点是可以较好的保障总体定位结果的正确性,但却难以保障各个关键特征点的精确定位,同时参数优化也导致 AAM 收敛速度较慢。尽管基于纹理差别预测参数变化可以提高优化速度,但这种线性近似关系在模型纹理与实际纹理差别较大不复存在,因此预测的结果往往会偏差较大。

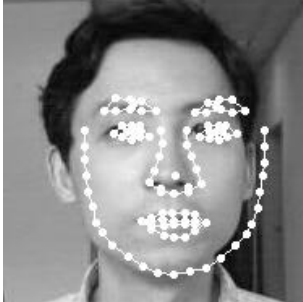
基于上述分析, 本文将二者结合起来, 提出了一种局部纹理约束的主动表观模型——LTC-AAM。该方法将 ASM 局部纹理匹配和全局形状约束引入 AAM 的全局纹理约束和基于纹理预测形状的启发式的参数优化过程中。从而使得 LTC-AAM 融合了 ASM 局部模型、AAM 的纹理约束和启发式的参数优化过程, 因而具有更佳的特征配准精度。

3.3.3 LTC-AAM 算法流程

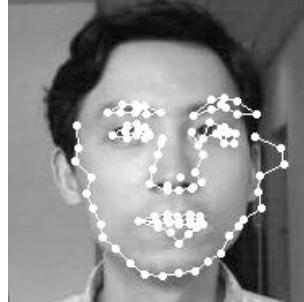
算法 3.3 给出了 LTC-AAM 的搜索流程, 图 3.5 给出了一个搜索过程的示例。需要说明的是, 简明起见, 该算法流程中没有给出对于相似变换参数的相应处理过程, 这些过程与 ASM 和 AAM 中的处理方式完全一致。

算法 3.3 : LTC-AAM 搜索流程

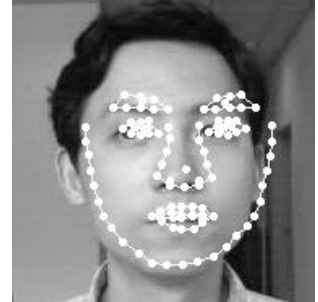
1. 初始化: $t=0$, 模型参数初始化为 $c_t=0$; 计算当前纹理与模型纹理之差 $dg = g_s - g_m$; 参见图 3.5 (a)。
2. 根据 c_t 计算 s_t , 对 s_t 中每个关键特征点进行局部纹理匹配, 其中面部轮廓点采用 3.3.1 节所述边缘约束方法, 得到 s'_t , 参见图 3.5 (b)。
3. 计算 s'_t 的最佳匹配模型形状 \hat{s}_t ; 令 $s'_t \leftarrow \hat{s}_t$, 并计算 c'_t ;
4. 计算当前纹理与模型纹理之差: $dg' = g_s - g_m$;
5. 如果 $\|dg'\| < \|dg\|$, 则接受 ASM 单步结果 $c''_t \leftarrow c'_t, dg'' \leftarrow dg'$;
否则不接受 ASM 单步结果 $c''_t \leftarrow c_t, dg'' \leftarrow dg$; 参见图 3.5 (c)。
6. 利用线性回归模型预测参数变化: $dc = R_c dg''$, 令 $k=1$;
7. 尝试新的模型参数 $c''_t = c''_t - kdc$, 计算误差函数 dg''
8. 比较 $\|dg''\|, \|dg''\|$
如果 $\|dg''\| < \|dg''\|$, 则转 10 接受 c''_t 作为新的模型参数; 否则尝试 $k=1.5; 0.5; 0.25$ 等转 7; 参见图 3.5 (e)。
9. $c_{t+1} \leftarrow c''_t, dg \leftarrow dg'', t=t+1$;
10. 如果第 5 步或者第 8 步中纹理差别已经收敛, 则算法结束;
如果迭代次数 t 超过预定次数, 则退出, 否则转 2



(a) 初始形状



(b) 局部模型匹配



(c) ASM 形状约束



图 3.5 LTC-AAM 的搜索步骤图示

Fig.3.5 Searching procedure of the LTC-AAM

概括地说, 该算法的搜索过程包括两大步骤:

- (1) 局部纹理模型匹配及形状统计模型约束局部搜索的过程 (对应算法 3.3 中的步骤 2-5, 参见图 3.5 (b-c))。这一过程对应于 ASM 中的局部搜索和形状调整过程, 不同的是本文在面部轮廓特征点的局部纹理模型匹配中引入了边缘信息以改善轮廓点的定位精度。同时, 在迭代过程中是否接受 ASM 的搜索结果由能量函数——当前纹理与模型纹理差别——是否下降来决定 (步骤 4 和步骤 5);
- (2) 根据纹理差线性回归预测并调整模型参数变化过程 (对应算法 3 中的步骤 6-8, 参见图 3.5 (d-e))。这一步骤对应于 AAM 中的单步迭代, 即在第一步 ASM 搜索的基础上, warp 出当前形状对应的当前纹理, 根据当前纹理与模型纹理的差别, 利用线性回归方法预测可能的参数变化, 以实现启发式的参数优化过程。

3.3.4 主要特点总结

尽管 LTC-AAM 是 ASM 与 AAM 二者的集成, 但并非是 ASM 与 AAM 的简单拼接, 而是二者优势互补的有机结合。这主要体现在以下几点:

- (1) 从对 ASM 改进的角度看, LTC-AAM 引入了明确的目标函数 (当前纹理与模型纹理的匹配程度), 这与 ASM 中隐含的优化目标——搜索形状与模型形状最佳匹配——相比, 其优化的目标函数更加合理明确;
- (2) 从对 AAM 改进的角度看, LTC-AAM 引入了局部纹理匹配过程, 一方面改善了 AAM 线性预测模型的“线性”有效范围较小因而初始模型不能有太大偏离的缺点, 加速了特征搜索的过程; 另一方面也使得特征点有更高的匹配精度, 尤其是对于面部轮廓点, 边缘强度加入到局部纹理匹配模型中, 更提高了其定位精度;
- (3) 从二者的衔接来看, 上述算法中步骤 4, 5 和 6 实现了良好的衔接, 而且 ASM 搜索过程和 AAM 参数优化过程采用了统一的目标函数, 因而可以保障优化过程的一致性。

3.4 实验与分析

本文采用的训练和测试图像共包括来自 FERET, Bern, 以及 CAS-PEAL 人脸库中的 500 幅准正面人脸图像, 其中男女比例大约为 3:2, 涵盖不同的表情 (包括笑、吃惊和愤怒), 少量的图像有光照条件的变化。全部人脸图像手工标定了 103 个关键特征点。图

3.6 给出了其中的两幅人脸图像及其手工标定的 103 个特征点的位置情况。

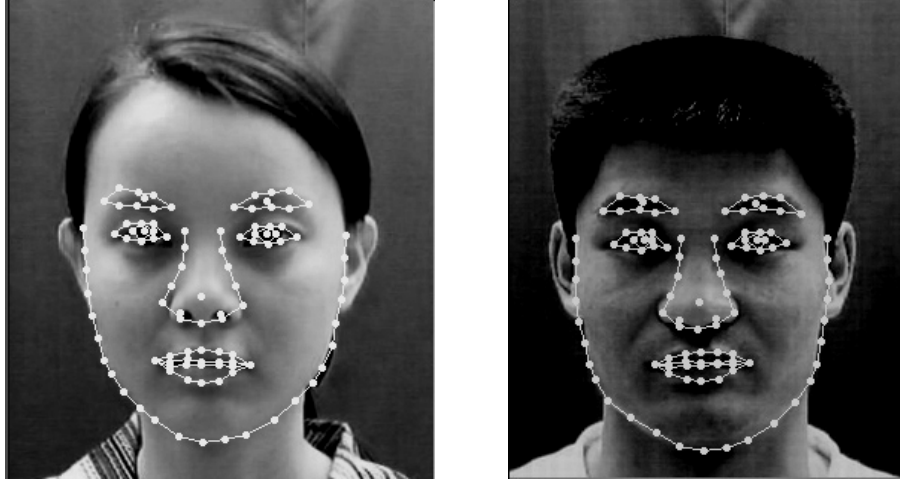


图 3.6 本文采用的 103 个特征点的情况

Fig. 3.6 The 103 manually labeled facial landmarks used in this thesis

如何合理地评价特征配准算法的性能到目前为止尚未形成公认的标准，主观评价法是最常用的，其局限性比较大。本文则采用了一种相对比较客观的评价策略：“平均定位误差法”，其定义如下：

$$d = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n \frac{\text{dist}(P_{ij}, P'_{ij})}{D_i} * 50 \right), \quad (3-13)$$

其中， m 为测试图像的总数目， n 为面部关键特征点的数目（本文中 n 为 103）， P_{ij} 分别为第 i 幅测试图像中手工标定的第 j 个特征点的坐标， P'_{ij} 则为算法自动定位的第 j 个特征点的坐标， $d(P_{ij}, P'_{ij})$ 则表示二者之间的直线距离， D_i 则表示第 i 幅测试图像中标定的双眼之间的直线像素距离。

公式 3-13 所定义的“平均定位误差”的物理含义是很直观的：它测量了假定图像中双眼距离为 50 像素的情况下，所有自动定位的特征点的坐标偏离手工标定“真实坐标”的平均距离。

本文实验中，我们从 500 幅训练图像中随机选择 350 幅作为训练图像，剩下的 150 幅图像作为测试图像，并如此进行了 10 次不同的实验。在本文实验中形状的初始化是根据眼睛位置调整平均形状完成的，其中眼睛的自动定位采用的是我们在文献[CS02]中给出的区域增长眼睛中心搜索算法。实验中主要参数的设置情况为：统计形状模型的维数为 20，纹理模型的维数为 110，二者融合后的表观模型参数为 80。

对每次实验都根据公式 3-13 计算平均定位误差，并与 ASM、AAM 的性能进行了对比，对比结果如表 3.1 所示，其中，表 3.1 中所列出的定位速度是算法在 PIV1.4G 计算机上的运行的平均速度。可见，尽管由于其引入了更多的约束因而 LTC-AAM 的运行速度比 ASM 有所降低，但却高于 AAM，而且总体而言 LTC-AAM 的定位精度比 ASM 和 AAM 都有了一定程度的提高。图 3.7 则给出了 LTC-AAM 的若干正确定位结果的示例，图 3.8 则给出了一些定位不够精确的图像示例。

表 3.1 平均定位误差的比较

Table 3.1 Performance comparison of different methods

定位方法	ASM	AAM	LTC-AAM
D (像素)	3.12	2.77	2.54
定位速度(秒)	0.35	0.72	0.47

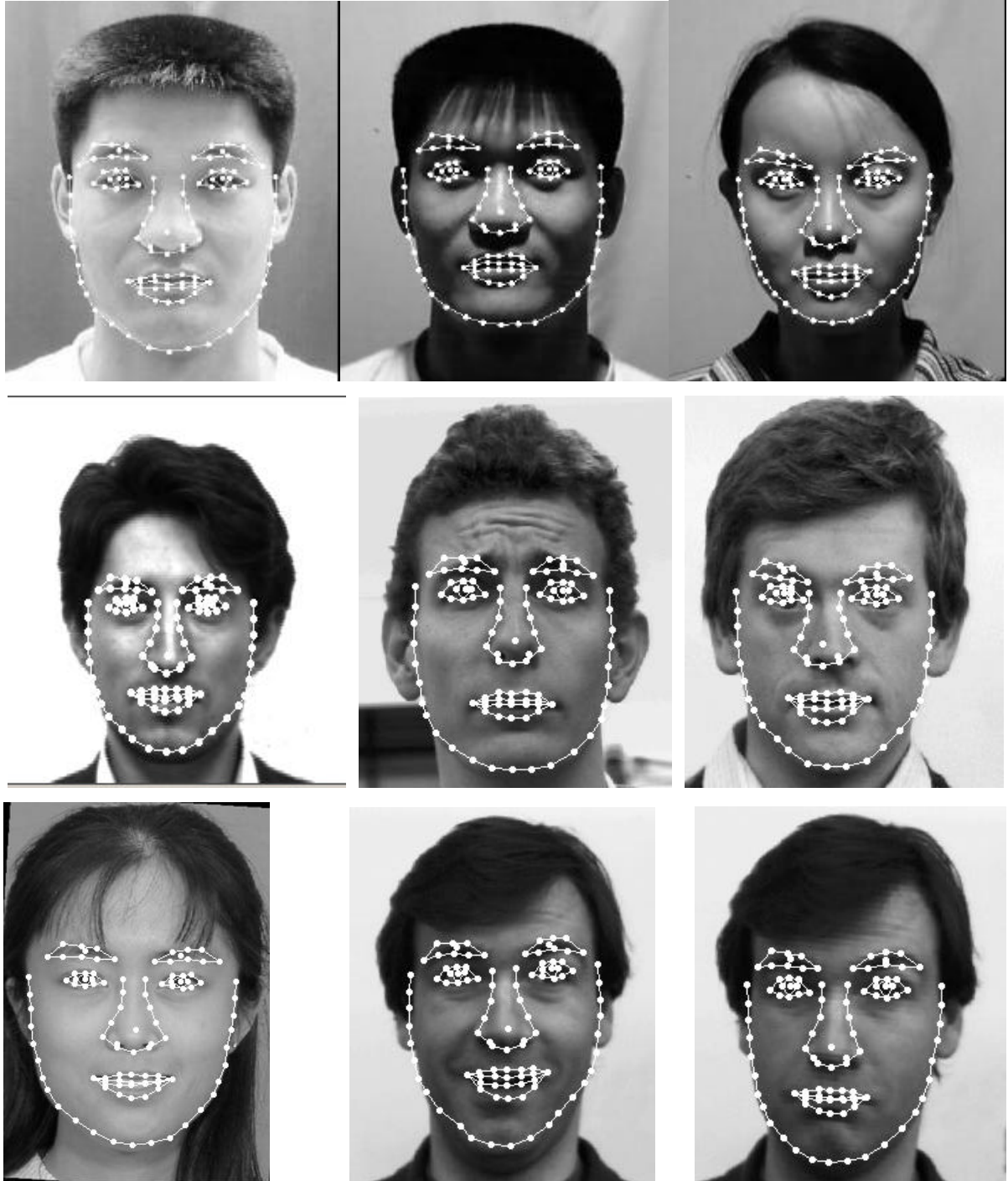


图 3.7 LTC-AAM 的特征定位结果示例

Fig. 3.7 Some experimental results of the LTC-AAM

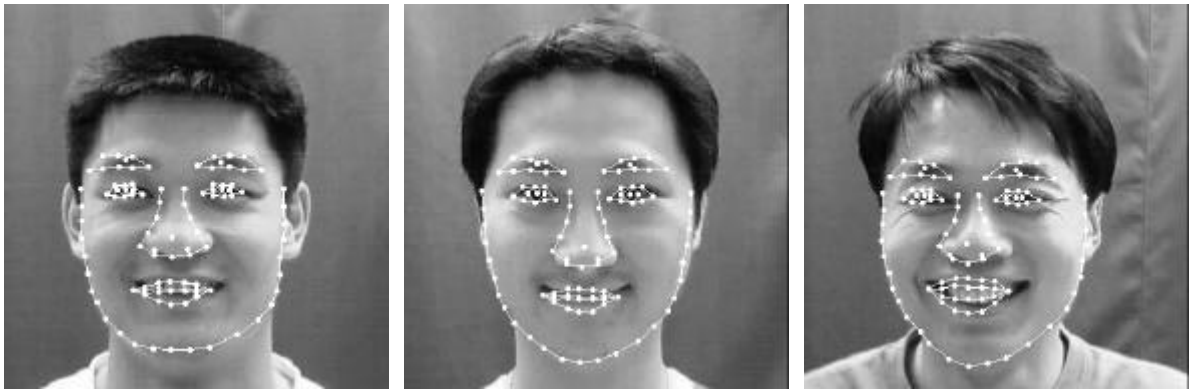


图 3.8 定位结果不理想的图像示例

Fig. 3.8 Some examples with poor alignment results

3.5 本章小结

本章在综合分析主动形状模型 (ASM) 和主动表观模型 (AAM) 优缺点的基础上, 考虑了它们的集成, 提出了一种局部纹理约束的主动表观模型 (LTC-AAM), LTC-AAM 的主要贡献包括:

- (1) 针对 ASM 对面部轮廓点定位不够理想的问题, 本文提出了将候选点边缘强度加入面部轮廓点局部纹理匹配模型中的策略, 提高了轮廓点的定位精度;
- (2) 在 AAM 中引入了 ASM 的局部纹理匹配模型更新特征点位置的局部搜索策略和基于统计形状模型的形状约束, 并继承了 AAM 中衡量当前纹理与模型纹理匹配度的目标函数作为进一步的约束条件, 从而统一了 ASM 和 AAM 的参数优化框架;
- (3) 在 ASM 单步搜索和调整后, 利用当前纹理来预测表观模型参数的可能变化, 从而继承了 AAM 的启发式的参数优化优点。

上述贡献使得 LTC-AAM 融合了 ASM 局部模型、AAM 的纹理约束以及基于线性回归的启发式参数优化过程, 从而在较大程度上继承了二者的优势。实验表明, LTC-AAM 能够比较精确的定位预定义的面部关键特征。

第四章 基于 AdaGabor 特征判别分析的人脸识别方法

人脸描述是人脸识别的核心问题之一，人脸识别的研究实践表明：在人脸三维形状信息难以准确获取的条件下，从图像数据中提取多方向、多尺度（频率）的 Gabor 特征是一种合适的选择。使用 Gabor 特征进行人脸识别的典型方法包括弹性图匹配方法（EGM）和 Gabor 特征判别分类法（GFC）。EGM 在实用中需要解决关键特征点的定位问题，而且其速度也很难提高；而 GFC 则直接对下采样的 Gabor 特征用 PCA 降维并进行判别分析，尽管这避免了精确定位关键特征点的难题，但下采样的特征维数仍然偏高，而且简单的下采样策略很可能遗漏了非常多的有用特征。本文针对这一问题，考虑如何对 Gabor 特征进行有效降维，将目前受到极大关注的 AdaBoost 算法创新性的应用于 Gabor 特征的选择上来，用于提取对识别最有利的特征（本文称这些选择出来的 Gabor 特征为 AdaGabor 特征），并最终通过对 AdaGabor 特征的判别分析实现识别（本文称该方法为 AdaGaDA 识别方法）。在 CAS-PEAL 和 FERET 人脸图像库上的对比实验表明：AdaGaDA 方法不但可以大大降低 Gabor 特征的维数，而且识别精度也有了较大的提高。

4.1 问题的提出

识别算法的精度和鲁棒性不仅取决于采用什么样的分类器，而在更大程度上取决于采用什么样的特征描述人脸，亦即人脸表示（representation）的问题。理论上讲，良好的人脸表示可以使得最简单的分类器具有良好的识别性能[DH03]。好的人脸表示应该具有以下特点：类内差别尽可能小，类间差别尽可能大；对不改变类别属性的变换具有鲁棒性；其提取应尽可能少的依赖手工标注。人脸三维形状信息和表面反射属性应该是比较好的人脸表示方法，但遗憾的是它们难以从 2D 图像数据中准确获取，因而并不实用。从本质上讲，目前的多数人脸识别方法，包括模板匹配、Eigenface、Fisherface、Bayesian 方法等均直接利用了 2D 图像表现（像素的灰度值）作为人脸表示，其缺点是受光照、视角、表情、饰物、遮挡等外界变化条件的影响比较大。

人脸表示的另一常用的思路是对 2D 图像表现做某种变换，比如从图像的空域表示变换到频域表示，常用的变换包括付立叶变换[LY99，赖阮 99，LY01]、离散余弦变换、各种小波变换等。其中，Gabor 变换是最受关注的一种，这主要是因为 Gabor 变换所采用的核(Kernels)与哺乳动物视觉皮层简单细胞 2D 感受野剖面(Profile)非常相似，具有优良的空间局部性和方向选择性，能够抓住图像局部区域内多个方向的空间频率（尺度）和局部性结构特征。这样，Gabor 分解可以看作一个对方向和尺度敏感的有方向性的显微镜[Dau85，Lee96]。同时，2DGabor 函数也类似于增强边缘以及峰、谷、脊轮廓等底层图像特征，这相当于增强了被认为是面部关键部件的眼睛、鼻子、嘴巴等信息，同时也增强了诸如黑痣、酒窝、伤疤等局部特征，从而使得在保留总体人脸信息的同时增强局部特性成为可能[LV93，WF97]。

前人工作也已经表明采用 Gabor 特征描述人脸能够取得非常好的识别效果。典型的方法包括：动态连接结构（Dynamic Link Architecture, DLA）[LV93]及其后续的弹性图匹配技术（Elastic Graph Matching, EGM）[WF97]，Gabor 小波网络（Gabor Wavelet Network, GWN）[Kr00]，Gabor 特征 Fisher 分类（Gabor Fisher Classifier, GFC）[LW02]等等。

弹性图匹配技术 EGM[WF97]将人脸表示为一个由若干预先定义的顶点和边构成的

属性图 $G = \{V, E\}; V = \{V_1, V_2, \dots, V_n\}; E = \{E_1, E_2, \dots, E_m\}$,图 G 的每个顶点 V_i 对应人脸上某个预先定义好的、具有特定高层语义的特征点位置, 顶点的属性为以该点为中心计算得到的多尺度、多方向 Gabor 变换特征形成的特征串 J_i (一般称为一个 *Jet*); G 的每条边 E_i 对应两个顶点之间的连接, 其属性为两点之间的距离 d 。识别则最终通过图之间的弹性

匹配来完成。通过合理的选择特征点的位置, 弹性图可以很好地同时对人脸的关键局部特征及其它它们之间的关系进行建模, 从而既利用了人脸的图像亮度分布特征, 又利用了面部结构信息。弹性图匹配方法的主要缺点是图的构造和匹配过程复杂度高, 特征点的精确定位难免会陷入局部最小, 从而使得构造的属性图不够精确而影响识别结果。

与弹性图匹配方法预先定义特征点位置及其 Gabor 小波参数设置的方法不同, GWN 方法[Kr00]既不预先定义特征点位置, 也不预先设定 Gabor 小波的参数, 而是从最佳描述的角度, 通过一个优化过程完全自动地确定特征位置及其相应的小波参数, 并最终用这些参数作为人脸的描述来进行人脸识别。GWN 是一种基于参数优化的建模方法, 计算复杂度很高, 其用于人脸识别的有效性还需要更多的实践检验。

使用 Gabor 特征的最直接的方式是在粗略对齐人脸图像的前提下, 逐像素计算每个像素点的多方向、多尺度 Gabor 变换特征, 形成 Gabor 特征的人脸表示, 然后进一步进行后端的特征提取和分类。这种思路的优点是计算简单, 只需要精确定位眼睛位置而不需要定位其他关键特征点。Liu 提出的 GFC 方法是该思路的典型代表[LW02], 为了解决逐像素 Gabor 特征维数过高的问题, Liu 采用了简单的下采样处理方法, 然后用 PCA 进行降维, 最后用增强的判别分析方法完成识别。尽管这一定程度上解决了维数问题, 但简单的下采样策略导致大量判别特征的丢失, 从而使得后续的分类精度下降。

针对逐像素 Gabor 特征人脸表示方法维数过高, 冗余度过大的问题, 本文提出了一种用 AdaBoost 对 Gabor 特征进行优化选择和判别分析的识别方法。在将 AdaBoost 应用于人脸识别时, 本文借鉴了 Baback 所提出的类内差和类间差思路[MP00], 从而解决了将两类学习算法应用于多类人脸识别问题的难点。尔后, 我们把 AdaBoost 作为纯粹的特征选择方法, 对选择出来的特征应用判别分析技术, 最后用最近邻进行分类, 实验结果表明其能够取得良好的识别效果。

4.2 前人相关工作介绍

4.2.1 Gabor 特征提取与性质

Daugman[Dau85]将 Gabor 滤波器扩展到了 2D 情况, 用以建模方向选择简单细胞的感受野。2D 形式的 Gabor 滤波器定义为一个用高斯包络函数约束的平面波:

$$y_{\vec{k}}(\vec{z}) = \frac{\vec{k}^2}{s^2} e^{-\frac{\vec{k}^2 \vec{z}^2}{2s^2}} [e^{i\vec{k}\vec{z}} - e^{-s^2/2}] \quad (4-1)$$

其中, 方括号中的第一项决定了 Gabor 核的震荡部分, 第二项则为补偿直流分量, 用以消除核函数响应对图像亮度绝对值变化的依赖性(以保证不同亮度值构成的均匀亮度区域的响应相同)。不难看出, 复数形式的 Gabor 核函数是奇(正弦波)偶(余弦波)项

的组合。其中，参数 \vec{k} 控制着高斯窗口的宽度、振荡部分的波长和方向，参数 s 则决定了窗口宽度和波长的比例关系，亦即高斯包络函数所影响的震荡数目。公式 4-1 实际上定义了一组在平移、旋转和尺度变化意义下自相似的小波函数，也就是说，该公式定义的所有滤波器，均可以通过对某一 Gabor 滤波器进行适当的平移、旋转和尺度变化来得到。因此公式 4-1 定义的所有 Gabor 核函数都具有相似的形状，只是方向和大小不同。图 4.1 给出了公式 4-1 定义的一个 Gabor 核函数的实部和虚部在空域中的形状。可见，其实部和虚部具有相似的形状，但相位不同。

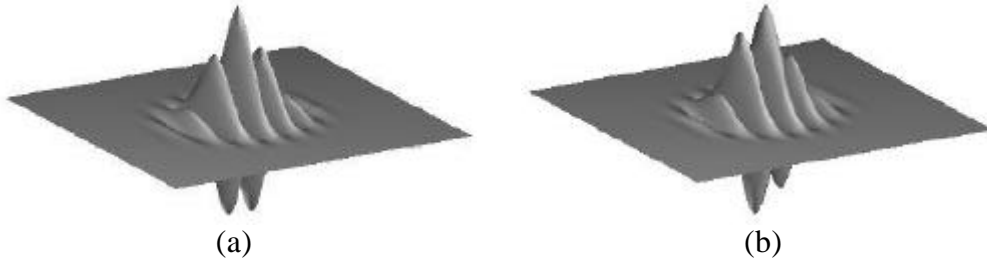


图 4.1 Gabor 核函数的实部和虚部在空域中的形状
(a)实部(余弦相位) (b)虚部(正弦相位)

Fig. 4.1 Shape of the Gabor kernels in the spatial domain
(a)Real part (Cosine phase); (b)Imaginary part (Sine phase)

上面定义的 Gabor 函数族对图像 I 的滤波过程则可以通过它们与图像的卷积来实现：

$$J_{\vec{k}, \vec{x}_0} = \int \mathbf{y}_{\vec{k}}(\vec{x}_0 - \vec{x}) I(\vec{x}) d^2x = (\mathbf{y}_{\vec{k}} * I)(\vec{x}_0) \quad (4-2)$$

实际运算时，该卷积过程必须既要在空间域内（图像位置 \vec{x}_0 ，比如逐像素）进行采样，也要对小波函数即频域（由 \vec{k} 控制的高斯窗宽、震荡的方向和波长）进行采样。小波函数采样最常用的采样方法是令[LV93，WF97]：

$$\vec{k} = k_v e^{i f_u} \quad (4-3)$$

其中：

$$k_v = k_{\max} / f^v, \quad f_u = \frac{u p}{8} \quad (4-4)$$

这里 f 是一个限定频域中核函数距离的间隔因子，通常取为 $\sqrt{2}$ 。最终的采样则通过设定 5 个不同的尺度 $v \in \{0,1,2,3,4\}$ 和 8 个不同的方向 $u \in \{0,1,2,3,4,5,6,7\}$ 来完成，另外两个参数的取值通常为 $s = 2p$ ， $k_{\max} = p/2$ 。结合公式 4-1 不难发现：这样定义的 Gabor 小波的高斯窗的实际窗宽为 $s' = s / k_v = s f^v / k_{\max} = 4 \times 2^{v/2}$ ，而 u 的取值变化则导致振荡方向的变化。图 4.2 和图 4.3 给出了这样采样得到的 Gabor 核函数的实部在空域中的图示，从中可以明显地看出尺度参数 v 和方向参数 u 的变化导致的滤波器变化情况。

上述的每个 Gabor 滤波器都可以对应于一个初等视觉皮层简单细胞的空间感受野的

信号处理过程,对图像处理而言,容易理解,Gabor 滤波函数将在与其振荡方向垂直的边缘处产生强烈的响应,而边缘对三维物体的识别是至关重要的。换句话说,Gabor 滤波可以看作一个对方向和尺度敏感的有方向性的显微镜,能够检测(响应)图像中一些具有相应的方向频率信息的、局部的显著特征,从而可以形成亮度图像的局部特征图谱,这些局部特征形成了原始输入图像的一种鲁棒、紧凑的特征表示。

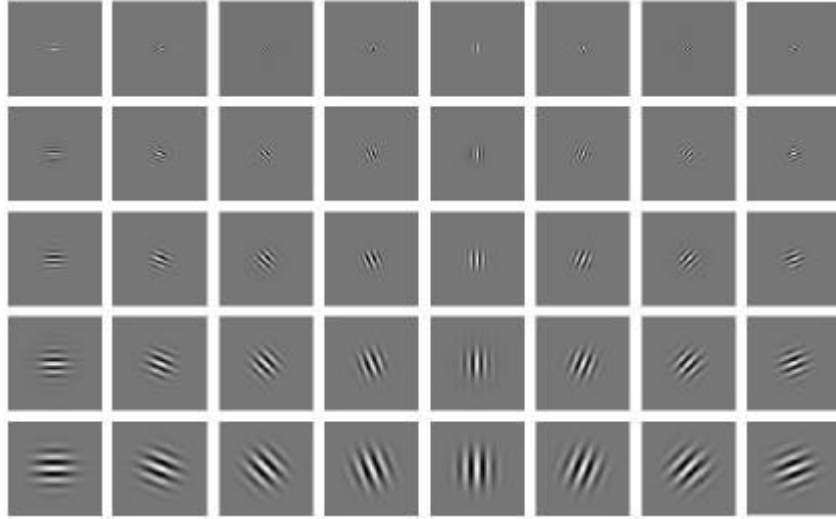


图 4.2 常用的 5 个尺度(自上而下 v 分别为 0 到 4)、8 个方向(自左而右 u 分别为 0 到 7) Gabor 核的实部图示

Fig. 4.2 Illustration of the real part of the Gabor kernels with 5 scales and 8 orientations

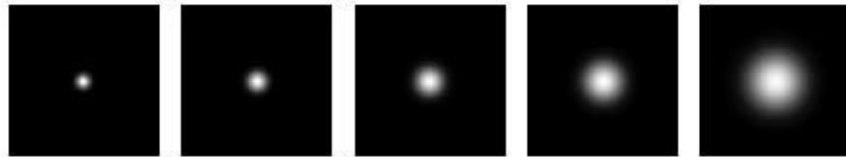


图 4.3 常用的 5 个尺度 Gabor 核函数的幅值图示(自左而右 v 分别为 0 到 4)

Fig. 4.3 Illustration of the magnitude of the Gabor kernels corresponding to 5 scales most commonly used

但是,Gabor 卷积过程实际产生由实部和虚部两个分量构成的复数响应,在边缘附近,Gabor 变换的实部和虚部会产生振荡,而不是一个平滑的峰值响应,因而不利于识别阶段的匹配。为此,一般的做法是退而求其次,抛弃 Gabor 变换本身具有的线性特征,而只保留 Gabor 响应的幅值(即实部和虚部平方和开根号),幅值信息实际反映了图像局部的能量谱,也可以理解为特定方向边缘的强度,而且在真实边缘附近具有良好的光滑性,有利于匹配识别。仅仅保留幅值信息其实摈弃了相位信息,相位信息是不稳定的,不利于识别,(但是可以用于精确的定位特征点的位置[WF97])。图 4.4 给出了对一幅人脸图像逐像素计算各个 Gabor 小波变换得到的 40 个不同的幅值图谱,其中(b)中所示变换后的图像分别对应于图 4.2 中相应位置的 Gabor 核函数与(a)中图像在相应位置进行卷积得到的幅值。

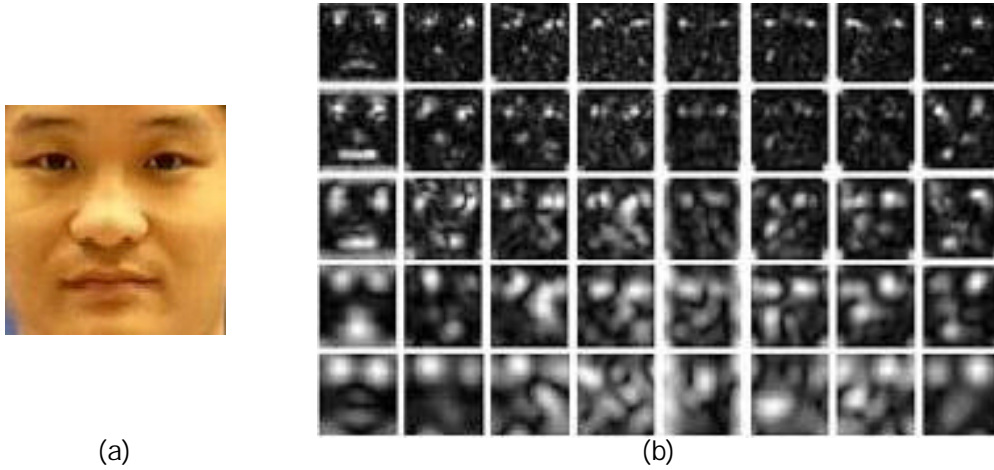


图 4.4 在人脸图像上逐像素计算得到的 40 个 Gabor 小波变换幅值图谱
(a)输入人脸图像 (b)40 个 Gabor 小波变换幅值图谱

Fig. 4.4 Gabor wavelet representation of face images computed pixel by pixel
(a)The input face image (b) Gabor wavelet representation of the a image.

通过上述方法在每个图像位置计算得到的40个幅值特征反映了以该图像位置为中心的局部区域频域内的能量分布特征，这40个幅值特征级联起来通常被称为一个 Jet [WF97]，简称为 J ，即图像位置 $p(x,y)$ 处的 Jet 为（参见公式4-2）：

$$J_{p(x,y)} = (J_{00,p(x,y)}, J_{10,p(x,y)}, \dots, J_{40,p(x,y)}, J_{01,p(x,y)}, J_{11,p(x,y)}, \dots, J_{47,p(x,y)}) \quad (4-5)$$

进一步将所有像素位置的 Jet 级联起来，就得到了输入图像 I 的 Gabor 特征表示：

$$\mathfrak{R}_I = \{J_{p(x,y)} \mid (x,y) \in \Omega_I\} \quad (4-6)$$

显然， \mathfrak{R}_I 这样的一种图像表示的特征维数是原始图像维数的 40 倍，如果原始输入图像为 64×64 大小，得到的 Gabor 特征维数是 163,840 维，直接对这样一个高维的特征向量进行分类识别将会非常困难。而这样的高维数据显然是有冗余的，而且这些特征对分类识别的贡献也是不同的，因此对其进行适当的降维是必不可少的。弹性图匹配技术采用的方法是预先定义一些面部关键特征点位置，每个特征点都有其高层语义[WF97]。这样的一种方法的优点是表示特征表示更紧凑，特征维数相对较低。缺点是（1）预定义特征点位置只能根据经验选取；（2）预定义特征点的自动定位复杂度高，而且也未必能够做到精确，从而影响最终的识别性能。Liu[LW02]提出的 GFC 方法则首先对 \mathfrak{R}_I 进行均匀下采样，然后进行主成分分析和线性判别分析。

为避免关键特征点预先定义的经验性及其它们的精确定位要求带来的复杂匹配问题，本文在对准双眼位置并对人脸图像进行尺度归一化的基础上，采用 AdaBoost 从公式 4-6 所定义的全部 Gabor 特征集中选择最优的 AdaGabor 特征，并进一步对选择出的 AdaGabor 特征进行子空间判别分析，以期在更低的时空复杂度的前提下，能够达到更高的识别性能。

4.2.2 AdaBoost 分类方法简介

AdaBoost(Adaptive Boosting 的缩写)是 Boosting 方法的一个典型特例。就 Boosting

本身而言,其目标是为了提高任何给定的学习算法的分类准确率。通常的做法是首先根据已有的训练样本集合训练出来一个分类器,要求这个分类器的准确率要比平均性能好,然后顺序地加入多个分量分类器,最终形成一个组合的总体分类器。Boosting 方法的这种机制可以保证最终的分分类器在训练样本集合上有任意高的准确率[DH03]。

而在 AdaBoost 方法中,每个训练样本都被赋予了一个权值,表明它能否被当前弱分类器(简单而分类性能相对差的分类器,例如简单的感知机就可以作为一个弱分类器)正确分类。AdaBoost 方法通过不断地加入弱分类器,最终达到某个预定的足够小的错误率。在加入某个新的弱分类器的某次迭代中,如果一个训练样本已经被正确分类,其权值就被下调,否则,其权值就要被调高,以便使下一轮的弱分类器专注于这些被错误分类的样本。最终得到的分类器则是训练得到的所有弱分类器的一个线性组合。依据这一基本思想,Freund 和 Schapire 首先给出了 Adaboost 算法的具体形式[FS97, FS99]。而针对人脸检测这个具体的应用问题,Viola 等人提出了一个修改的 AdaBoost 算法[VJ01, VJo01],Viola 的算法将弱分类器与弱特征(大量可选特征中的某一个特征)等价起来,每个弱分类器仅通过一个特征实现分类,并通过 AdaBoost 算法组合弱分类器,从而同时实现了弱特征的选择与组合。针对人脸检测问题,Viola 等人定义了大量矩形特征(简化的 Harr 特征,可以从 20×20 的图像中提取数十万的特征供挑选和组合)。Viola 提出的这个 AdaBoost 算法已经成为人脸检测领域的经典成果得到了更多的继承和发展[ZZ02, LZ04]。Viola 等人还将该算法应用到了人体检测领域,同样取得了良好的效果。

Viola 等人也尝试了将同样的思路用于人脸识别,但实验效果却不明显,至少没有取得比传统方法更好的识别性能。究其原因,本文认为主要其原因不在 AdaBoost 分类策略上,而在于人脸识别需要区分的模式类别数目庞大,而且类别之间的相似度很大,简单 Harr 特征并不适合于人脸识别。本文注意到 Gabor 特征表示方法的有效性,因此采用了 Gabor 特征表示作为原始特征集,采用 AdaBoost 方法对它们进行选择,然后进行判别分析,取得了比传统方法更好的识别性能。

4.3 AdaGaDA 人脸识别方法

AdaBoost 在人脸检测上的巨大成功促使我们考虑如何将其应用到人脸识别中去。注意到 Gabor 特征表示方法的有效性,我们采用了 Gabor 特征表示作为原始特征集,借鉴作差法形成两个“差空间”：“类内差”空间和“类间差”空间[MP00],从而将多类问题转化为 AdaBoost 可以解决的两类问题,然后采用 AdaBoost 方法对高维的 Gabor 特征进行特征选择(选择出的特征称为 AdaGabor 特征),然后对选择出来的 Gabor 特征进行判别分析从而实现最终的识别,取得了比传统方法更好的识别性能。方便起见,我们称该方法为 AdaGaDA 人脸识别方法,该算法的总体训练框架如图 4.5 所示。在图 4.5 中,为了尽可能的利用全部可能的 Gabor 特征,原始人脸图像在 Gabor 表示空间的特征 \mathcal{R} 是通过公式 4-10 计算出来的,没有经过下采样。

4.3.1 多类问题的二类化

AdaBoost 本质上是一个二类分类器学习算法,用于人脸识别时,首先要解决的一个问题就是如何将多类问题转化为二类问题。解决该方法的方法有一对多法(one-to-rest)、一对一法(one-to-one)等,这两种方法对 N 类识别问题分别需要构造 N 和 $N \times (N-1)/2$ 个二分类器,非常复杂,而且为了增加一个新的类别,就需要重新构造分类器组合,很不实用。考虑到实用的人脸识别系统的分类器不能过分复杂,增加新的人脸类别也必须方

便快捷，本文借鉴了 Baback 提出的将多类问题转化为“类内差”空间和“类间差”空间的思想[MP00]，并在这两个空间内应用 AdaBoost 学习能够将二者最佳的分开那些 Gabor 特征。

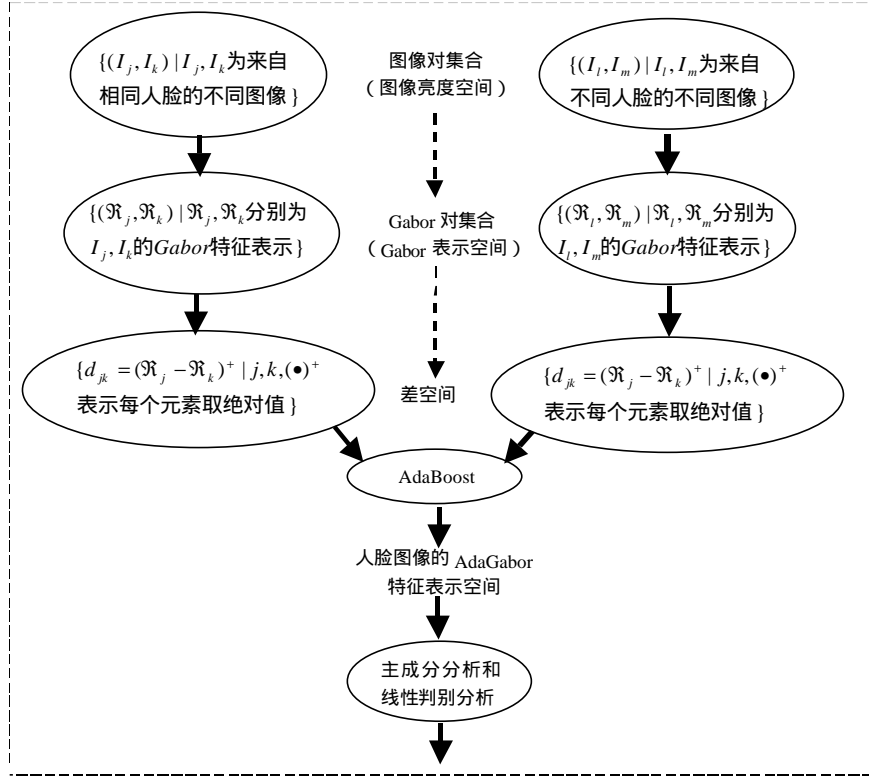


图 4.5 AdaGaDA 人脸识别方法总体框架

Fig. 4.5 Framework of the AdaGaDA face recognition method

在作差法中，为了保证“类内差”空间和“类间差”空间具有不同的中心，本文没有直接采用 Baback 所采用的向量差方式（ $x-y$ 和 $y-x$ 形成不同的差向量，因此类内差和类间差的中心都在原点 0 处），而是对 Gabor 特征表示采用了一种“街区差”。设任意两幅输入人脸 I_j, I_k ，它们的 Gabor 特征表示分别为 R_j, R_k ，则它们之间的街区差定义为：

$d_{jk} = (R_j - R_k)^{\parallel}$ ，其中 $(x)^{\parallel}$ 表示对 x 的所有元素取绝对值。在这样一种定义下，任意两个样本之间的“街区差”是与顺序无关的，这样可以确保“类内差”的中心更接近 0 向量，而类间差的中心则会偏离 0 向量。

4.3.2 基于 Boosting 技术的 AdaGabor 特征选择

应用图 4.5 所示框架训练 AdaBoost 分类器时必然碰到的一个关键问题是：“类内差”（定义为正例）样本集合与“类间差”（定义为反例）样本集合数据规模的不均衡问题。假设训练集中共有类别（人）数为 m ，每个类别有 k 个不同样本，则根据作差法，形成的正例样本集和反例样本集的数目将分别为： $N^+ = C_m^1 C_k^2 = mk(k-1)/2$ 和

$N^- = C_m^2 C_k^1 C_k^1 = k^2 m(m-1)/2$ ，这样，二者之比为： $R = N^- / N^+ = k(m-1)/(k-1)$ 。通常

情况下, m 会比较大, 而 k 会比较小, 比如 $m=500$, $k=5$, 则正反例样本集规模分别为 5,000 和 3,118,750, 也就是说它们之间在数量上差了 624 倍。这种不均衡的学习集非常不利于分类器的设计。为解决该问题, 从大规模的反例样本池中随机选取适量的反例样本参与训练是解决样本不平衡问题的一种自然的选择。但问题是随机选取实际上难以保证样本的代表性, 更难以保证得到的分类器对剩余样本的错误率控制在一定的水平上。为解决该问题, 结合随机采样方法, 本文进一步借鉴了 Boosting 的思路, 提出了一种基于反例样本集随机重采样的级联 AdaBoost 学习策略。该策略的工作流程如图 4.6 所示, 其细化的抽象描述见算法 4.1。

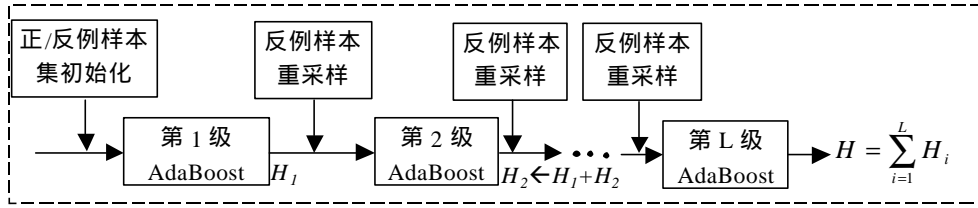


图 4.6. 基于反例样本集随机重采样的级联 AdaBoost 学习策略

Fig. 4.6. Cascade of AdaBoost based on re-sampling from the negative samples pool

算法 4.1 基于反例样本重采样的级联分类器学习算法

输入: $(x_1, y_1), \dots, (x_n, y_n)$ 是“类内差(正例)”“类间差(反例)”学习集,

$x_i \in X, y_i \in Y = \{0, 1\}$;

初始化: 所有的正例样本均被加入训练集中, 此后一直保持不变; 从反例样本池中随机选取反例样本, 数量保持为正例样本数目的固定倍数; 形成初始正反例样本集 S^+, S_0^-

For $l = 1, \dots, L$

Begin

(1) 初始化样本权重, 设: 正例样本和反例样本分别有 N^+, N^- 个

$$w_{1,i} = \begin{cases} \frac{1}{2N^+} & y_i = -1 \\ \frac{1}{2N^-} & y_i = 0 \end{cases} \quad (4-7)$$

(2) 调用 AdaBoost 学习算法, 形成当前层 AdaBoost 分类器, 并将其与前面已学习到的分类器融合为一个统一的 AdaBoost 分类器 H_l :

$$H_l(x) = \text{sign}\left(\sum_{t=1}^T \mathbf{a}_t h_t(x) - \frac{1}{2} \sum_{t=1}^T \mathbf{a}_t\right) \quad (4-8)$$

其中: $T = \sum_{j=1}^l T_j$, T_j 为第 j 层 AdaBoost 学习到的弱分类器数目(通过控制学

习精度来得到), $\mathbf{a}_t = \log \frac{1}{b_t}$ 为其所在层 AdaBoost 中相应弱分类器的权值。

(3) 反例样本重采样：从剩余的反例样本集中随机选取一个样本 x ，如果它被分类器：

$$H_l(\mathbf{x}_l; x) = \text{sign}(\sum_{i=1}^T \mathbf{a}_i h_i(x) - \frac{1}{2} \mathbf{x}_l \sum_{i=1}^T \mathbf{a}_i) \quad (4-9)$$

错分 (公式 4-9 所表示的分类器是公式 4-8 分类器的一个松弛版本，其中， $\mathbf{x}_l < 1$ ，且随着 l 的增大而增大)，则将其加入新的反例样本集中，否则就抛弃。

新的反例样本集大小保持为 N^- 。

End

H_L 为最终学习到的 AdaBoost 分类器，即：

$$H_L(x) = \text{sign}(\sum_{i=1}^T \mathbf{a}_i h_i(x) - \frac{1}{2} \sum_{i=1}^T \mathbf{a}_i) \quad (4-10)$$

其中： $T = \sum_{j=1}^L T_j$ ， T_j 为第 j 层 AdaBoost 学习到的弱分类器数目， $\mathbf{a}_i = \log \frac{1}{b_i}$ 为其所在层 AdaBoost 中相应弱分类器的权值；

为了加速人脸检测，Viola 采用了一种类似的 Cascade 策略。但本文的方法与之有很大不同，Viola 的 Cascade 策略是一种多层验证方法，每一层都会排斥掉大量反例样本，而正例样本必须得以逐级通过所有层次的验证，其主要目的是快速的抛弃大量的非人脸，从而加速人脸搜索过程。用于人脸识别时，不需要类似的过程，因此本文的方法没有在中间层排斥任何反例样本，而是正反例样本都要通过所有的弱分类器。本文的方法从重采样形成弱分类器的角度讲更接近标准的 Boosting 的做法。

尽管 AdaBoost 本身是一种通过自适应的组合弱分类器来得到增强分类器的方法，而不仅仅是一种特征选择的手段，但毋庸置疑，上述 AdaBoost 学习策略提供了一种有效的特征选择手段。也就是说，它可以从高维的 Gabor 特征表示中选择那些对分类最有效的特征用来识别。这本身就是一种非常有效的降维方法。为了叙述的方便，我们将上述 AdaBoost 学习过程选择出的 T 个 Gabor 特征定义为 AdaGabor 特征。该方法在 FERET 人脸库上进行实验，在弱分类器数目 (亦即弱特征数) $T=1000$ 维左右时就可以取得良好的识别效果。这说明原始的高维 Gabor 特征是高度冗余的，只需要其中的少部分就可以很好的区分不同人脸。

4.3.3 基于 AdaGabor 特征的子空间判别分析

尽管 AdaBoost 本身提供了一种通过组合这些弱特征进行分类识别的方法，但 AdaBoost 的这种组合方式并没有改变这些特征，也就是说，分类器仍然是在原始特征空间设计完成的。同时，AdaGabor 特征的维数相比原始 Gabor 特征维数已经有了很大的降低，但对于分类器设计而言，1000 多维的特征仍然是高维的，容易带来维数灾难问题，例如对于线性判别分析技术来说，1000 多维很容易带来奇异矩阵分解困难。因此，本文还进一步尝试了对 AdaGabor 特征进行最优的线性判别分析的方法：即首先采用主成分分析技术对 AdaGabor 特征进行降维，然后采用 Fisher 线性判别分析方法对其进行判别

分析。在模式识别中，Fisher 线性判别分析(FLD)广泛应用于特征提取和降维。FLD 力图找到一组最佳的投影方向，在这些投影方向上，可以最好的区分训练集中属于不同类别的样本。这种可分性，在数学上可以用不同的标准来衡量。其中比较常用的一种定义为：在投影后的训练集上，类间离散度矩阵的行列式和类内离散度矩阵的行列式的比值最大化。FLD 能够在增大类间离散度的同时减小样本的类内离散度，从而使得在投影空间中，不同类别的样本将会尽可能的分开，而属于同一类别的样本将会聚合在一起，因而更适于识别。其形式描述如下：

对于一个 c 类问题，类间和类内离散度矩阵分别定义为：

$$S_b = \sum_{i=1}^c \Pr(\Omega_i) (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \quad (4-11)$$

$$S_w = \sum_{i=1}^c (\Pr(\Omega_i) \times \frac{1}{N_i} \sum_{y_k \in \Omega_i} (y_k - \mathbf{m}_i)(y_k - \mathbf{m}_i)^T) \quad (4-12)$$

其中， $\Pr(\Omega_i)$ 是先验类概率密度， \mathbf{m}_i 是类 Ω_i 的样本均值， \mathbf{m} 是所有类的样本均值， N_i 是类 Ω_i 中训练样本的个数。

如果 S_w 是非奇异的，那么投影矩阵 W_{lda} 可以通过求解下述优化问题得到：

$$W_{fld} = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|} \quad (4-13)$$

W_{lda} 最终可以通过解如下广义特征值问题而得到：

$$S_b W = S_w W \Lambda \quad (4-14)$$

但是在人脸识别领域，类内离散度矩阵 S_w 在绝大多数情况下是奇异的。这是因为：

由 S_w 的定义可知， $S_w \in R^{n \times n}$ ， S_w 的秩最大为 $N - c$ （ N 为训练集中样本数目）， N 通常又远远小于每幅人脸图像的特征维数 n ，所以 $N - c \ll n$ ，因此 S_w 通常是奇异的。为了解决这个困难，人们进一步提出了增强法、扰动法、零空间法等不同的方法，而一般首先使用主成分分析(PCA)将原始特征的维数从 n 降低到 $N - c$ 或者更少。解决该问题的方法不同会导致最终系统的识别性能不同，但一般不会有本质的差别，本文仅讨论了首先采用 PCA 进行降维然后用 FLD 进行判别分析的策略。

4.4 实验与分析

4.4.1 测试数据库情况介绍

为了使得评测结果是统计显著的，我们选择了人脸图像规模在 1000 人以上的

FERET 人脸库和 CAS-PEAL 人脸库，下面简单介绍这两个人脸库的情况。

4.4.1.1 FERET 人脸库

FERET 人脸库是由美国 DARPA 资助的 FERET 计划构建[PM00]，其基本构成如表 4.1 所示。为了有效的验证算法的性能，本文按照 FERET 标准测试方法的规定将 Fisherface 方法的训练数据集合 L (training set) 与测试数据集合 T (testing set) 作了严格区分，其中后者包括已知人脸原型图像集合 G (gallery) 和探测图像集合 P (probe set)[PM00]。

表 4.1 FERET 测试标准人脸库集合情况表

Table. 4.1 Face database structure of the standard FERET testing protocol

数据库		人数	图像数	特征描述
训练集 L (Training set)		429	1002	部分人的部分图像来自参考集 G 和测试集 P
测试集 T	参考集 G (Gallery)	1196	1196	准正面，均匀光照条件
	探测集 P (Probes)	FB	1195	同一天采集，准正面，均匀光照，表情不同
		fc	194	同一天采集，准正面，不同的摄像头
		Duplicate I	243	1031 天内不定期采集，准正面，有光照变化
		Duplicate II	75	Duplicate I 的子集，至少 12 个月之后采集

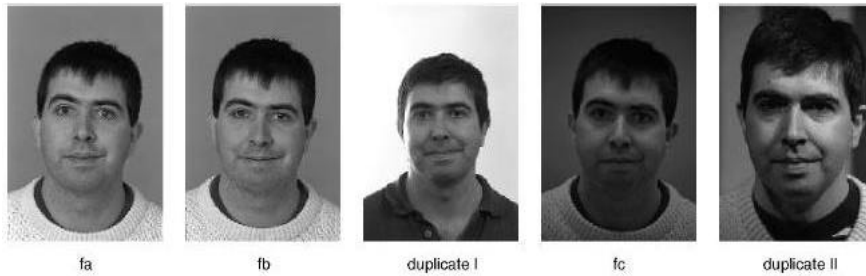


图 4.7 FERET 人脸图像库示例图像

Fig. 4.7 Example face images in FERET face database.

FERET 人脸库中典型的示例人脸图像如图 4.7 所示。FERET 提供了所有这些人脸图像的精确眼睛位置。

4.4.1.2 CAS-PEAL 人脸库

CAS-PEAL 共享人脸数据库包含 1,040 人的 30,900 幅图像，分为两个子库：正面图像子库和姿态图像子库。本文仅使用了其中的正面图像子库，包含 1,040 人的共 9,060 幅图像，其中涉及的不同成像条件图像的分布情况如表 4.2 所示，示例图像见图 4.8。

CAS-PEAL 第一版的共享人脸库中提供了三个标准的图像集用以测试不同识别算法的性能，即：训练集合 (Training Set)、原型图像集合 (Gallery) 和测试集合 (Probe Sets)。CAS-PEAL 训练集合 (Training Set) 供算法训练识别模型和参数用，它是从正面子库中随机选择 300 人、每人随机选择 4 幅正面图像，共 1,200 幅图像构成的。原型图像集合 (Gallery) 是提供给算法为待识别人建立原型的图像子集，包含 1,040 人，每人 1 幅标准正面图像(正面平视，环境光照，中性表情，无饰物)。测试集合 (Probe Set) 是提供给算法进行识别的图像子集。对正面子库，测试集包括除去训练集和原型集以外余下的 6,992 幅图像，分为表情变化、光照变化、饰物变化、背景变化、距离变化和时

间跨度变化 6 个子集。

表 4.2 CAS-PEAL-R1 的正面图像子库图像情况

Table 4.2 Frontal sub-database of CAS-PEAL face database

		变化种类	人数	图像数目
正面子库	标准	1	1,040	1,040
	表情	5*	377	1,884
	光照	≥ 9	233	2,450
	饰物	6	438	2,646
	背景	2-4	297	650
	距离	1-2	296	324
	时间	1	66	66
正面子库图像总数:			9,060	

* 中性表情的标准正面图像没有包括在内

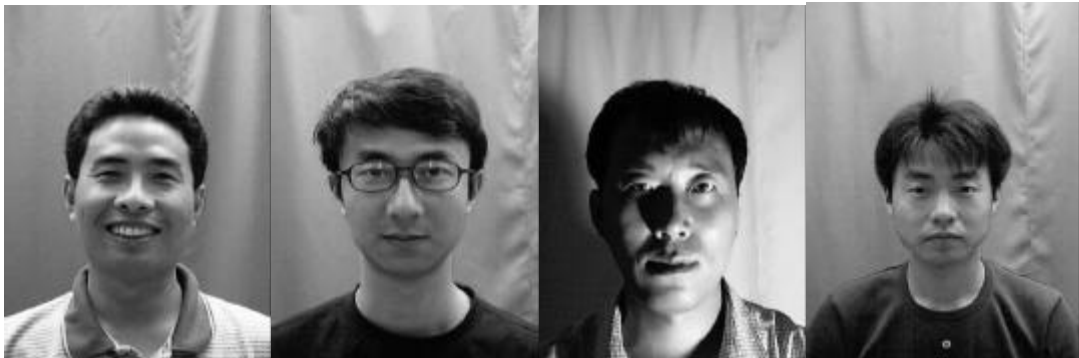


图 4.8 CAS-PEAL-R1 正面子库中的图像示例

Fig. 4.8 Example face images in the CAS-PEAL face database

4.4.1.3 图像归一化与预处理

为了进一步消除背景、头发等非人脸本质因素对识别性能的影响，根据给定的双眼位置，本文采用如下图 4.9 所示的方式对人脸图像进行几何归一化，其中，图 4.9(a)中的裁剪参数 h , t , b 分别为 0.64, 0.43 和 1.85，最终所有的图像归一化为 64×64 像素。为了在一定程度上消除不同的光照条件的影响，所有图像均进行了直方图均衡化，并进一步对裁剪出的人脸区域叠加掩模(Mask)以便减小背景和发型变化对人脸识别的不利影响。本文以下讨论所涉及的算法性能均是在该尺寸图像上的测试情况。



图 4.9 本文实验中采用的人脸图像归一化方法

Fig. 4.9 Face normalization method used in our experiments

4.4.2 AdaGabor 特征的统计特性分析

AdaGabor 特征应该是最有利于识别的特征。为了考察这些特征的特性，在 FERET 人脸库上进行实验，采用最简单的线性感知机作为 AdaBoost 中的弱分类器，共选出了 $T=1000$ 维 AdaGabor 特征，本文对这些特征从以下几个方面进行了实验和统计分析：

1、判别能力最强的 AdaGabor 特征情况

图 4.10 给出了选择出的头四个最强 AdaGabor 特征的情况，它们所对应的 Gabor 核函数被叠加到了一幅人脸图像上，从中可以看出它们的位置、尺度和方向参数的情况。不难发现，这四个最强“弱特征”分别提取了两个眼睛、鼻子和嘴巴处的判别特征。这与直观的印象也是基本一致的。需要注意的是，眼睛位置的两个判别特征的中心位置并没有落在眼睛中心，这主要是因为本文采用的归一化方法是对齐所有图像的眼睛中心，从而导致眼睛中心区域之间的差别比较小造成的。



图 4.10 通过 AdaBoost 选择出的头四个最强“弱特征”

Fig.4.10 The first four Ada-Boosted Gabor features

2、AdaGabor 特征中 40 种特征的频数统计

不同类型的 Gabor 特征对于识别的有效性是不同的。图 4.11 给出了前 1000 维 AdaGabor 特征中 40 种不同的 Gabor 滤波器出现的频度情况（图中横坐标标示的 40 种 Gabor 滤波器类型分别对应本文图 4.2 从左到右、自上而下进行编号的 Gabor 核）。从中可以看出：（1）不同的 Gabor 滤波器类型对于判别的贡献率差别是比较大的；（2）对判别贡献率最大的 Gabor 滤波器类型是 1 号（水平方向 $u=0$ ，尺度为 $v=0$ ）、5 号（竖直方向 $u=4$ ，尺度为 $v=0$ ）和 13 号（竖直方向 $u=4$ ，尺度为 $v=1$ ）。

AdaBoost 选择出的 40 种 Gabor 特征分布情况

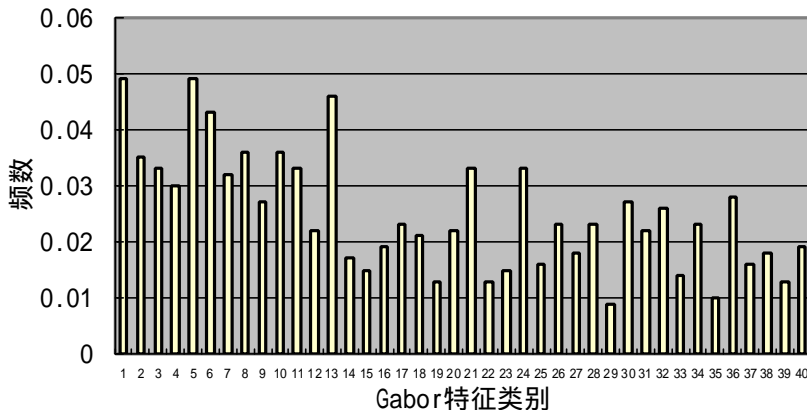


图 4.11 前 1000 维 AdaGabor 特征中 40 种不同的 Gabor 特征频度情况

Fig.4.11 The distribution of the learned leading 1000 AdaGabor features

3、AdaGabor 特征的尺度分布统计

不同尺度的 Gabor 特征对于识别的贡献也是有较大不同的。图 4.12 给出了本文的实验中选择出的前 100、500 和 1000 维 AdaGabor 特征中不同尺度的 Gabor 特征所占比例的分布和变化情况。从中不难看出较小尺度的 Gabor 特征对于识别是更有利的，尤其是当需要区分人脸之间的细微差别时，这种趋势更加明显（前 1000 维中后来选择出来的特征更多的是小尺度特征）。这一点也是与直观的感觉相吻合的。

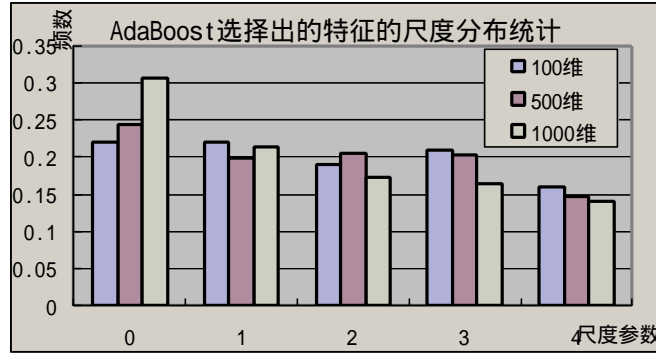


图 4.12 Ada Gabor 特征中不同尺度的分布和变化情况

Fig.4.12 The scale distribution and variation of the leading AdaGabor features

4、AdaGabor 特征的方向分布特征

类似的，不同方向的 Gabor 特征也具有不同的贡献。图 4.13 给出了本文的实验中选择出的前 100、500 和 1000 维 AdaGabor 特征中不同方向的 Gabor 特征所占比例的分布和变化情况。尽管其分布比较平坦，但从中仍然不难看出：竖直方向（ $v=4$ ）的 Gabor 特征判别能力最强，能够有效的区分大多数的人脸差别，而区分细微差别时，其他各种不同方向的特征的作用都增加。但总体而言，45 度方向（ $v=2$ 和 $v=6$ ）的 Gabor 特征的判别能力最弱。

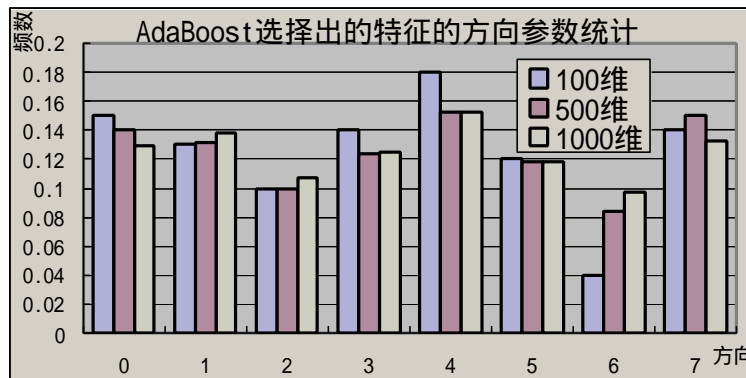


图 4.13 AdaGabor 特征中不同方向的分布和变化情况

Fig.4.13 The orientation distribution and variation of the leading AdaGabor features

4.4.3 对比算法设计

为了对比本文算法与其他算法性能上的差别，我们同时实现了 Fisherface 方法[BH97] 和 GFC 方法[LW02]。其中，GFC 算法实现的是一个基于 Gabor 特征均匀下采样，然后进行主成分分析和 Fisher 判别分析的方法。下采样的比例为 4×4 ，亦即每个 4×4 小栅

格内的所有 Jet 计算均值，从而下采样后 Gabor 特征的维数为 $15 \times 15 \times 40 = 9,000$ 维。与 GFC 算法不同的是我们采用的是 PCA 降维策略而没有采用 Liu 方法中的增强判别分析方法。

4.4.4 识别性能对比分析

采用本章论述的 AdaGaDA 方法，分别在 FERET 和 CAS-PEAL 人脸库上进行了实验，通过选择得到的 AdaGabor 特征在两个人脸库上分别保留了 1884 维和 2500 维。表 4.3 和表 4.4 分别给出了 AdaGaDA 方法与 Fisherface、GFC 方法在这两个人脸库上的首选识别率的对比情况，需要注意的是，公平起见，表中列出的 PCA 特征维数 D_{pca} 是相应方法取得最佳识别率时候的维数，而 FDA 特征维数 D_{lda} 是不同子集最佳维数的均值。

表 4.3 AdaGaDA 与 Fisherface、GFC 方法在 FERET 人脸库上的性能对比

Fig. 4.3 Performance comparisons of AdaGaDA, Fisherface and GFC on FERET

训练集合	FERET Training CD			测试集合		FERET 人脸库	
识别方法	参数情况			测试子集合（ 首选识别率 ）			
	D _{ori}	D _{pca}	D _{lda}	FB	fc	DupI	DupII
Fisherface	4096	300	210	94.4%	65.5%	59.3%	36.3%
GFC	9000	500	250	96.3%	81.4%	68.8%	49.1%
AdaGaDA	1884	250	200	97.2%	77.8%	64.5%	41.0%

表 4.4 AdaGaDA 与 Fisherface、GFC 方法在 CAS-PEAL 人脸库上的性能对比

Fig. 4.4 Performance comparisons of AdaGaDA, Fisherface and GFC on CAS-PEAL

训练集合	CAS-PEAL 训练集			测试集合			CAS-PEAL 测试库		
识别方法	参数情况			测试子集合（首选识别率）					
	D_{ori}	D_{pca}	D_{lda}	表情	饰物	光照	背景	老化	距离
Fisherface	4096	400	210	80.2%	71.0%	28.8%	97.5%	77.3%	97.5%
GFC	9000	500	250	92.9%	85.1%	44.3%	98.9%	93.9%	100%
AdaGaDA	2500	250	200	98.2%	87.5%	29.6%	99.6%	97.0%	99.3%

由表 4.3 可以看出，AdaGaDA 方法在 FERET 四个标准测试集合上的识别性能要明显优于 Fisherface 方法。而在原始特征维数远低于 GFC 的 9000 维的情况下，其在 FERET 人脸库四个标准测试集合上的识别性能与 GFC 是可比的，在 FB 测试库上其首选识别率则略高于 GFC 算法。需要指出的是：FERET 人脸库的标准测试集中，FB 和 DupI 的图像规模是较大的，而 fc 和 DupII 两个测试集合规模较小，因此这种识别率的差别并非是统计显著的，例如对 fc 测试集合而言，尽管 AdaGaDA 方法的识别率低于 GFC，但其实际识别错误的图像也不过比 GFC 多 8 幅。

而通过表 4.4 不难看出，AdaGaDA 方法在 CAS-PEAL 人脸库六个测试子集上的识别性能均较明显的优于 Fisherface 方法。而与 GFC 方法相比，在原始特征维数大大低于 GFC 方法的情况下，除了光照子集外，AdaGaDA 方法在其他几个测试集合上的识别性能要优于 GFC 方法。尤其是表情和饰物子集，AdaGaDA 方法的识别率都比较明显的高于 Fisherface 和 GFC 方法。

值得特别说明的是：与GFC方法相比，AdaGaDA方法在CAS-PEAL光照子集上的性能较差，这并不能说明方法对光照变化不鲁棒。由于AdaGaDA方法采用AdaBoost学习策略学习AdaGabor特征，因此，选择出来的特征必然在一定程度上依赖于训练集中图像的分布情况。而CAS-PEAL人脸库的训练集中缺少光照变化图像，在CAS-PEAL全部300人的共1200幅训练图像中，只有200幅包含光照条件变化的图像。因此，选择出来的AdaGabor特征没有能够良好的覆盖不同的光照变化导致的类内变化，并导致在光照子集上的识别性能不佳。如果能够构建一个包含更多光照条件变化的训练图像库，通过AdaBoost选择出来的AdaGabor特征应该可以更好的适应光照条件的变化，从而使得AdaGaDA方法的光照鲁棒性得到提高。

以上对比分析表明，AdaGaDA方法的识别性能总体上要明显优于Fisherface方法，而与GFC方法相比，在FERET和CAS-PEAL共十个测试子集的性能比较中，二者的识别性能难分伯仲，但由于AdaGaDA方法的特征维数远低于GFC方法，因此其特征提取的时间低于GFC方法，从而可以加速识别系统的识别速度。在实际应用系统设计中，由于AdaGabor特征的维数大大低于GFC的特征维数，这使得采用更大的学习集用来进行判别分析或者其他的分类器设计更加方便，从而使得最终的识别系统的性能优于GFC方法。

4.5 本章小结

本章研究了高维 Gabor 特征的有效降维问题，首次成功地将 AdaBoost 应用到了人脸识别中，提出了一种 AdaGaDA 人脸识别方法。该方法将 AdaBoost 算法应用于 Gabor 特征选择中，以提取对识别最有利的 AdaGabor 特征，并最终通过对 AdaGabor 特征的判别分析实现识别。AdaBoost 本质上是一个两类分类器学习算法，为了将其应用于多类人脸识别问题，本文借鉴了类内差和类间差的思想，将多类问题转化为两类问题，从而可以直接应用 AdaBoost 选择那些对识别有利的 Gabor 特征用于识别。在 AdaBoost 学习算法设计方面，考虑到类内差样本集和类间差样本集在规模上的严重不均衡，本文提出了一种基于反例样本重采样的级联式反例样本重采样策略，从而保证了 AdaBoost 过程能够学习到足够多的 AdaGabor 特征。在 CAS-PEAL 和 FERET 人脸图像库上的对比实验表明：AdaGaDA 方法不但可以大大降低 Gabor 特征的维数，而且识别精度也比传统方法有了较大的提高。

第五章 人脸识别中的“误配准灾难”问题研究

现有的多数人脸识别系统都要依赖于面部特征（比如眼睛位置）的严格配准来归一化人脸以便提取人脸描述特征，但面部特征配准的准确度如何影响人脸识别算法的性能却没有得到足够的重视。本章首次针对这一问题进行了系统的研究，发现 Fisherface 的识别率会随着误配准的增大而急剧下降——本文称这一现象为“误配准灾难”问题。在分析“误配准灾难”问题产生原因的基础上，本章探讨了其可能的四类解决方案，即：精确再配准方法、不变特征方法、误配准学习方法和误配准不变距离度量方法。为了能够量化比较不同算法对误配准的鲁棒性，本章提出了一套考虑了误配准概率分布的识别性能评估方法和误配准鲁棒性度量策略。针对 Fisherface 的“误配准灾难”问题，本章提出了一种对误配准进行强化训练的 E-Fisherface 方法，该方法通过在模型训练阶段加入人为配准偏差来提高判别分析方法对误配准的鲁棒性。接下来，本章还进一步考察了上一章研究的 Gabor 特征人脸识别方法对误配准的鲁棒性。上述方法在 FERET 和 CAS-PEAL 人脸图像数据库上的对比实验表明：Gabor 特征比单纯的灰度特征具有更强的鲁棒性；对误配准进行强化训练的 E-Fisherface 方法同样可以大大提高识别系统的鲁棒性。

5.1 引言

对实用的人脸识别系统而言，人脸检测、面部特征配准和人脸建模与比对是三个不可或缺的步骤。针对这三个相对独立的问题，研究人员已经开展了大量的研究。但这三个步骤之间的无缝集成问题却没有得到足够的重视。特别是研究人脸建模与比对问题的研究人员，他们的实验中往往假定面部特征是精确配准的（实验中采用手工标定的双眼位置）。问题是：面部特征配准是否已经非常好地解决了？答案是否定的，尤其是对用户不配合的非限定条件应用系统中，自动提取的面部特征位置往往会有很大的偏差。另一方面，对研究面部特征提取的研究人员而言，对一般大小（比如 92×112 像素）的人脸而言，面部特征点的误配准（如眼睛中心位置）可能高达 5 个像素——而这往往都被认为是“正确”的。更为严重的是：在输入人脸图像质量较差时，面部特征点精确位置的定义往往是歧义的，图 5.1 以眼睛中心为例说明了这一问题，在瞳孔中心与眼睛中心不重合、图像模糊、眼镜反光或者遮挡的情况下，这一问题更为突出，即使是人也难以准确判断所谓的眼睛位置，更不用说自动定位算法了。

上述分析表明：一方面需要进一步研究更加精确的面部特征自动定位算法，另一方面，人脸建模与比对方法不应该过分要求面部特征配准的精度，而必须对误配准具有较强的鲁棒性，方能实现稳定可靠的人脸识别。实际上，在模式识别领域内，对不改变模式类别属性的各种变换（如平移等相似变换）的不变性研究，很早就引起了人们的重视，尤其在 OCR 领域，采用切距离（Tangent Distance）度量[SC98]可以提高识别性能已经成为共识，相关的文献较多[KM04]，而在人脸识别领域，该问题尚未得到足够的重视，更没有得到系统的量化研究。即使在 OCR 文献中，也没有给出识别算法性能随着误配准程度不同而变化的曲线（参见图 5.3），更没有给出在考虑误配准鲁棒性的条件下如何评价不同算法的性能并进行量化的比较，本章针对人脸识别中双眼特征的误配准问题进行实验分析和理论研究，揭示了眼睛位置的准确度如何影响识别算法性能，给出了识别

算法对误配准鲁棒性的度量方法，研究了如何提高识别算法对误配准鲁棒性的问题，并提出了两种不同的解决方案，以实现更加鲁棒实用的人脸识别系统。

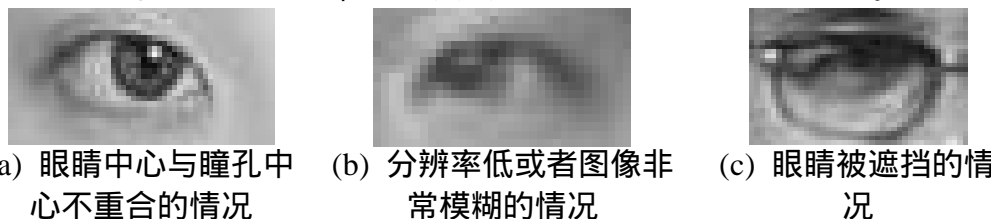


图 5.1 某些特征点的精确位置本身可能是歧义的，以眼睛中心的定义为例

Fig. 5.1 Correct alignment for some landmarks is essentially ambiguous. Taking eye center as example.

5.2 “误配准灾难”问题的提出

Fisherface 是最重要的人脸识别方法之一[BH97]，实验也表明其在眼睛位置精确定位的条件下具有良好的识别效果。然而，基于 Fisherface 方法的人脸识别应用系统的实际效果却往往会出人意料的差。为了分析这一现象，我们通过记录错误识别案例的实验方法收集了大量的错误样本进行分析，结果表明：大多数被错误识别的样本完全可以通过精确调整眼睛的位置得到正确的识别结果，也就是说，识别性能的下降很大程度上是由于自动定位的眼睛位置不够准确造成的。这使得我们在进一步研究眼睛的精确定位算法的同时，开始认真对待识别算法本身对于眼睛位置定位误差的鲁棒性问题，考虑如何降低识别算法本身对于特征误配准的敏感性以弥补特征定位不准确带来的系统性能下降。为了系统的分析误配准对识别性能的影响，本章设计了评测实验来显式的揭示该问题的严重性[SC04]。

5.2.1 实验设计——Fisherface 在 FERET 人脸库上的测试

本章采用的 Fisherface 方法是 Belhumeur[BH97]提出的方法的一种实现，具体方案是首先采用主成分分析（PCA）进行降维，然后采用线性判别分析（LDA）对降维后的特征进行判别分析，最后采用最近邻方法进行分类识别。为了使得评测结果是统计显著的，我们选择了人脸图像规模在 1000 人以上的人脸库，该人脸库有关信息参见[PM00]及本论文 4.4.1.1 小节介绍，方便起见，本文后面的实验只采用了 FERET 中最大的标准测试集 FB。为了有效的验证算法的性能，实验中严格按照 FERET 标准测试方法[PM00]的规定将 Fisherface 方法的训练数据集 L 与测试数据集 T 作了严格区分，其中 PCA 和 LDA 的变换矩阵 W_{pca} 和 W_{lda} 都是在训练集合 L 上学习得到，然后对 G 和 P 中的所有人脸图像通过降维变换到 LDA 特征空间 P 中任意图像的识别则通过在 LDA 特征空间中与 G 中所有已知样本比对来完成。

在实验中所采用的几何归一化和灰度归一化方法同本论文 4.4.1.3 小节的介绍，在归一化的 64×64 像素人脸图像中，双眼距离 D 为 28 像素。需要指出的是，本章以下讨论所涉及的误配准均指在该尺寸大小图像上的误差情况。

为了系统而定量地评估识别方法对于误配准的敏感性，本文采用对探测图像(Probe)眼睛位置人为加入误配准的实验方法。容易理解，在不考虑 3D 姿态变化导致的误配准的条件下，实际上眼睛位置偏差导致的误配准最终反映为归一化后的人脸图像与真实图

像之间的平移 (T_x, T_y)、尺度 (s) 和旋转 (q) 等相似变换参数的变化。为简明起见, 本文仅针对这些参数分别进行实验。图 5.2 给出了这些人为误配准对归一化人脸图像模板的影响情况, 可以看出这些误配准在较大程度上影响了人脸图像的表现(Apppearance)。

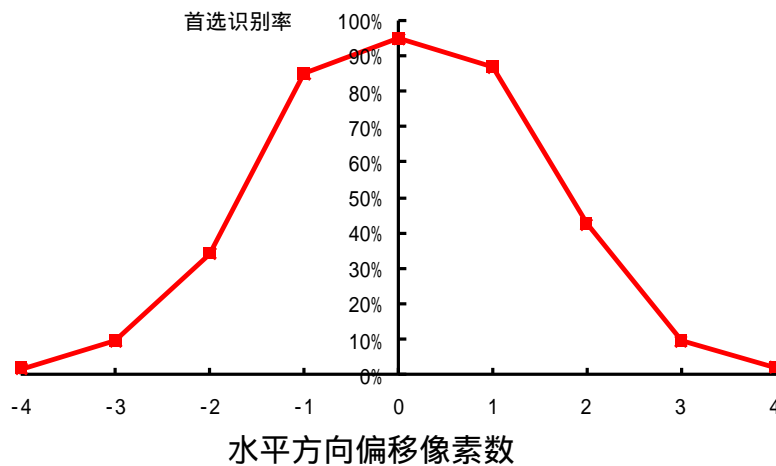


图 5.2 误配准对归一化人脸图像模板的影响示例

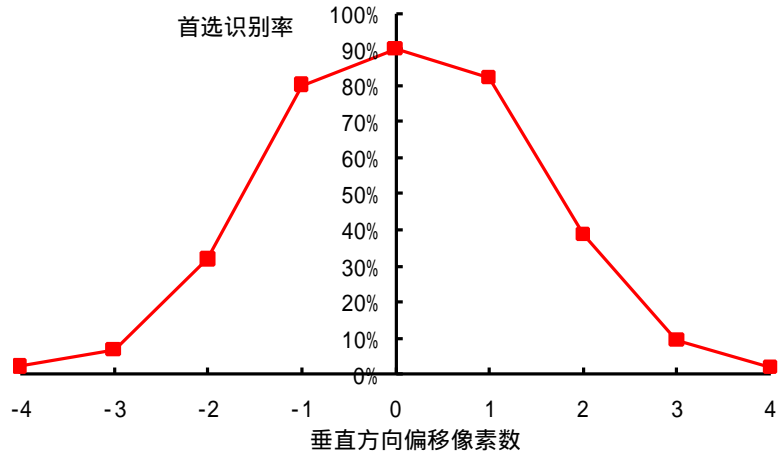
Fig. 5.2 Normalization error due to mis-alignment

5.2.2 “误配准灾难”问题

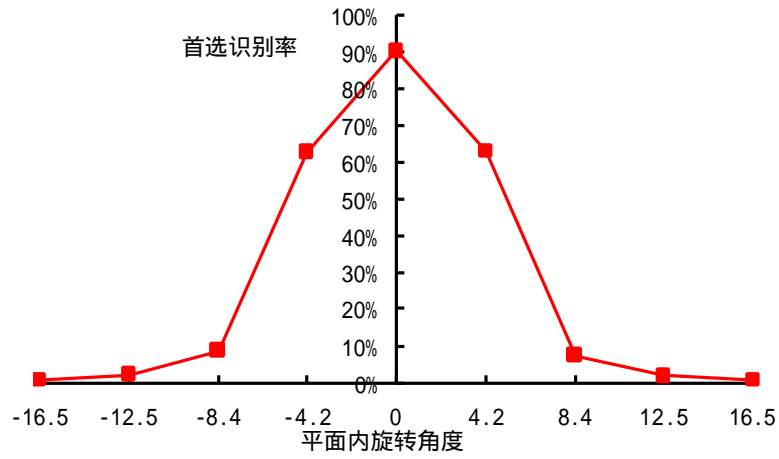
采用上节所描述的测试方法, 对 Fisherface 方法在 FERET-FB 探测集上测试由于误配准而导致一定平移、旋转和尺度变化时其首选识别率的变化情况, 结果如图 5.3 所示。其中, 图 5.3(a)是双眼位置在水平方向向相同方向平移若干像素的情况; 图 5.3 (b)是双眼位置在竖直方向向相同方向平移若干像素的情况; 图 5.3 (c)为双眼位置竖直方向错位从而导致一定角度平面内旋转时识别率的变化情况; 图 5.3 (d)为双眼位置在水平方向向相反方向平移导致尺度变化时的情况。不难看出, Fisherface 的识别率会随着误配准的增大而急剧下降, 即使在只有一个像素平移偏差的情况下, 识别性能也下降了 10 个百分点; 而平面内旋转 4.2 度 (左、右眼在竖直方向朝相反方向各移动 1 个像素) 时, 识别性能则下降了 20 多个百分点; 尺度发生 0.07 的变化 (左、右眼在水平方向朝相反方向各移动 1 个像素) 时, 首选识别率陡然下降了 30 个百分点。显然, 在发生如此小的误配准时, 如此剧烈的性能下降是实用的人脸识别系统根本不能接受的, 尤其是在全自动的人脸识别系统中, 面部特征定位算法往往很难达到百分之百的准确度, 偏差一两个像素往往是非常正常的, 这就意味着实用识别系统的识别性能会很不稳定。基于此, 本文把人脸识别算法的识别性能随误配准增大而急剧下降的现象称为“误配准灾难”, 以强调该问题的重要性。



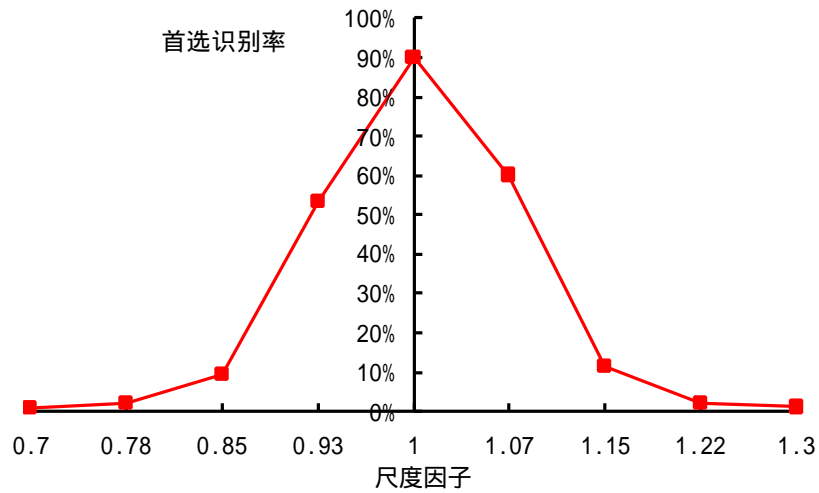
(a)



(b)



(c)



(d)

图 5.3 在 FERET-FB 测试集上，Fisherface 方法的识别率与误配准的关系
 Fig. 5.3 Relationship between the recognition rate of the Fisherface and the mis-alignment

5.3 “误配准灾难”产生原因分析及解决方案归类

特征配准的目的本质上是要建立两幅图像各像素在“高层语义”上的对应关系，亦即使得待匹配的两个像素具有相同的高层语义。误配准就意味着图像之间像素高层语义不能严格相同，这对基于表观的识别方法来说，显然会造成“牛唇不对马嘴”的错误匹配。图 5.4(a-c)以一个极端的情形作例子，生动的阐述了“误配准灾难”问题产生的原因，假设图中实线矩形为原始图像模式（二值边缘图像），虚线为由于误配准而产生的图像模式，很显然，误配准将使得待匹配的两个像素具有完全不同的“高层语义”，如果采用基于表观的方法进行二者的匹配，对应像素灰度值之间的相似度根本上就不能反映输入模式之间的真实相似度，因而也就不可能做到良好的分类。对于基于表观匹配的人脸识别方法而言，情形尽管没有如此直观和极端，但也可以通过图 5.4(d-f)窥见一斑，其中 d, e, f 分别是图像发生平移、旋转和尺度变化后图像作差（取绝对值）的结果。不难看出，误配准的出现，导致了很大程度的无效匹配，因而计算出来的距离或者相似度就很难反应实际的人脸之间的相似度，并最终导致了识别性能的急剧下降。

从分类的角度讲，误配准显然增大了归一化的图像空间中各人脸类别的类内散度，也在一定程度上减小了不同人脸之间的类间散度，从而导致分类困难。特别的，对 Fisherface 方法而言，由于训练过程并没有考虑对误配准的学习，因而 Fisherface 方法得到的判别向量只在精确配准条件下满足类间散度与类内散度比值的最小化，而不能适应误配准条件，因而导致了识别性能的急剧退化。

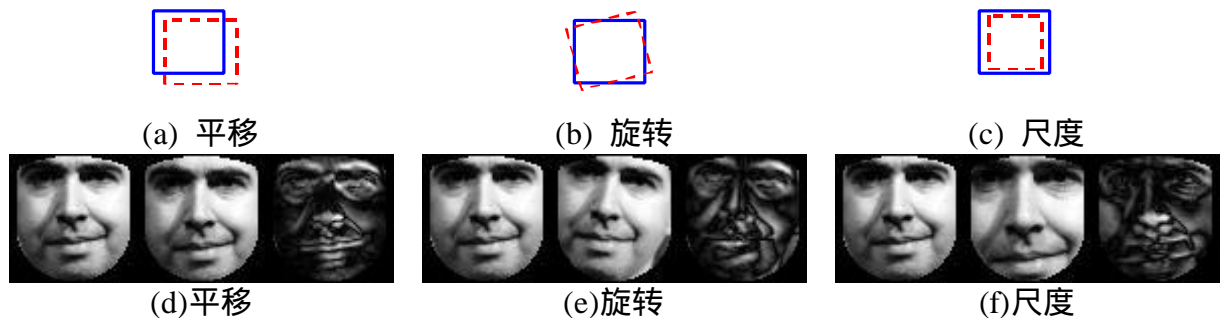


图 5.4 导致“误配准灾难”产生的根源：无效匹配图示

Fig. 5.4 Root of “Curse of mis-alignment”: meaningless pixels matching

即使是微小的特征配准误差都引起了人脸图像表观模式的非线性变化，从而导致了模式类内变化不够紧凑，类间模式重叠增多，因而带来了分类困难。参考光学字符识别（OCR）等传统模式识别领域对变换不变分类问题的研究，本文将“误配准灾难”问题的解决方案划分为四大类，即：精确再配准方法、误配准不变的特征方法、配准误差学习方法、变换不变距离度量方法。当然，这四类方法实际上还可以进行结合，得到对误配准具有更强的鲁棒性的识别方法。这四类方法的基本思路及其相关的前人工作分别简述如下。

5.3.1 精确再配准方法

这类方法通常在特征配准算法的基础上，利用从分类器反馈回来的匹配信度或识别结果的信度进一步调整特征的配准，以便得到更高的匹配/识别信度，最终完成更加鲁棒的识别。例如，[廖 03]采用了一种基于遗传算法的邻域搜索策略并以识别距离为目标函

数对眼睛的中心位置进行进一步的优化，取得了较好的效果。这类方法由于引入反馈导致了算法时间复杂度的升高。

5.3.2 误配准不变的鲁棒特征方法

这类方法试图从输入人脸图像模式中提取对误配准不敏感的人脸描述特征，即提取的特征不会随误配准的增大而发生大的漂移，因而模式类别的表达更加紧凑(compact)，从而使得后续的分类算法能够容易地实现鲁棒的识别。Gabor 小波特征使具有方向和尺度选择性的局部特征，也已经被证明具有更佳的误配准鲁棒性能[BL90, LV93, WF97, Wur97, SG04]。弹性图匹配技术[LV93, WF97]则在采用 Gabor 特征方面更进了一步，通过图匹配技术在精确定关键特征点的基础上进一步增强识别算法对配准误差的鲁棒性。Lampinen[LO95]则采用多层的自组织映射神经网络进一步分类 Gabor 特征。高斯函数导数(Gaussian Derivatives)特征具有与 Gabor 特征类似的特性，同样得到了在变换不变的模式识别中得到了应用[SC00]。另外一种常被采用的是所谓的高阶自相关特征(High-Order Autocorrelation, HOA)或者矩特征[PL92]等，这些特征被认为具有一定的仿射变换(平移、尺度和旋转)不变性。

5.3.3 误配准学习方法

这类方法试图在分类器的训练阶段学习可能的配准误差[HB97, Ma02, SG03]，从而使得最终得到的分类器不受误配准的影响。Martinez 的扰动(Perturbation)[Ma02]方法就属于这一类，该方法通过扰动扩大原型图像集(Gallery)，从而增加了每个待识别人脸的图像数目，然后采用高斯混合模型对每个人的图像进行建模，因此，该方法每个人脸都需要较大的存储空间来建模，时间和空间复杂度都比较高。本文提出的方法则是在训练阶段通过加扰扩大训练集，而并不需要扩大原型图像集，因此其时间和空间复杂度都很低，更加实用有效。

5.3.4 变换不变距离度量方法

模式识别最终在本质上要计算模式描述之间的距离或者相似度，对外部变换不变的距离度量是一种自然的思路。切距离(Tangent Distance)及其相关技术[SC98, KM04, SG03]是这类方法中最典型的，也是 OCR 领域广泛采用的技术。切距离方法认为一个原型模式经过各种仿射变换产生的模式在空间中扩展为一个流形结构，该流形在原型模式点处可以用一个线性切空间来近似，所谓切距离就是输入模式到该原型模式线性切空间的距离。而[SG03]则采用了 PCA 方法对切空间进行建模。鲁棒匹配(Robust Correlation)[JK02]和全局仿射变换(Global Affine Transform, GAT)[WK01]也可以归类为变换不变的距离度量方法，其基本思路都是试图在进行相关量计算之前进行图像之间的仿射变换校准。

5.4 考虑误配准鲁棒性的识别性能评价方法

不同算法的识别性能对误配准会有不同的鲁棒性，对误配准过分敏感的识别算法是难以实用化的，因此，单纯采用单一的识别率指标来评价某一识别算法的性能是不够的，

我们需要一种新的评价体系来评价识别算法的性能。例如，考虑三种人脸识别算法 A 和 B，在相同的测试条件下（已知眼睛标准位置），其识别性能随误配准变化的情况如图 5.5 所示：A 算法在没有误配准的情况下识别率为 92%，B 算法在没有误配准的情况下识别率为 100%，C 算法则是将 A 算法性能曲线人为下调 10 个百分点后的结果。直观上，B 算法似乎明显优于 A 算法。但事实真得如此吗？该如何合理地评价这两种识别算法的实际性能呢？本节对此进行了深入探讨，考虑了存在配准误差的情况下如何量化比较它们的系统识别性能。

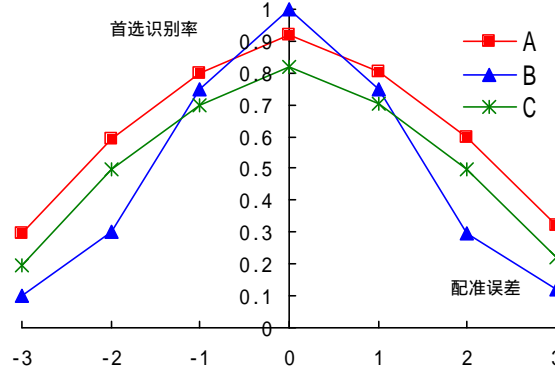


图 5.5 三种识别算法 A, B, C 的识别率与误配准的关系

Fig. 5.5 Relationship between the mis-alignment and the recognition rates of three methods A, B, C

我们进一步假设对 A, B, C 算法采用同样的眼睛定位算法进行系统集成。如前所述，人脸特征对齐算法往往很难完美的定位面部的关键特征点，为此，本文考虑对特征对齐算法的精度以概率分布的方式进行建模。设特征点（本文特指眼睛中心）实际位置为 P^* ，而自动特征定位算法给出特征点位置为 P ，则可以合理的假定 P 偏离 P^* 的程度满足正态分布，即：

$$p(\partial) \sim N(m, s^2),$$

其中 $\partial = d(P, P^*)$ 。

下面给出本文提出的考虑了误配准敏感性的识别算法性能评价方法：

定义 1 考虑误配准鲁棒性的正确识别率 r^* 定义为：

$$r^* = \int_{\Omega} p(\partial) r(\partial) d\partial, \quad (5-1)$$

其中： ∂ 为配准误差， Ω 为可能的配准误差值域范围， $p(\partial)$ 为眼睛定位算法的配准误差分布概率密度函数， $r(\partial)$ 为识别算法在误配准为 ∂ 时的识别率。

定义 r^* 实际是把误配准的分布作为权值计算识别算法的加权平均识别率，反映了识别系统的期望识别率，因而可以很好地反映识别算法在与特征对齐算法集成后的实际系统的识别性能，比单一的识别率具有更大的应用价值。

r^* 实际上同时考虑了识别性能与鲁棒性，但有时候单独考虑鲁棒性也是非常必要的。例如，对图 5.5 所示的算法 A 和 C 而言，C 算法的识别性能曲线是 A 曲线简单向下平移 10 个百分点而得到的，尽管 C 算法的平均识别性能下降了 10 个百分点，但直觉上二者对误配准的敏感性/鲁棒性应该是没有差别的（因为它们的下降速度一致），为了度量这一点，本文进一步提出下面的鲁棒性及其敏感性度量准则。

定义 2：识别算法对误配准的鲁棒性度量 R 定义为：

$$R = \int_{\Omega} p(\partial) \frac{r(\partial)}{r_0} d\partial = \frac{r^*}{r_0}. \quad (5-2)$$

其中， r_0 表示识别算法在没有误配准时时的识别率，其他符号含义同公式 5-1。

R 反映了识别算法的识别性能受误配准的影响程度，其值域范围为 $(0, 1)$ ，该值越大表示识别算法越鲁棒，等于 1 是鲁棒性的理想情况：识别算法的性能不受误配准的影响。为了叙述的方便，本文还同时引入了识别算法对误配准的敏感性度量，取 R 的倒数是一个自然的选择。

定义 3：识别算法对误配准的敏感性度量 S 定义为：

$$S = 1/R = r_0/r^*. \quad (5-3)$$

S 的值域范围 $(1, \infty)$ ，该值越大表示该算法对误配准越敏感，相反，该值越接近 1，则表示识别算法的鲁棒性更强！

上述三个度量的引入，使得我们能够有效的对识别算法的总体识别性能进行量化的评估。例如，对图 5.5 所示的 A, B 算法，假设它们所使用的眼睛定位算法的定位误差满足 0 均值 1 方差的高斯分布，即： $p(\partial) \sim N(0,1)$ （值得指出的是：这意味着如果规定定位误差在 3 个像素以内为正确定位，定位正确率应为 99.7%），则根据公式 5-1~公式 5-3 计算出来的 A, B 算法的识别性能及其对误配准的鲁棒性和敏感性度量如表 5.1 所示。其结果表明：在误配准满足 0 均值 1 方差正态分布的条件下，算法 A 的实际识别性能要优于算法 B，算法 A 也具有更强的鲁棒性，算法 B 则对误配准更敏感——这一结论应该是更为合理的，也是合乎对图 5.5 的直觉感受的。

另外， R 和 S 的引入使得我们能够有效的抛开识别率的高低而单独评估识别算法对误配准的鲁棒性，例如对图 5.5 所示的算法 A 和 C，计算得到的 R, S 是相同的，这非常吻合我们对他们鲁棒性能的直观感觉。

表 5.1 考虑误配准敏感性后的 A, B, C 算法识别性能比较

Tab. 5.1 Performance comparison between A, B, and C after considering mis-alignment sensitivity

算法	r_0 (%)	r^* (%)	R	S
A	92	82.3	0.895	1.12
B	100	79.5	0.795	1.26
C	82	72.3	0.895	1.12

5.5 基于配准误差学习的增强判别分析方法（E-Fisherface）

误配准并不改变人脸图像类别属性，但却增大了归一化的图像空间中人脸类别的类内散度，也在一定程度上减小了不同人脸之间的类间散度，从而导致分类困难。因此，一种直观的解决方案就是将可能的误配准加入到训练集合中，通过加强训练的方法来对误配准进行建模。基于此，本文提出了增强的 Fisherface 人脸识别方法，简称 E-Fisherface 方法。简单的说，该方法通过人工扰动给定的眼睛位置来产生若干存在误配准的“虚拟”人脸样本，并把这些样本作为正常样本用来训练 LDA，以使得 LDA 能够学习到误配准导致的类内/类间变化，从而增强 Fisherface 方法对误配准的鲁棒性。该方法的具体过程如下：

（1）PCA 变换矩阵的计算

与标准 Fisherface 方法一样，首先采用 PCA 进行高维图像特征的降维处理。需要指出的是，计算 PCA 的训练集有两种可能的选择：一种是采用“扰动”扩展后的训练集，另一种是只采用原始的未扩展训练集。考虑到计算的复杂度问题，本文采用了后一种策略来产生 W_{pca} ；

（2）增强的 LDA 训练过程

对训练集合中的任意图像，获取其眼睛特征位置后，采用图 5.6 所示的 8 邻域扰动方法扩展 LDA 的训练集合。也就是说，左右眼睛分别有 9 种不同的位置选择（包括精确位置），这样二者组合可以产生 81 种不同的误配准扰动图像，也就是说，训练集合中的每一幅图像都最终产生 81 幅归一化的人脸图像加入到训练集合中，图 5.7 给出了若干图像的示例，这些扰动图像进一步用前面计算得到的 PCA 变换矩阵 W_{pca} 进行降维，然后训练 LDA 变换矩阵得到加强训练后的 W_{lda} 。由于这些存在配准误差的训练图像的加入，学习得到的线性判别分析模型可以在很大程度上将误配准造成的模式差别学习为类内变化而不是类间变化，从而提高识别的精度；

（3）最近邻分类过程

得到 W_{pca} 和 W_{lda} 之后，识别过程与标准 Fisherface 方法一样，就变得相对简单了。即首先计算 Gallery 中的所有图像（眼睛位置相对精确已知）的 LDA 特征。然后对探测集中的任意图像，同样计算其 LDA 特征并与 Gallery 中所有已知 LDA 特征进行相似度计算排序以得到识别结果。

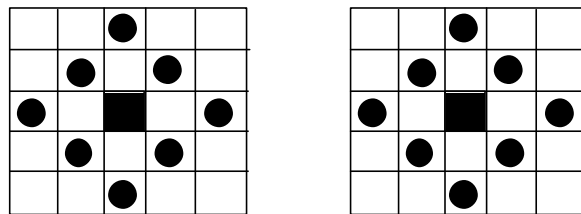


图 5.6 眼睛中心位置配准扰动的 8 邻域选择

Fig. 5.6. The eight-neighbors of the eye center perturbation.

需要指出的是，本文提出的 E-Fisherface 方法与 Martinez 的扰动(Perturbation)方法[16]是很不相同的：Martinez 的扰动方法扩大的是参考集（Gallery）而不是训练集合，参考集中的每个原型人脸图像都被加扰，从而增加了每个待识别人脸的原型图像数目，然后采用高斯混合模型对这些原型图像进行建模，因此，该方法每个人脸都需要较大的存储空间来建模，时间和空间复杂度都比较高。而本文的方法并不增大 Gallery 集，而只是

增大训练集，因此空间复杂度更低，时间耗费也主要是在训练过程，一旦训练完成，可以实现非常快速的特征提取和特征比对。



图 5.7 根据单一输入图像生成的多个虚拟样本

Fig. 5.7 Examples of virtual image from training sample for augmenting the training set of LDA

5.6 Gabor 特征的误配准鲁棒性分析

如前所述，配准误差导致了归一化的人脸图像对应像素高层语义的不同并最终导致了识别性能的下降。因而提取对配准误差具有良好不变性的人脸表示特征是解决误配准灾难问题的重要方法之一。Gabor 特征提取的是像素邻域内的局部特征，也就是说，在图像空间上相邻的像素点会有非常相似的 Gabor 特征（幅值），因而具有比灰度特征更强的空间相关性，这种更强的空间相关性可以在一定程度上降低一定范围内的配准误差造成的高层语义的不匹配程度，也就是说，Gabor 特征对误配准应该具有更为鲁棒的性能。[BL90, LV93, WF97, Wur97]等前人工作也表明 Gabor 特征是一种良好的人脸表示方法，用于人脸识别可以实现更为精确的人脸识别。

尽管如此，但却没有文献针对性地探讨 Gabor 特征对于“误配准灾难”问题的有效性，也没有文献给出 Gabor 特征鲁棒性能的量化评价方法，以及与其他表示方法的量化比较结果，本文对此进行了补充，通过实验考察了基于 Gabor 特征的人脸识别方法对于平移、旋转和尺度变化的鲁棒性，从误配准鲁棒性的角度分析了基于 Gabor 特征的人脸识别方法的有关性质。具体而言，在 5.7 小节的实验部分我们考察了第四章所提出的 AdaGaDA 方法以及 GFC 人脸识别方法的误配准鲁棒性，通过在 FERET 人脸库上进行误配准测试绘制误配准-识别率曲线，并计算 5.4 节所定义的性能评价量化指标，与 Fisherface, E-Fisherface 进行了对比。

5.7 实验评估与分析

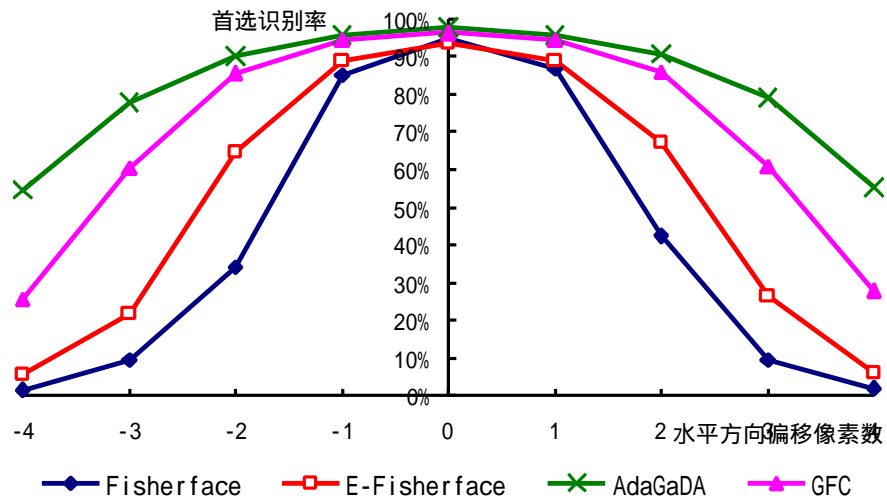
为了量化比较上述识别方法对误配准的鲁棒性，分别在 FERET-FB 测试库和 CAS-PEAL 正面人脸测试库上进行了两组实验。在 FERET 上的实验沿用了 5.2.1 小节所描述的方法，目的是能够分别测试这些方法对于误配准引起的平移、旋转和尺度变化的鲁棒性，并应用本文 5.4 节所定义的性能评价指标对它们的性能进行量化和比较。而在 CAS-PEAL 人脸库上的实验则是一个完整的系统测试，其中的眼睛位置是自动算法定位的结果，由此可以很好地体现本文所提出的方法的重要性。

5.7.1 在 FERET 上的对比实验

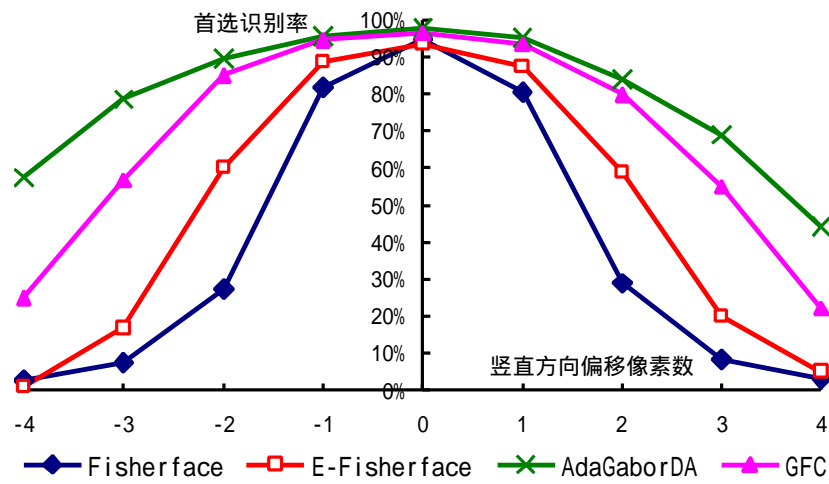
采用与 5.2.1 节同样的测试方法，在 FERET 人脸数据库的 FB 测试集上进行实验。图 5.8 给出了 Fisherface、E-Fisherface、GFC、AdaGaDA 几种不同的人脸识别方法的首选识别率随误配准引起的平移、旋转和尺度等参数的变化而变化的情况，其中(a-d)分别给出了水平平移、竖直平移、平面内旋转和尺度四种不同的情况。表 5.2 则给出了在假定眼睛定位误差满足 $p(\partial) \sim N(0,1)$ 情况下，几种算法在考虑了误配准后的综合识别性能和误配准鲁棒性/敏感性的对比情况。需要注意的是这些图表都是在双眼距离为 28 像素的条件下得出的。

从图 5.8 不难看出，与 Fisherface 方法相比，E-Fisherface 方法的识别性能随配准误差增大而下降的趋势更加平缓，尤其是对于尺度和旋转情况。从表 5.2 则可以更方便的看出二者性能的量化比较情况，尤其对于旋转和尺度变化，E-Fisherface 的鲁棒性评测指标远远好于 Fisherface 方法。这充分说明本文提出的 E-Fisherface 方法具有更强的误配准鲁棒性。值得指出的是，E-Fisherface 方法在没有配准误差的情况下识别性能 (r_0) 略有下降，这主要是因为加入扰动数据导致类内模式覆盖扩大，不同类别模式之间重叠的可能性也因此增大。尽管判别分析技术可以在一定程度上降低类内散度、增大类间散度，但还不够理想，因此导致了较精确配准条件下识别性能的略微下降。但这与该方法对配准误差鲁棒性的优点相比是可以大体忽略不计的，我们的实践经验也表明了这一点。

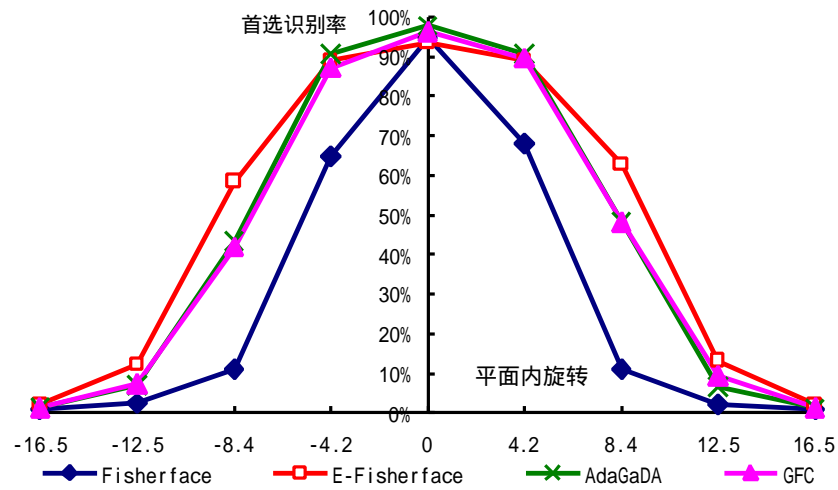
图 5.8 和表 5.2 同样表明：基于 Gabor 特征的人脸识别方法 GFC 和 AdaGaDA 都具有更好的误配准鲁棒性，GFC 和 AdaGaDA 的鲁棒性都远远好于 Fisherface，尤其是对于水平方向和竖直方向的平移情况，其鲁棒性要比 E-Fisherface 更加鲁棒。这在一定程度上验证了 5.6 小节的分析，也更进一步说明了 Gabor 特征的优越性。但也必须注意的是：Gabor 特征对旋转和尺度还比较敏感，尽管要比 Fisherface 鲁棒，但要比 E-Fisherface 略差。这与 Gabor 特征的方向和频率（尺度）选择性有直接的关系，当人脸图像发生平面内旋转、尺度变化的时候，Gabor 特征会发生相应的变化，因而导致 Gabor 特征的“对齐”出现偏差，最终影响识别性能。



(a)



(b)



(c)

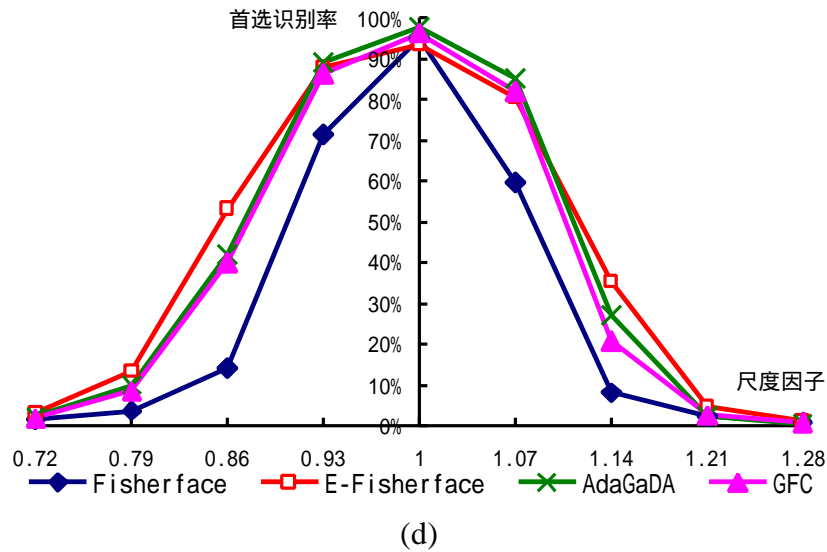


图 5.8 E-Fisherface 和 Fisherface 方法识别性能随配准误差变化情况比较
Fig. 5.8 Performance-Misalignment curve comparisons between Fisherface and E-Fisherface

表 5.2 考虑了误配准后的综合识别性能和误配准鲁棒性/敏感性的对比
Table 5.2 Performance comparisons between Fisherface and E-Fisherface when considering mis-alignment robustness

误配准	算法	r_0 (%)	r^* (%)	R	S
X 平移	Fisherface	94.8	83.6	0.881	1.13
	E- Fisherface	93.4	87.5	0.937	1.07
	GFC	96.3	93.8	0.974	1.03
	AdaGaDA	97.2	95.6	0.979	1.02
Y 平移	Fisherface	94.8	80.2	0.846	1.18
	E- Fisherface	93.4	86.4	0.925	1.08
	GFC	96.3	93.2	0.968	1.03
	AdaGaDA	97.2	95.1	0.973	1.03
旋转	Fisherface	94.8	71.2	0.751	1.33
	E- Fisherface	93.4	87.0	0.931	1.07
	GFC	96.3	86.1	0.894	1.12
	AdaGaDA	97.2	87.8	0.899	1.11
尺度	Fisherface	94.8	70.8	0.747	1.34
	E- Fisherface	93.4	82.9	0.887	1.13
	GFC	96.3	82.5	0.857	1.17
	AdaGaDA	97.2	84.9	0.869	1.15

图 5.8 和表 5.2 同时还表明：无论哪种识别方法，它们对水平方向平移、竖直方向平移、平面内旋转、尺度变化四种不同的误配准情况的鲁棒性都是依次下降的，也就是说，水平方向的平移对于识别系统识别性能的影响是最弱的，而尺度变化引起得性能下降则最明显。因此，后续的研究尤其需要注意解决旋转和尺度不变性问题。

5.7.2 在 CAS-PEAL 上的对比实验

为了进一步验证本文提出的问题及其解决方案的有效性,我们在 CAS-PEAL 人脸库上评测了上述四种算法在与同一人脸检测、眼睛定位模块集成得到全自动的人脸识别系统时候的性能。其中人脸检测、眼睛定位方法分别采用了文献[MY99]和[CS02]的方法。实验结果见图 5.9 和表 5.3。

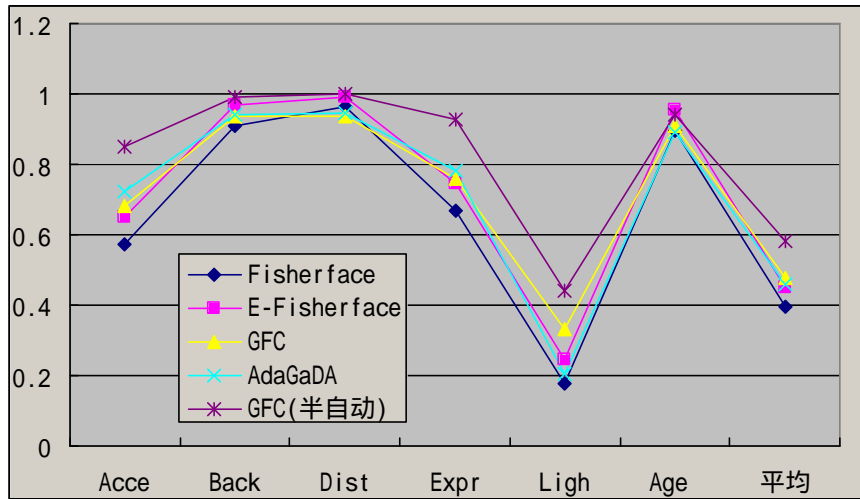


图 5.9 在 CAS-PEAL 人脸库上的全自动测试性能比较

Fig. 5.9 Comparisons of the performance of several automatic face recognition systems

表 5.3 全自动人脸识别算法在 CAS-PEAL 人脸库上的性能比较

Table. 5.3 Comparisons of the performance of several automatic face recognition systems

	Fisherface	E-Fisherface	GFC	AdaGaDA	GFC(半自动)
表情	0.669	0.745	0.761	0.783	0.929
饰物	0.571	0.648	0.681	0.721	0.851
光照	0.179	0.244	0.333	0.205	0.443
背景	0.91	0.967	0.937	0.942	0.989
老化	0.895	0.955	0.909	0.893	0.939
距离	0.963	0.989	0.938	0.945	1.00
加权平均	0.396	0.449	0.478	0.461	0.583

图 5.9 和表 5.3 中,除了列出了全自动的测试结果之外,还同时给出了眼睛位置手工标定时 GFC 算法的识别性能作为参照系。可以看出,由于全自动系统中眼睛的位置定位存在一定的误差,造成了一定程度的配准误差,因此全自动 GFC 系统的识别性能明显低于半自动 GFC 系统(眼睛位置手工标定)的性能,其平均识别率下降了 11 个百分点。而对比全自动的 Fisherface 和 E-Fisherface 可以看出:由于对配准误差进行了学习, E-Fisherface 的识别率比 Fisherface 提高了 5.3 个百分点。而基于 Gabor 特征的全自动 GFC 和 AdaGaDA 方法的识别性能更优于 Fisherface 方法。这些实验进一步说明了误配准灾

难问题的重要性以及本文所研究的方法的有效性。

而 E-Fisherface 与 GFC、AdaGaDa 相比，尽管 E-Fisherface 的总体识别率略低，但由于其没有非常耗时的 Gabor 特征提取过程，因此其人脸注册和识别速度都可以明显比 GFC 和 AdaGaDa 更快，从这一点上来看，其性能也是可以接受的，尤其是针对那些需要快速识别和更少存储空间的应用系统而言，这一点更加突出。同时，可以相信，同样的误配准学习策略用于 GFC 和 AdaGaDa，则应该可以使得他们对于误配准有更好的鲁棒性。

5.8 本章小结

为了揭示现有典型识别算法的识别性能对特征配准准确度的敏感程度，本文通过对眼睛位置人为加扰的方法对 Fisherface 算法的识别性能随平移、旋转和尺度改变而变化的情况进行了实验评估，结果表明：Fisherface 的识别性能随着误配准的增大而急剧下降——我们称这一现象为“误配准灾难”。本文进一步提出了一种考虑了误配准鲁棒性的识别性能评估方法，给出了识别算法对误配准的鲁棒性度量策略；最后提出了一种对误配准进行强化学习的“误配准灾难”解决方案 E-Fisherface。在 FERET 人脸图像数据库上的实验表明该方法可以有效的提高识别算法对误配准的鲁棒性。本文研究内容对于设计鲁棒实用的人脸识别系统而言是非常重要的。总之，本章的主要贡献包括：

1. 首次提出并系统研究了人脸识别中的“误配准灾难”问题。首先，通过实验评估了眼睛位置在偏离手工标定真实位置的情况下 Fisherface 方法性能的变化情况，结果发现：Fisherface 方法的识别性能随着误配准（反映为平移、旋转和尺度等仿射变换参数的变化）的增大而急剧下降，本文称这一现象为“误配准灾难”问题；
2. 提出了一种考虑了误配准分布的人脸识别系统识别性能评价方法，以弥补识别率变化曲线难以量化比较不同算法性能的不足，并给出了衡量识别算法对误配准鲁棒性和敏感度的相应度量；
3. 针对 Fisherface 方法的“误配准灾难”问题，提出了一种增强的判别分析方法 E-Fisherface，该方法在模型训练阶段通过对误配准进行学习，可以有效的降低原有判别分析方法得到的判别特征对误配准的敏感性，大大提高了 Fisherface 算法的实用程度；
4. 从误配准鲁棒性的角度，进一步考察了基于 Gabor 特征的人脸识别（AdaGaDa 方法和 GFC 方法）方法，结果表明：AdaGaDa 方法和 GFC 方法都具有良好的误配准鲁棒性，尤其对于误配准导致的平移变换，而对旋转和尺度变化其鲁棒性稍弱，需要进一步的研究。

第六章 人脸识别应用系统设计

在人脸识别技术尚未达到完全成熟阶段的今天，系统设计的作用非常突出，良好的系统设计可能起到事半功倍的效果，而拙劣的系统设计可能会葬送最好的核心识别算法的前途。鉴于此，本章初步探讨了人脸识别应用系统设计中需要考虑的一些共性关键问题，如现场环境设计、摄像设备选择与安装、注册原型图像的挑选、核心算法选择标准、学习集选择、模型在线更新等。最后，基于前面章节提出的核心技术和关键问题解决方案，给出了人脸照片比对系统、模拟的 XP 用户自动登录系统和机场登机旅客身份验证系统三个典型 AFR 应用原型系统的设计方案和系统情况。

6.1 问题的提出

一个实用的人脸识别应用系统能否有良好的识别性能当然主要依赖于核心识别算法的性能，但系统设计是否合适同样在很大程度上影响着应用系统的成败，尤其是在目前情况下，人脸识别技术并非成熟技术，还面临着诸多关键问题的挑战，系统设计的作用更加突出，良好的系统设计可能起到事半功倍的效果，而拙劣的系统设计可能会葬送最好的核心识别算法的前途。本章初步探讨了人脸识别应用系统设计中需要考虑的一些共性关键问题，并给出了两大类三个典型应用系统的设计方案。

6.2 应用系统设计共性关键问题初探

6.2.1 现场环境设计与改造

人脸识别的研究现状表明（参见本文第二章），目前最好的人脸识别系统也只有在限定条件下才能取得最佳的识别效果。因此，在条件允许的情况下，必须认真设计应用系统的现场环境，必要时对现场环境进行改造。设计时需要考虑的主要要素包括：

（1）系统应用时段内光照条件应尽量一致

尤其要尽量保证系统注册原型图像的采集条件应该尽量与系统实际运行时候的图像采集条件一致。这对于完全通过人造光源照明的应用现场而言是可以比较容易地做到的，但对于人造光源与自然环境光源（例如通过窗户）混合照明的现场来说，保持全天候二十四小时照明条件尽量一致并非易事，必要的时候要通过人造光源进行补偿，以便消除这些影响。

（2）避免逆光、侧光、高光和曝光不足

逆光、侧光、高光和曝光不足对于现有的多数人脸识别系统而言都是灾难性的（很多系统甚至不能检测到人脸），因此，除非万不得已，应尽量避免出现这些情况。例如，尽量不要将摄像设备对准室外、窗户等，同时应该尽量使得被识别对象是面向光源的。必要的时候要通过反光板、毛玻璃灯罩等人造光源进行补偿，以便消除逆光、侧光、曝光不足等问题。

6.2.2 摄像设备选择与安装

摄像设备是人脸识别系统的“眼睛”，其采集到的人脸图像的质量和属性在很大程度上决定了人脸识别系统的最终性能，需要慎重挑选，并进行合适的安装。

(1) 摄像设备选择

并非功能最全或者最昂贵的摄像设备就是最好的，对不同需求的应用系统应该采用不同种类、不同功能的摄像设备。需要考虑的摄像设备要素主要包括：

a) 感光方式和成像质量

目前主流的摄像设备主要采用 CCD 和 CMOS 图像传感器。对多数应用而言，应该尽量选用 CCD 摄像头，尽量避免选用价格较低但成像质量较差的 CMOS 摄像设备；另外，对于需要在较暗的环境下工作的应用而言，应选择最小照明度小的产品。

b) 镜头焦距与变焦镜头

除非有相应的形变矫正策略，应尽量使用标准焦距镜头，而避免采用形变较大的广角甚至鱼镜头。这主要是因为人脸识别算法开发和模型训练所采用的图像多数是标准镜头摄制的，鱼镜头带来的形变会导致识别模型的失效。对于一些用户不会配合的应用场合，选择具有变焦镜头的摄像设备可能是非常有必要的，例如可以在检测到人脸比较小的情况，自动变焦以摄取分辨率更高的人脸图像用于识别。

c) 光圈、自动光圈、白平衡与增益调整功能

对于需要在自然环境光源条件下、全天候运行的应用系统而言，具备自动光圈、白平衡和增益功能的摄像设备是必备的。自动光圈和增益调整功能使得应用系统可以根据输入图像的明暗程度调节光圈或增益以避免过度曝光或者欠曝光。而白平衡则可以抵消外界环境光源带来的色温变化，这一点对于基于肤色的人脸检测系统尤为关键。

(2) 摄像设备安装

a) 视角：尽量保证正面视角图像。现有的人脸识别算法的姿态鲁棒性都比较差，因此摄像设备的安装应该尽量保证用户能够基本正面镜头。对于固定摄像机的场合而言，身高不同会导致视角的变化过大，针对该问题，可以考虑的策略包括放置多个不同位置的摄像设备，在识别时进行多路识别结果的融合。

b) 距离：人与摄像设备距离的设定依赖于镜头的焦距情况，一般应该保证人脸在图像的中央位置，大小不应该少于 32×32 像素。还需要注意的一点是人与摄像头的距离不要过近，以免出现严重的桶形失真。

(3) 摄像设备的一致性

由于不同的摄像设备具有不同的内部参数，因此不同的摄像设备摄制的图像会有很多隐含的差别，会在一定程度上影响人脸识别系统的性能。FERET97 测试也表明了这一点。因此，如果能够保证训练模型图像、注册图像、现场识别图像的采集设备是统一的，至少尽量保证是同一型号的，识别效果会有更大的保障。

6.2.3 增强人脸检测功能

人脸检测技术日臻成熟，对多数用户比较配合的应用系统而言，现有的人脸检测算法已经能够有 95% 以上的检测率，可以保证应用系统的需求。但实践表明，对很多室外

应用而言（比如视频监控系统），尤其是在日照充足的中午，逆光、侧光会使得检测系统的性能下降很多（高原地区更甚）。在这些场合下，应用相对比较成熟的运动检测、背景减等方法，进一步降低误检率、提高检测率是必须考虑的策略。

提高人脸检测算法性能的另一种策略是对现场重新进行学习。典型的做法是现场采集典型的人脸图像，加入到原有的训练图像集中，重新训练检测模型，从而提高系统对新环境的适应能力。

6.2.4 选择合适的注册原型图像

对有些应用系统而言，例如罪犯库照片检索系统、敏感人物实时监控系统等用户不会主动配合的场合，每个待识别人脸的原型图像是预先通过其他途径获得的，通常只有少量甚至一幅原型图像。而对门禁、考勤等用户需要配合的系统而言，用户是愿意配合的，原型图像是在用户注册时现场采集的，对这类系统而言，选择合适的原型图像是非常关键的。这主要包括以下关键因素：

（1）原型图像的数量

一般而言，较多的原型图像可以使识别模型具有更强的内插能力，从而提高系统的识别性能。但如果人脸表示模型和分类算法不够理想，原型图像增多也会在一定程度上提高系统误识率，同时还会使得系统的空间、时间复杂度上升，用户注册时间、识别系统响应时间都会增加，系统的可用性下降。因此，原型图像的数量一般可以根据不同应用控制在 3 到 15 幅。

（2）原型图像的性质

如果可能，多幅原型图像应该覆盖不同的变化条件，比如不同的表情、一定程度的姿态变化、不同的摄像距离等等。但需要注意的是，如果核心算法不具备姿态估计和校正能力，加入姿态变化过大的原型图像效果可能适得其反。

（3）原型图像的有效性

对于实用的人脸识别系统而言，全自动的采集注册图像几乎是必需的。而人脸检测、面部特征定位是难以保证 100% 的正确率的，如果要避免用户手工修改这一耗时而不方便的程序，系统必须具备自动的原型图像有效性测试功能。在本文的系统中，通过计算新输入图像到人脸子空间的距离来判断原型图像的有效性。

（4）原型图像的差异性

容易理解，两幅非常相似的原型图像对识别性能的贡献是很微弱的。如前所述，全自动的采集注册原型图像又往往是必需的，如果不加入相应的控制策略，很容易会采集到连续多幅非常相似、对识别性能无意义的原型图像。最简单的策略就是对每一幅新加入的原型图像都与该人脸的其他原型图像进行匹配，只有差别超过一定阈值而且有效的原型图像才被加入。

6.2.5 核心识别算法的选择

识别率高低和识别速度并非是选择核心识别算法时的全部依据。针对一个具体的应用，在选择算法时至少要综合考虑以下关键因素：

- （1）要求的识别率：这当然是需要关注的第一要素。条件允许的话，应该构建一个与该应用实际环境匹配的人脸库，进行相关的测试实验，以便确定哪些算法最能够满足应用需求，实际上，FRVT 测试就是为了这一目的而进行的大规模测试。
- （2）系统响应时间：识别率很高但识别速度根本不能满足具体应用需求的算法实际上

是不能被应用的。反过来说，应该尽量选择刚好能够满足系统响应时间需求，又具有较高的识别率的算法。需要注意的是：系统响应时间一方面与算法的复杂度有关，另一方面与数据库中待识别人脸数及其原型图像数目有关，在计算时应该首先估计算法单次匹配一幅测试图像与一幅原型图像的时间，然后将其与全部待识别人脸的全部原型图像相乘以得到总的系统响应时间。

- (3) 存储空间限制：不同算法中原型人脸图像的表示不同，其所需要的存储空间也不相同，可能从几十个字节到几百千字节不等。对于待识别人脸数目较少，原型人脸图像也较少的应用系统而言，需要的存储空间的差别可能不大，但对有些应用系统而言，例如需要在几百万乃至上千万人脸图像中进行检索的应用系统而言，就必须认真考虑存储和检索策略问题。
- (4) 系统扩展能力：这包括两方面的内容，其一，增加新人脸类别的方便程度和复杂度；其二，增加某一类别新原型图像的方便程度和计算复杂度。例如，基于 SVM 的方法对很多人脸识别应用而言就不是非常理想的方法，无论是“一对一”还是“一对多”策略，每增加一个新的人脸类别甚至增加一个新的原型图像就需要重新训练大量分类器，这对很多应用系统而言都是不可接受的。

6.2.6 识别输出策略

很多应用人脸识别系统都是摄像设备时刻联机捕获人脸图像并进行识别的，在这种情况下，考虑到单次识别可能由于各种原因出现错误，每次都输出单帧人脸图像的识别结果并非明智的选择，可以考虑对多帧人脸图像的识别结果进行适当的融合以便输出更为鲁棒的识别结果。可以考虑的融合策略有：

- (1) 平均法：即将连续若干帧识别的匹配结果（比如相似度）进行平均，输出最大平均相似度的识别结果。这种方法的优点是可以保证输出结果的平滑性。
- (2) 最大相似度法：即直接输出连续若干帧识别中匹配相似度最大的作为最终的识别结果。
- (3) 最大信度法：首先定义一个识别信度，然后输出连续若干帧识别中信度最大的作为最终的识别结果。例如，可以将单次匹配的首选相似度与次选相似度之比作为信度进行排序。

应用多帧识别结果融合时需要注意的问题是：必须通过跟踪策略保证参与融合的多帧图像的确来自同一人脸，否则就会出现识别结果的混乱。

6.2.7 学习集的构建

毋庸置疑，现有的多数人脸识别算法都具有一定程度的“学习集”依赖性，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，其实际系统中的泛化能力也并不能得以保障。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) 学习集与测试集“人群”特性的一致性。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) 学习集与“测试集”成像条件应尽量保持一致。这包括上面提及的摄像设备、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。

- (3) 只要学习算法的学习能力足够强，学习集的规模要尽可能得大。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。

6.2.8 识别模型更新与在线学习

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纪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纪录的“测试集”加入到学习集中，重新训练识别模型（例如线性判别分析方法的线性变换矩阵），并相应的更新所有原型图像的人脸表示特征。这种模型更新对于系统性能的提升效果通常是最明显的。
- (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。
- (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。
- (4) 自动在线学习。上面三种模型更新策略均需要一定的脱机交互工作，而自动在线学习则是对识别模型进行联机更新的方法。其最理想的模式是根据尽可能少的用户交互过程（例如仅给出识别正确或者错误的奖惩导师信号）进行全自动的模型更新。目前尚未发现对该问题的成熟的解决方案。

6.2.9 活体判别问题

对门禁等需要高安全系数的应用来说，活体判别（即判断输入是“照片”还是真实的活体人脸）是必需的。可以考虑的解决方案包括：

- (1) 眨眼判别。对于可以要求用户配合的应用系统，要求用户眨眼一到两次，系统就可以根据自动判别得到的眼睛的张合状态的变化情况来区分照片和人脸。具体的实例可参见文献[邓 03]。
- (2) 嘴部张合判别。与眨眼判别类似，要求用户张开、闭合嘴巴一到两次，系统据此区分照片与真实人脸。
- (3) 视差分析方法。无论眨眼还是张闭嘴都必须用户配合方可，视差分析方法则是不需要用户配合的方法。其基本思路是采用两个平行放置的摄像头，计算检测到的人脸区域内的视差情况，据此给出人脸区域表面点距离摄像头光心的距离，从而可以根据距离的分布情况区分当前输入是照片还是真实人脸。

6.3 基于面像的身份识别系统

确定输入人脸图像的身份是人脸识别的基本功能，也是一大类应用系统需要完成的

功能。如本文第一章所述，识别系统实际上有两类：闭集识别和开集识别。下面分别以作为闭集识别应用实例的人脸照片检索系统和作为开集识别应用实例的 Windows XP 系统模拟用户自动登录系统两个应用为例，介绍采用前面章节的技术方法设计应用系统的有关问题。

6.3.1 人脸照片比对系统

人脸照片比对是公安系统内最典型的人脸识别应用之一。公安部门抓获的很多记录在案的在逃犯罪嫌疑人往往并不透露自己的真实身份，因而公安人员必须在其所拥有的庞大罪犯嫌疑人面像数据库中进行检索以便确定其真实身份。这类应用系统的特点主要包括：

- (1) 对识别率的要求偏低，不要求首选识别率很高，按照相似度排序即可；
- (2) 系统响应时间相对并不重要，必要时可以采用高性能计算机并行运算；
- (3) 人脸图像数据库规模庞大，每人仅一幅原型图像，图像来源复杂，因而质量参差不齐；
- (4) 因数据库规模庞大，人脸描述特征应尽可能占用较小的存储空间；

基于上述几点考虑，我们可以选择 E-Fisherface 作为核心识别算法，系统的总体框架如图 6.1 所示。其中产生 PCA、LDA 变换矩阵的学习集是由 CAS-PEAL、FERET 等多个人脸库中的人脸组合而成，在训练产生 LDA 时，采用了本文第五章所提出的误配准学习方法以增强系统对于误配准的鲁棒性。最终 PCA 和 LDA 的维数分别保留为 500 和 200 维。系统除了需要存储统一的 PCA 和 LDA 变换矩阵之外，每幅原型人脸图像的 LDA 描述特征经过简单的量化压缩处理仅占用 400 字节的存储空间。这样，单次比对可以在毫秒级内完成，最后通过最近邻完成识别。结果的呈现方式为按照相似度对库中人脸进行重新排序，参见图 6.2 的系统界面图示。图 6.2 界面中，左侧上部为摄像头当前视频，左侧中部为输入查询图像，右侧的图像自上而下从左到右为按相似度排序的结果，每幅图像下部数字为其与查询图像的相似度。图 6.2 中所示首选结果即是正确的。

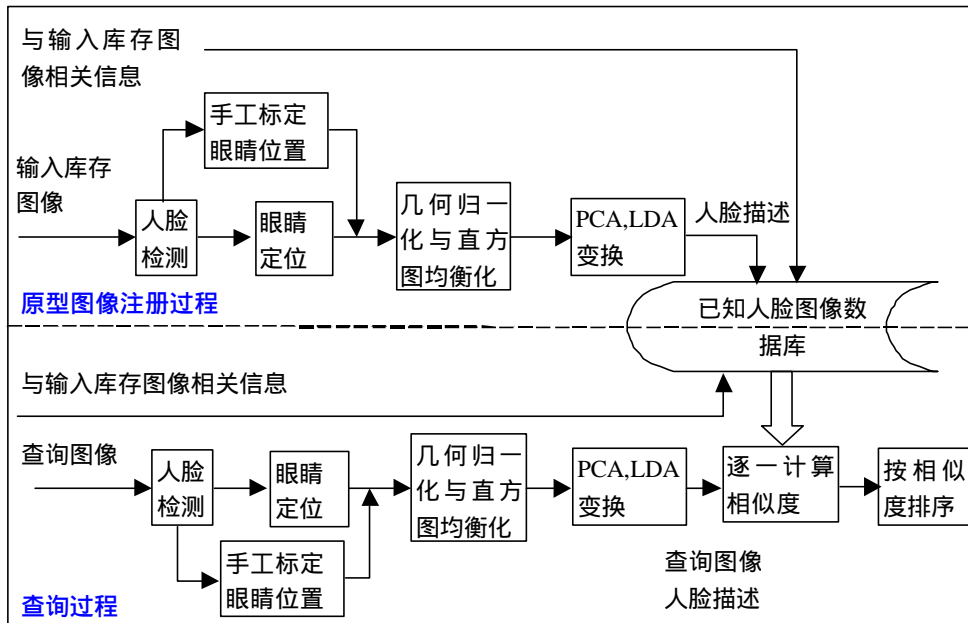


图 6.1 基于 E-Fisherface 的人脸照片比对系统框架

Fig. 6.1 Framework of facial photo matching system based on E-Fisherface



图 6.2 人脸照片比对系统界面

Fig. 6.2 Interface of facial photo matching system based on E-Fisherface

考虑到这类应用系统的用户是愿意通过增加适当的交互操作来换取更高的比对精度的，该系统还提供了以下功能：

- (1) 提供了接口供使用者通过人工交互提供更多查询信息，例如人脸的性别、年龄、种族、是否配戴眼镜等其他辅助信息，这些辅助查询信息可以进一步缩小搜索空间，从而提高系统的总体识别性能。
- (2) 除了全自动的比对功能之外，还提供接口供用户手工调整关键面部特征的位置，以便减少误配准带来的影响。
- (3) 批处理原型图像注册过程保留了所有人脸检测、眼睛定位准确度的信度信息，因此使用者可以在批处理完毕后，对可能存在眼睛定位问题的原型图像的眼睛位置进行人工修改。其中涉及到的信度信息是由定位到的眼睛区域特征及其据此得到的人脸区域特征的子空间约束得到的。

6.3.2 模拟 Windows XP 用户自动登录系统

智能人机交互是人脸识别技术的重要应用领域之一，能够识别用户身份的自动系统可以大大提升人机接口系统的自动化程度和可用性。典型的应用包括基于面像的用户自动登录系统、屏幕保护程序等。本节以模拟的 Windows XP 用户自动登录系统为例介绍它们的系统设计问题。这类应用系统的应用需求主要包括：

- (1) 对识别率的要求较高，尤其是涉及信息安全问题时，必须要有非常低的误识率；

- (2) 系统响应时间比较重要, 必须能够在所有系统平台上快速运行, 需要的系统资源尽可能少;
- (3) 注册用户一般较少, 人脸图像数据库规模不大, 每人可以捕获很多张原型图像, 质量能够保障;
- (4) 因数据库规模比较小, 人脸描述特征允许占用较多的存储空间;

为了保障识别精度和安全性, 我们可以选择第四章所提出的 AdaGaDA 方法作为核心识别算法, 系统的总体框架如图 6.3 所示。其中 AdaGabor 特征选择、PCA/LDA 过程所用的学习集是来自 CAS-PEAL、FERET 等多个人脸库中的人脸组合而成, AdaGabor 特征选择了 1800 维, PCA 和 LDA 的维数分别保留为 400 和 250 维。系统除了需要存储 PCA 和 LDA 变换矩阵之外, 每幅原型人脸图像的判别描述特征经过简单的量化压缩处理仅占用 500 字节的存储空间, 最终通过最近邻分类方法完成识别。

系统识别结果是按照开集识别的基本要求操作的, 为保证系统的安全性, 还要对首选用户的相似度进行验证, 看是否超过了预设的阈值, 如果超过阈值则系统自动登录该用户, 否则继续下一帧识别。连续多帧识别结果为同一候选用户但不能超过预定阈值时, 则对全部用户按照相似度进行重新排序, 要求用户选择用户名并输入密码, 这对用户很多的多用户系统尤其有用, 参见图 6.4 的系统界面图示。

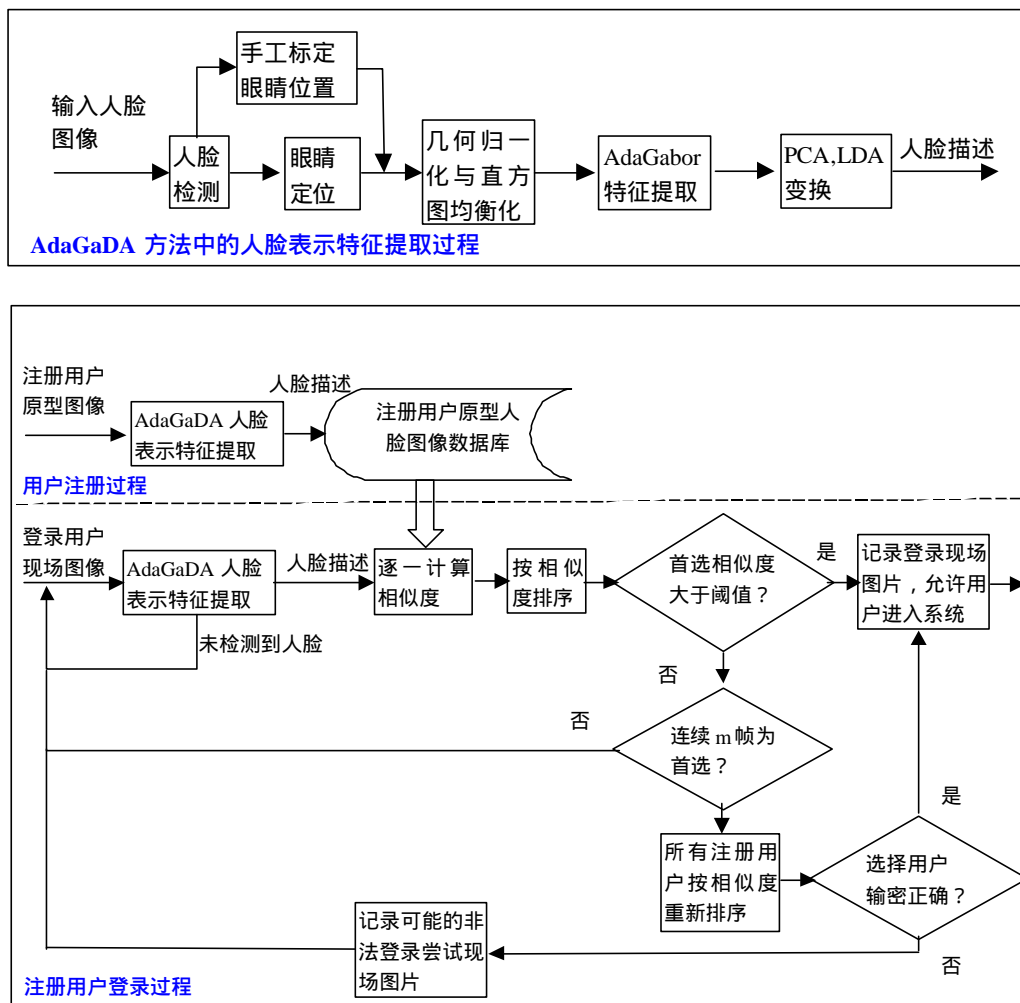


图 6.3 基于 AdaGaDA 人脸识别方法的 WindowsXP 用户登录系统框架
Fig. 6.3 Framework of WindowsXP face logon system based on AdaGaDA



图 6.4 基于 AdaGaDA 人脸识别方法的 WindowsXP 用户登录系统界面示例

Fig. 6.4 Interface of WindowsXP face logon system based on AdaGaDA

设计此类应用系统时,需要认真考虑以下问题:

- (1) 精细选择用户注册图像。这类系统通常允许相对比较繁琐的用户注册过程(原因包括注册用户少,安全性高,用户愿意配合),因此,可以通过较多的手工交互保障注册图像的数量、覆盖的变化条件、图像质量、眼睛定位的精度等。
- (2) 合理设定确认阈值。为保证安全性,这些“类门禁”应用系统的典型特点是要求要有非常低的误识率,而拒识率可以稍高。通常的做法是创建一个有一定代表性的校验图像数据库,绘制系统的 ROC 曲线,确定一个可以接受的误识率,比如 0.5%,则可以得到需要设定的参考阈值,同时也可以大体估计该系统的拒识率。
- (3) 记录系统日志。如第一章所述,人脸识别系统的一个重要优点是具有非常强大的事后追踪能力,这种优势是需要通过记录系统日志来支持的。参见图 6.3 的系统框架,不但可以记录每一次正确登录时的现场面像,在有“非法用户”试图进入系统时也可以进行记录,增强系统的事后追踪能力。

6.4 基于面像的身份验证系统

上面介绍的两个应用系统是典型的识别系统应用案例,人脸识别的另一类重要应用是基于面像的身份验证系统。本节通过介绍机场登机旅客身份验证系统为实例介绍这类系统的特点及其应用系统设计问题。

6.4.1 机场登机旅客身份验证系统

目前多数民航机场的旅客登机流程基本都是：办理登机牌——安检进入隔离区——登机。其中办理登机牌和安检进入隔离区都需要严格的查验身份证、护照等证件，而进入隔离区后登机时却只查验登机牌，这其中隐藏的一个安全隐患是：可能存在蓄谋地交换登机牌因而交换登机的问题。我们设计的机场登机旅客身份验证系统是解决该问题的一种措施。这类应用系统的应用需求主要包括：

- (1) 对确认率要求较高，必须要有相对低的拒识率；
- (2) 系统响应时间要短，必须能够在桌面平台上快速运行；
- (3) 注册用户一般都在 500 人以下，原型人脸图像数据库规模不大；
- (4) 数据库规模比较小，人脸描述特征允许占用较多的存储空间；

本文设计的系统在旅客办理登机牌 (Check In) 时全自动捕获该乘客的最多 5 幅原型人脸图像，其中用到了 6.2.4 小节所讨论的有关技巧，必要时要求乘客予以配合。这些原型图像及其它它们的人脸描述特征被存储到相应航班的数据库中。在旅客登机时，登机牌扫描过程可以获得登机牌号码，用以从数据库中调出该乘客办理登机牌时采集的原型图像及其描述特征，并与登机现场采集的图像进行自动比对，以验证旅客身份的合法性。

为了保障识别精度和安全性，仍然选择第四章所提出的 AdaGaDA 方法作为核心识别算法，系统总体框架如图 6.5，演示系统 Check In 端和 Boarding 端的界面分别如图 6.6 (a) 和 (b) 所示。其中 AdaGabor 特征选择、PCA/LDA 过程所用的学习集是来自 CAS-PEAL、FERET 等多个人脸库中的人脸组合而成，AdaGabor 特征选择了 1800 维，PCA 和 LDA 的维数分别保留为 400 和 250 维。系统除了需要存储 PCA 和 LDA 变换矩阵之外，每幅原型人脸图像的判别描述特征经过简单的量化压缩处理仅占用 500 字节的存储空间。

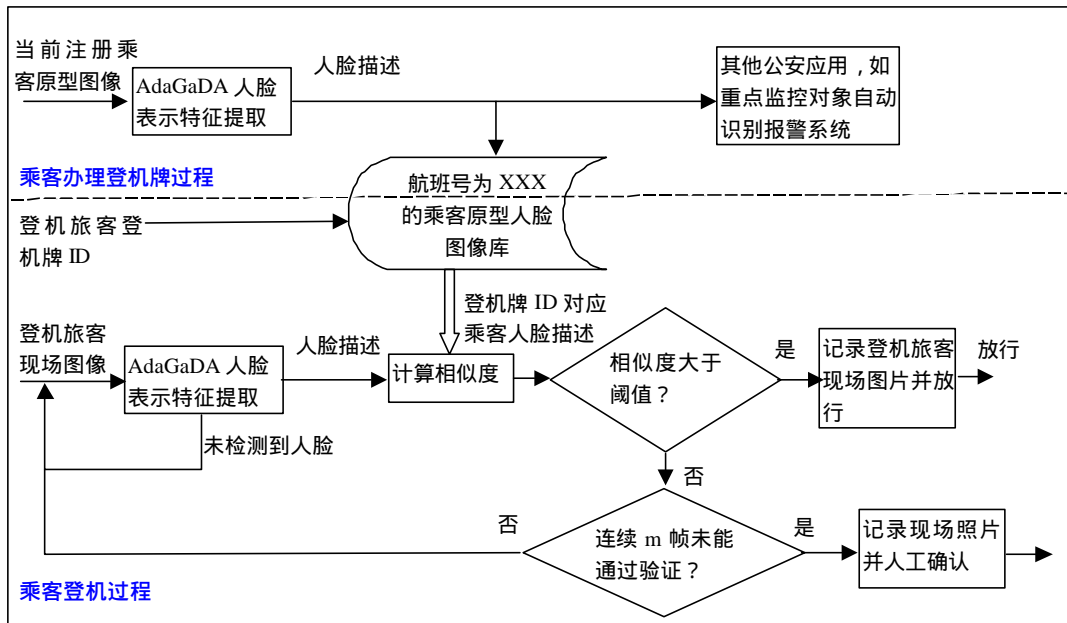


图 6.5 基于 AdaGaDA 人脸识别方法的机场登机旅客身份验证系统

Fig. 6.5 Airport boarding passenger verification system based on AdaGaDA method



(a) Check In System

(b) Boarding System

图 6.6 机场登机旅客身份验证系统界面

Fig. 6.6 Interface of airport boarding passenger verification system

6.5 本章小结

本章基于上述章节提出的人脸识别核心技术与关键问题解决方案，初步探讨了设计实用的人脸识别应用系统时需要考虑的共性关键问题，并以人脸照片比对系统、模拟的XP用户自动登录系统和机场登机旅客身份验证系统为例介绍了典型应用系统的设计问题，给出了它们的总体系统框架，演示系统界面和主要的技术特点等。

第七章 结束语

自动人脸识别(AFR)研究具有重要的理论和应用价值。经过三十多年的发展,AFR技术取得了长足的进步,但实践表明一般意义上的人脸识别技术还远未成熟,要开发出真正鲁棒、实用的AFR应用系统还需要更为鲁棒的核心算法并需要解决大量的关键问题,本文对其中涉及到的若干关键问题进行了初步探讨。本章对本论文研究内容的提出,解决的关键问题及其主要贡献进行了总结。在此基础上,探讨了本论文工作可能的后续扩展。

7.1 研究的问题及其研究意义

本论文研究的是自动人脸识别(AFR)问题,即要赋予计算机根据面孔辨别人物身份的能力。该项研究无论在理论层面还是在应用层面都具有重要的研究价值。尤其是近年来,各类安全问题日益凸显,准确可靠的身份验证逐渐成为社会生活诸多领域的必须环节,也因此受到了越来越多的研究人员的重视。

经过三十多年的发展,AFR技术取得了长足的进步,目前最好的AFR系统在理想情况下已经能够取得可以接受的识别性能,并已经出现了若干商业人脸识别系统,人脸识别技术获得了初步的应用。但应用时间和各类测试都表明:非理想条件下的人脸识别技术还远未成熟!要开发出真正鲁棒、实用的AFR应用系统还需要解决大量的关键问题,本文基于这一背景,研究了其中的几个关键问题,主要包括:作为识别必要前提条件的面部关键特征精确定位问题;高效的人脸描述特征及其相应的高精度核心识别算法;如何提高AFR系统对不可避免的配准误差的鲁棒性问题。本文最后还初步探讨了AFR应用系统设计层面的诸多共性关键问题及其参考解决方案。

本论文的研究成果不但在人脸识别研究的理论层面具有一定的参考价值,而且本文的研究内容对于设计AFR应用系统也具有一定的借鉴意义,所提出的若干关键技术已经获得了实际应用。

7.2 本文研究内容及主要贡献

本文的目标是研究开发出鲁棒、实用的人脸识别系统。基于此,本文研究了开发鲁棒实用的人脸识别系统所需要的核心技术和关键问题解决方案,重点探讨了与面部特征精确配准、鲁棒的人脸表示方法以及分类判别方法有关的几个关键问题:

- (1) 作为识别必要前提条件的面部关键特征精确定位问题。面部关键特征精确定位(即面部特征精确配准)是鲁棒实用的人脸识别系统的基本前提,而且面部关键特征定位的精度直接影响了后续的人脸描述特征提取和分类判别的精确度,可以断言,没有精确的特征配准,就没有高性能的识别;
- (2) 高效的人脸描述特征及其相应的高精度核心识别算法。AFR算法的精度和鲁棒性不仅取决于采用什么样的分类器,而且在很大程度上取决于采用什么样的特征描述人脸,亦即人脸表示的问题。
- (3) 如何提高AFR系统对不可避免的配准错误的鲁棒性问题。对实用的人脸识别系统而言,面部特征配准是不可或缺的步骤,现有的多数人脸识别系统都要依赖于面

部特征（比如眼睛位置）的严格配准来归一化人脸以便提取人脸描述特征，但面部特征配准的精度如何影响人脸识别算法的性能、如何提高识别系统对不可避免的特征配准误差的鲁棒性等问题却没有得到足够的重视。

- (4) 应用系统设计层面的共性关键问题。核心识别算法的性能当然是实用人脸识别应用系统识别性能的决定性因素，但在人脸识别技术尚未完全成熟的今天，系统设计的作用也非常突出。

本文针对上述问题开展研究工作，所取得的主要研究成果和对人脸识别研究的主要贡献总结如下：

7.2.1 全面综述了自动人脸识别研究的发展状况

本文综述部分首先探讨了 AFR 的一般计算模型，并按照其发展历史，将 AFR 研究划分为三个阶段，对每个阶段研究的特点以及代表性方法进行了较为详尽的总结，并从面部特征自动定位、人脸表示模型、分类和识别算法三个角度对人脸识别领域的主要方法进行了分类整理。在此基础上，结合近年来主要的人脸识别评测，对人脸识别研究的现状进行了阐述。此外，还总结了现有的主要商业人脸识别系统以及主要的公共人脸图像库的情况。最后讨论了人脸识别领域目前仍然面临的主要开放问题和一些需要关注的技术趋势。

7.2.2 提出了局部纹理约束的主动表观模型 LTC-AAM

主动形状模型（ASM）和主动表观模型（AAM）是图像分析与理解领域主流的算法。在对他们的优缺点进行综合分析的基础上，本文将二者进行集成，提出了一种局部纹理约束的主动表观模型 LTC-AAM，其主要特点包括：(1) 继承了 ASM 的局部纹理匹配模型，并针对 ASM 对面部轮廓点定位不够理想的问题，提出了将候选点边缘强度加入面部轮廓点局部纹理匹配模型中的策略，有效提高了轮廓点的定位精度；(2) 在局部纹理匹配的更新特征点位置的基础上，不但继续采用了 ASM 的统计形状约束，进一步引入了 AAM 中衡量当前纹理与模型纹理匹配度的目标函数，从而将 ASM 搜索过程统一为与 AAM 相同的参数优化框架；(3) 在 ASM 单步搜索和调整，借鉴 AAM 的思想，利用当前纹理来预测表观模型参数的可能变化，从而实现了启发式的参数优化过程。这些特点使得 LTC-AAM 融合了 ASM 局部模型、AAM 的纹理约束以及基于线性回归的启发式参数优化过程，从而在较大程度上继承了二者的优势。实验表明，LTC-AAM 能够比较精确的定位面部关键特征，为后续的人脸识别与确认打下了良好的特征基础。

7.2.3 提出了基于 AdaGaDA 人脸识别新方法

人脸描述是人脸识别的核心问题之一，人脸识别的实践表明：在人脸三维形状信息难以准确获取的条件下，从图像数据中提取多方向、多尺度（频率）的 Gabor 特征是一种合适的选择。使用 Gabor 特征进行人脸识别的典型方法包括弹性图匹配方法（EGM）和 Gabor 特征判别分类法（GFC）。EGM 在实用中需要解决关键特征点的定位问题，而且其速度也很难提高；而 GFC 则直接对下采样的 Gabor 特征用 PCA 降维并进行判别分析，尽管这避免了精确定位关键特征点的难题，但下采样的特征维数仍然偏高，而且简

单的下采样策略很可能遗漏了非常多的有用特征。本文针对这一问题,考虑如何对 Gabor 特征进行有效降维,将目前受到极大关注的 AdaBoost 算法创新性的应用于 Gabor 特征的选择上来,用于提取对识别最有利的特征(本文称这些选择出来的 Gabor 特征为 AdaGabor 特征),并最终通过对 AdaGabor 特征的判别分析实现识别(本文称该方法为 AdaGaDA 识别方法)。在 CAS-PEAL 和 FERET 人脸图像库上的对比实验表明:AdaGaDA 方法不但可以大大降低 Gabor 特征的维数,而且识别精度也有了较大的提高。

7.2.4 提出并系统研究了人脸识别中的“误配准灾难”问题

现有的多数人脸识别系统都要依赖于面部特征(比如眼睛位置)的严格配准来归一化人脸以便提取人脸描述特征,但面部特征配准的准确度如何影响人脸识别算法的性能却没有得到足够的重视。本文首次针对这一问题进行了系统的研究,发现 Fisherface 的识别率会随着误配准的增大而急剧下降——本文称这一现象为“误配准灾难”问题。在分析“误配准灾难”问题产生原因的基础上,本章探讨了其可能的四类解决方案,即:精确再配准方法、不变特征方法、误配准学习方法和误配准不变距离度量方法。为了能够量化比较不同算法对误配准的鲁棒性,本文提出了一套考虑了误配准概率分布的识别性能评估方法和误配准鲁棒性度量策略。针对 Fisherface 的“误配准灾难”问题,本文提出了一种对误配准进行强化训练的 E-Fisherface 方法,该方法通过在模型训练阶段加入人为配准偏差来提高判别分析方法对误配准的鲁棒性。接下来,本文还进一步考察了上一章研究的 Gabor 特征人脸识别方法对误配准的鲁棒性。上述方法在 FERET 和 CAS-PEAL 人脸图像数据库上的对比实验表明:Gabor 特征比单纯的灰度特征具有更强的鲁棒性;对误配准进行强化训练的 E-Fisherface 方法同样可以大大提高识别系统的鲁棒性。

7.2.5 初步探讨了人脸识别应用系统设计问题

核心识别算法的性能当然是实用人脸识别应用系统最终识别性能的决定性因素,但系统设计是否合适同样在很大程度上影响着应用系统的成败,尤其在人脸识别技术尚未完全成熟的今天,系统设计的作用更加突出,良好的系统设计可能起到事半功倍的效果,而拙劣的系统设计可能会葬送最好的核心识别算法的前途。鉴于此,本文初步探讨了人脸识别应用系统设计中需要考虑的一些共性关键问题,如现场环境设计、摄像设备选择与安装、注册原型图像的挑选、核心算法选择标准、学习集选择、模型在线更新、活体判别策略等。最后,基于前面章节提出的核心技术和关键问题解决方案,给出了人脸照片比对系统、模拟的 XP 用户自动登录系统和机场登机旅客身份验证系统三个典型 AFR 应用原型系统的设计方案。

本论文所提出的若干关键技术和关键问题解决方案已经获得了初步的实际应用,取得较好的效果。

7.3 后续工作

人脸识别是一项极具挑战性的研究课题,本论文仅针对其中的几个关键问题开展了研究,但即使针对这几个问题也还有很多工作需要进一步深入,主要包括:

（1）在面部特征配准方面的下一步工作

由图 3.8 的结果可以看出，本文方法对于嘴部、眼睛区域的定位效果有时候不尽人意，尤其是面部表情比较夸张的情况下，LTC-AAM 还难以正确收敛，这一方面可能因为目前采用的形状模型训练样本不够丰富，但我们认为这在很大程度上也与 ASM、AAM 对于局部纹理的建模不够精细造成的。因此，本文后续工作中将尝试对嘴角、眼角等关键的控制点用更为复杂的局部模型进行建模，以便能够得到更为精确的关键特征配准。一个典型的思路是利用弹性图匹配中的位移估计方法来对关键的控制点进行局部搜索，并借以约束其他非控制点的局部搜索。

目前的 LTC-AAM 中，所有特征点的局部匹配信度是均匀设置的，因此，个别点局部搜索错误过分严重的时候往往会使得参数调整后的结果不得不为这些个别的严重错误做出牺牲，从而减缓搜索过程。本文另一后续工作将要考虑的是对不同的特征点位置设置不同的信度，信度的估计则包括两点：（1）先验知识；（2）局部模型匹配的相似度。

另外，目前 LTC-AAM 对于姿态和光照的适应能力还比较差，如何将其扩展到多姿态人脸特征的自动定位上去，并使之具备良好的光照适应能力将是本文的另一个重要后续工作。

（2）AdaGaDA 人脸识别方法下一步工作

概括而言，AdaGaDA 方法的有三部分的核心内容，即：高维 Gabor 特征集合，作为特征选择器的 AdaBoost 学习算法，基于主成分分析的判别分析技术。本文的后续工作也可以围绕这三部分分别展开。首先，原始特征集合可以进一步扩展包含更多其它特征。尽管实践表明 Gabor 特征是一种非常有效的人脸表示方式，但并不是唯一的，也没有证据表明其是最优的。因此，可以考虑加入更多的表示特征供后续的特征选择器进行选择。其次，特征选择方法也完全可以采用 AdaBoost 之外的其他方法，例如遗传算法等来代替，以寻求更优的 Gabor 判别特征集合。最后，针对分类识别方法，对选择出来的特征进行 Bayesian 决策、核判别分析等也是可以考虑尝试的后续研究工作。

从应用设计的角度来看，进一步将 AdaGaDA 人脸识别方法应用到更多实用系统中，需要考虑在大规模学习集上进行 AdaGaDA 模型训练问题，这涉及到两方面的问题：其一是如何构建一个足够规模、覆盖足够图像变化的学习集；其二是 Gabor 特征表示和 AdaBoost 的学习能力是否能够保证得到收敛的训练模型，并保障得到的识别模型具有良好的新样本推广能力。这也将是我们的另一后续工作。

另一方面，本文的研究还局限于 Gabor 特征在正面人脸识别中的应用，如何将其扩展到多姿态人脸识别上去，或者对 Gabor 特征作相应的调整和适应，也是可以考虑的下一步工作。

（3）误配准灾难问题研究的下一步工作

本文提出的误配准灾难问题还有大量的研究空间可以进一步拓展。如本论文 5.3 小节所总结的，解决误配准问题有四大类方法。本文仅考察了其中的不变特征方法和误配准学习方法，而没有考虑变换不变距离度量方法，也没有考虑这些方法的进一步组合。因此，可以考虑的一个直接的后续工作是基于误配准学习的 AdaGaDA 识别方法，其基本思路是在 AdaGabor 特征学习算法中，通过扰动的方法加入更多的训练样本，尤其是要加入更多的旋转和尺度变化样本，可以预期如此训练得到的 AdaGabor 特征将对误配准具有更强的鲁棒性。此外，我们还将考虑将切距离度量与 E-Fisherface 或 AdaGaDA 结合，从距离度量不变性的角度进一步增强算法对误配准的鲁棒性。

误配准带来的图像表现变化是一种高度非线性的流形（Manifold）数据分布，子空

间法或者切距离法采用线性模型对原型样本点附近的流形结构进行近似，实际上是一种比较粗糙的近似方法。近年来非线性流形学习已经取得了重要的研究进展，如 ISOMAP，LLE 等，这些技术进步为误配准问题的非线性解决方案的提出提供了可能。这将是本文的另一重要后续工作。

（4）应用系统设计方面的下一步工作

本文只是初步考虑了应用系统设计中涉及到的共性关键问题，但构建鲁棒、实用的人脸识别应用系统需要更多的工程技术解决方案和实践开发经验的支持。因此，加强与有关产业部门的合作，尤其是加强与传感器研制部门的合作从信号层面展开工作保障数据的可靠性，也应该是重要的未来工作之一。

最后，需要进一步指出的是，尽管人脸识别并非一个新的研究课题，但仍然是一项充满挑战问题的研究课题（参见本论文 2.7.1 小节），需要投入大量精力对其进行更深入地研究，尤其需要深入研究更好的人脸表示方法以及相应的分类识别算法，也需要针对光照、姿态和老化等关键问题展开专门的系统研究，需要借鉴非线性流形学习、3D 人脸建模等领域的最新进展（参见本论文 2.7.2 小节）。相信经过研究人员深入细致的研究，自动人脸识别问题最终一定能够得到圆满地解决！

参考文献

- [AG96] J.Atick, P.Griffin, N.Redlich, Statistical Approach to Shape from shading: Reconstruction of Three Dimensional Face Surfaces from Single Two Dimensional Image, Neural Computation, Vol 8, pp. 1321-1340, 1996
- [AM97] Y.Adini, Y.Moses, S.Ullman, Face Recognition: The Problem of Compensating for changes in illumination Direction, IEEE Trans. On PAMI, 19(7), pp721-732, 1997
- [AN98] J.B.Are, D.Nandy, A Volumetric/Iconic Frequency Domain Representation for Objects With Application for Pose Invariant Face Recognition, IEEE Transaction on PAMI, 20(5), pp.449-457, 1998. 5
- [AY02] Olugbenga Ayinde, Yee-Hong Yang, Face recognition approach based on rank correlation of Gabor-filtered images, Pattern Recognition 35 (2002) 1275–1289
- [BB03] 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 BANCA database and evaluation protocol. In Audio- and Video-Based Biometric Person Authentication (AVBPA), pages 625–638, 2003.
- [Be95] David Beymer, Vectorizing Face Images by Interleaving Shape and Texture Computation, A.I.Memo No. 1537, 1995.9
- [BH96] 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, ECCV'96, pp45~56, 1996
- [BH97] P. N. Bellhumer, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition, 17(7): 711--720, 1997
- [BH01] A.U.Batur, M.H.Hayes, Linear subspace for illumination robust face recognition, Proc. of CVPR01, 2001
- [边 00] 边肇祺, 张学工等《模式识别》(第二版) 清华大学出版社, 2000.
- [BJ01] R.Basri, D.Jacobs, Lambertian Reflectance and Linear Subspaces, ICCV2001, Beckman Institute. Vol.2, p383-390, 2001
- [BK98] P. Belhumeur and D. Kriegman, What Is the Set of Images of an Object under All Possible Illumination Conditions, Int'l J. Computer Vision, vol. 28, no. 3, pp. 245-260, July 1998.
- [BL90] J.Buhmann, M.Lades, von der Malsburg, Size and distortion invariant object recognition by hierarchical graph matching. In: Proceedings of IEEE Intl. Joint Conference on Neural Networks. pp.411-416, San Diego, 1990
- [BL98] M.S.Bartlett, H.M.Lades and T.J.Sejnowski. "Independent Component Representations for Face Recognition". Proc. of SPIE, 2399, pp.528-539, 1998.
- [BP92] R. Brunelli and T. Poggio, Face Recognition through **Geometrical Features**, Proceedings of ECCV '92, pp792-800
- [BP93] R.Brunelli, T.Poggio, Face recognition: features vs. templates. IEEE Trans. on PAMI, 15(10): 1042-1052, 1993

- [BP95] D. Beymer and T. Poggio, Face Recognition from One Example View, Proc. Int'l Conf. Computer Vision, pp. 500-507, 1995
- [BPo95] D.Beymer, T.Poggio, Image Representations for Visual Learning, Science, Vol.272, pp1905-1909, 1995.
- [BR02] Volker Blanz, Sami Romdhani, and Thomas Vetter, Face Identification across different Poses and Illuminations with a 3D Morphable Model, Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, 2002, pp202-207.
- [BT03] V.Blanz and T.Vetter Face Recognition Based on Fitting a 3D Morphable Model, TPAMI, vol25, no9, pp1063-1075, 2003
- [BV99] Volker Blanz, Thomas Vetter, A Morphable Model For the Synthesis of 3D Faces, SIG'GRAPH'99, 1999
- [CB65] H.Chan and W.W.Bledsoe, A man-machine facial recognition system: some preliminary results, Technical report, Panoramic Research Inc., Cal, 1965
- [CB00] H. Chen, P. Belhumeur, and D. Jacobs, In Search of Illumination Invariants, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 254-261, 2000.
- [CE98] T.F.Cootes, G.J.Edwards, C.J.Taylor, Active Appearance Models, Proc. European Conf. Computer Vision, vol. 2, pp. 484-498, 1998.
- [陈 99] 陈熙霖, 山世光, 高文. 多姿态人脸识别. 中国图像图形学报, 1999, 4(10): 818~824.
- [CL93] G.Chow, X.B.Li Towards a system for automatic facial feature detection. Pattern Recognition, 26(12), pp1739-1755. 1993
- [CL00] 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, vol. 33, no.10, pp. 1713-1726, 2000.
- [CM02] R. Cappelli, D. Maio and D. Maltoni, "Subspace classification for face recognition", in proceedings Workshop on Biometric Authentication - ECCV'02 (BIOW2002), Copenhagen, pp.133-141, June 2002
- [CS02] Bo Cao, Shiguang Shan, Wen Gao, Debin Zhao, "Localizing the Iris Center by Region Growing Search," Proceedings of IEEE International Conference on Multimedia and Expo, Vol. 2, pp129-132, 2002.
- [CT92] I.Craw, D.Tock, & A.Bennett, "Finding face features", In Proceedings 2nd European Conference on Computer Vision, pp92-96, Italy, 1992
- [CT95] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, Active Shape Models - Their Training and Application., Computer Vision and Image Understanding, Vol. 61, No. 1, January, pp. 38-59, 1995.
- [CW95] R.Chellappa, C.L.Wilson, Saad Sirohey, Human and Machine Recognition of faces: A survey, Proceedings of the IEEE, vol.83, no.5, 1995.5
- [CW00] T.F.Cootes, K.Walker, C.J.Taylor, View-based Active Appearance Models, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp227-232, Grenoble, France, 2000.3
- [CW02] Jen-Tzung Chien, Chia-Chen Wu, Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition, IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 24, NO. 12, pp1644-1649, DECEMBER 2002

- [Dau85] J.G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," J. Optical Soc. Amer., vol. 2, no. 7, pp. 1,160-1,169, 1985.
- [DC01] O.Deniz, M.Castrillon, M.Hernandez. "Face Recognition Using Independent Component Analysis and Support Vector Machines". Int. Conf. on Audio- and Video-Based Person Authentication. Lecture Notes, pp. 59-64, 2001.
- [邓 03] 邓刚, 曹波, 苗军, 高文, 赵德斌, 基于支持向量机眼动模型的活性判别算法, 计算机辅助设计与图形学学报, 2003
- [DG01] Hyoja-Dong, and Nam-Gu. Asian Face Image Database PF01. Pohang University of Science and Technology.
- [DH03] R.O.Duda, P.E.Hart, D.G.Stork 著, 李宏东, 姚天祥等译, 模式分类, 机械工业出版社, 2003
- [丁 02] 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, Vol.30, No.9: 1292-1294, 2002
- [EC99] G. Edwards, T. Cootes, and C. Taylor, Advances in Active Appearance Models, Proc. Int'l Conf. Computer Vision, pp. 137-142, 1999.
- [FG02] R. Feris, J. Gemmel, K. Toyama and V. Krueger, "Hierarchical Wavelet Networks for Facial Feature Localization", Proc. Int'l Conf. Automatic Face and Gesture Recognition, 2002.
- [FS97] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci., 55(1):119--139, 1997.
- [FS99] Y.Freund and R. Schapire, "A short introduction to boosting," Journal of Japanese Society for Artificial Intelligence, vol. 14, no. 5, pp. 771-780, 1999.
- [FY00] G.C.Feng, P.C.Yuen, J.H.Lai, Virtual View face image synthesis using 3D Spring-based Face Model from a single Image, Fourth IEEE Face & Gesture Recognition 2000, 530-535
- [FY01] G.Feng, P.C.Yuen, Multi-cues eye detection on gray intensity images, Pattern Recognition, 34 (5), 2001
- [GB00] A.S.Georghiades, P.N.Belhumeur, D.J.Kriegman, From Few to Many: Generative Models for Recognition Under Variable Pose and Illumination, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp277-284, Grenoble, France, 2000.3
- [GB01] Athinodoros S.Georghiades, Peter N.Belhumeur, David J.Kriegman, From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose, IEEE Trans. on PAMI, 23(6) pp.643660-139, 2001.6
- [GC04] Wen Gao, Bo Cao, Shiguang Shan, DeLong Zhou, Xiaohua Zhang, Debin Zhao. The CAS-PEAL Large-Scale Chinese Face Database and Evaluation Protocols. Technical Report No. JDL_TR_04_FR_001, Joint Research & Development Laboratory, CAS, 2004. <http://www.jdl.ac.cn>
- [GH71] A.J.Goldstein, L.D.Harmon, and A.B.Lesk. Identification of human faces. Proceedings of the IEEE, 1971, 59(5): 748~760.
- [GK98] A.S.Georghiades, D.J.Kriegman and P.N.Belhumeur, "Illumination Cones For Recognition Under Variable Lighting: Faces". Proc. of IEEE CVPR, pp52-58, 1998
- [GL96] F.Goudail, E.Lange, T.Iwamoto, K.Kyuma, and N.Otsu, Face recognition system

- using local auto-correlation and multi-scale integration, IEEE Trans. Pattern Anal. Machine Intell. 18(10), 1996, 1024-1028
- [GL00] 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 Recog., pp.196-201, Grenoble, 2000.3
- [GL02] Y.Gao and M.Leung, "Face Recognition Using Line Edge Map", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No.6, pp.764-779, 2002.
- [GM02] R. Gross, I. Matthews, and S. Baker. Eigen light- fields and face recognition across pose. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, May 2002.
- [Gr00] Grudin, "On Internal Representations in Face Recognition Systems," Pattern Recognition, vol. 33, 2000
- [GZ01] G. D. Guo, H. J. Zhang, and S. Z. Li, Pairwise Face Recognition, IEEE Int'l Conf. On Computer Vision, Vol 2, pp.282-287, 2001.
- [Ha94] P.W.Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, Seattle, WA, pp995~998, 1994
- [Ha95] P. Hallinan, A Deformable Model for Face Recognition under Arbitrary Lighting Conditions, PhD thesis, Harvard Univ., 1995
- [HB97] Thien M. Ha and Horst Bunke, Off-Line, Handwritten Numeral Recognition by Perturbation Method, IEEE Transactions On Pattern Analysis and Machine Intelligence, VOL. 19, NO. 5, pp535-539, MAY 1997
- [HB03] B.-W. Hwang, H. Byun, M.-C. Roh, and S.-W. Lee. Performance evaluation of face recognition algorithms on the asian face database, KFDB. In Audio- and Video-Based Biometric Person Authentication (AVBPA), pages 557~565, 2003.
- [HC92] C.L.Huang, C.W.Chen, "Human facial feature-extraction for face interpretation and recognition". Pattern Recognition, 25(12), pp1435-1444, 1992
- [HH77] L.Harmon and W. Hunt, Automatic recognition of human face profile, Computer Graphic and Image Process, vol. 6, pp.135-156, 1977
- [HJ01] R.-L. Hsu and A.K. Jain, "Face modeling for recognition," IEEE Int'l Conf. Image Processing, vol. II, pp. 693-696, Oct. 2001
- [HK81] L.Harmon, M.K. Khan et al. Machine identification of human faces. Pattern Recognition, 13(2):97~110. 1981
- [HL01] X.W. Hou, S.Z. Li, H.J. Zhang, Q.S. Cheng. "Direct Appearance Models". In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Hawaii. December, 2001.
- [Ho91] Z.Hong. Algebraic feature extraction of image for recognition. Pattern Recognition, 24(3),pp211-219, 1991
- [洪 92] 洪子泉, 杨静宇. 用于图像识别的图像代数特征抽取. 自动化学报, 1992, 18(2): 233~237.
- [洪 94] 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.
- [黄 98] 黄修武, 基于代数方法的人脸图像特征提取与识别, 南京理工大学博士论文, 1998

- [HZ00] Fujie Huang, Zhihua Zhou, Hongjiang zhang, Tsuhan Chen, Pose Invariant Face Recognition, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp245-250, Grenoble, France, 2000.3
- [JD00] A.K.Jain, Robert P.W.Duin, Jianchang Mao. Statistical Pattern Recognition: A Review, IEEE Trans. On PAMI, 22(1): 4-37, 2000.1
- [荆 98] 荆晓远,模式识别分类技术在人脸识别中的应用,南京理工大学博士论文,1998
- [JK02] K. Jonsson, Josef Kittler, Yongping Li, Jiri Matas: Support vector machines for face authentication. Image and Vision Computing. 20(5-6): 369-375 (2002)
- [JM00] 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, pp208-213, Grenoble, France, 2000.3[SVM]
- [Ka73] T. Kanade. Picture processing system by computer and recognition of human faces. Ph.D Dissertation. Kyoto Kyoto University, 1973.
- [KB76] Gerald J. Kaufman and Kenneth J.Breeding. The automatic recognition of human faces from profile silhouettes. IEEE Transactions on Systems, Man, and Cybernetics, 6(2), pp113-121, 1976.2
- [KM04] Daniel Keysers, Wolfgang Macherey, Hermann Ney, Jorg Dahmen, Adaptation in Statistical Pattern Recognition Using Tangent Vectors. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 26, Number 2, pages 269-274, 2004.2
- [Kr00] V. Krueger. Gabor wavelet networks for object representation. . DAGM Symposium, Kiel, Germany, 9, 13-15, 2000.
- [KS90] M.Kirby and L.Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces, IEEE Trans. PAMI, 12(1), 103-108, 1990
- [KT00] C.L.Kotropoulos, A.Tefas, IPitas. Frontal Face Authentication Using Discriminating Grids with Morphological Feature Vectors, IEEE trans. On Multimedia, Vol.2, No.1, pp14-26 March, 2000
- [KW88] M.Kass, A.Witkin, and D.Terzopoulos. Snakes: Active contour models. Int. Journal of Computer Vision, pp.321-331,1988
- [赖阮 99] 赖剑煌,阮帮志,冯国灿,频谱脸:一种基于小波变换和 Fourier 变换的人像识别新方法,中国图像图形学报,(A 版),4(10),811-817,1999
- [赖 99] 赖剑煌,小波分析、图像模式识别及其在人像识别上的应用,中山大学博士论文,1999
- [LB00] A. Leonardis, Horst Bischof. Robust recognition using Eigenimages. Computer Vision and Image Understanding, 2000, 78:99~118.
- [LC93] K.Liu, Y.Cheng, J.Yang. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. Pattern Recognition, 26(6): 903~911, 1993
- [Lee96] T. S. Lee. Image Representation Using 2d Gabor Wavelets. IEEE Trans. Pattern Analysis and Machine Intelligence, 18(10):959--971, 1996
- [LG97] Steve Lawrence, Lee Giles, Ah Chung Tsoi, Andrew D. Back, Face Recognition: A Convolutional Neural-Network Approach, IEEE Trans. On Neural Network, Vol.8, No.1, Jun. 1997
- [LH01] K.C.Lee, J.Ho, D.Kriegman, Nine points of Light: Acquiring Subspaces for Face Recognitjoin under variable lighting, CVPR01

- [LH02] Q.S.Liu, R.Huang, H.Q.Lu and S.D.Ma, "Face Recognition Using Kernel Based Fisher Discriminant Analysis", Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 197-201, 2002.
- [LH03] K.Lee, J.Ho, M.Yang, D.Kriegman, Video-Based Face Recognition Using Probabilistic Appearance Manifolds, IEEE Conf. on Computer Vision and Pattern Recognition, 2003, vol.1, pp.313-320
- [廖 03] 廖频, 基于统一概率模型的人脸识别技术, 中科院计算所博士论文, 2003 年
- [梁 01] 梁路宏, 人脸检测与跟踪研究, 清华大学工学博士学位论文, 2001
- [梁 02] 梁路宏, 艾海舟, 人脸检测研究综述, 计算机学报, 25 (5): pp.449-458, 2002.
- [Liu99] C.Liu. Statistical and evolutionary approaches for face recognition. George Mason University, Ph.d Dissertation, 1999.
- [Liu03] C. Liu: "A Bayesian Discriminating Features Method for Face Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no. 6, pp. 725-740, 2003.
- [刘 97] 刘明宝, 人脸检测与跟踪的研究, 哈尔滨工业大学工学博士论文, 1997
- [刘 03] 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 年
- [LJ04] Stan Li, Anil Jain (Ed). Handbook of Face Recognition, Springer, 2004
- [LL99] S. Li and J. Lu, Face Recognition Using Nearest Feature Line, IEEE Trans. Neural Networks, vol. 10, no. 2, pp. 439-443, Mar. 1999.
- [LO95] J. Lampinen and E. Oja. Distortion tolerant pattern recognition based on self-organizing feature extraction. IEEE Transactions on Neural Networks, 6(3):539-- 547, 1995.
- [LR03] Mun Wai Lee, Surendra Ranganath, Pose-invariant face recognition using a 3D deformable model, Pattern Recognition, Volume 36, Issue 8, pp. 1835-1846, 2003
- [LT94] A. Lanitis, C. J. Taylor, and T. F. Cootes. An automatic face identification system using flexible appearance models. In British Machine Vision Conference, BMVA Press, 1994, 1:65~74.
- [LT95] A. Lanitis, C. Taylor, and T. Cootes, A Unified Approach to Coding and Interpreting Face Images, Proc. Int'l Conf. Computer Vision, pp. 368-373, 1995.
- [LT97] A. Lanitis, C. Taylor, and T. Cootes. Automatic interpretation and coding of face images using flexible models. IEEE Transactions on Pattern Analysis & Machine Intelligence, vol.19, no.7, pp.743-56, July 1997
- [卢 98] 卢春雨, 人脸自动识别若干问题研究与系统实现, 清华大学博士论文, 1998
- [卢 99] 卢春雨, 张长水, 基于区域特征的快速人脸检测法, 清华大学学报, vol.39, 1999.1
- [LV93] M.Lades, J.C.Vorbruggen, J.Buhmann, J.Lange, C.v.d.Malsburg, R.P.Wurtz, W.Konen, Distortion Invariant Object Recognition in the Dynamic Link Architecture, IEEE Trans. On Computers, 42(3), pp 300-311, 1993
- [LW99] C.J.Liu and H. Wechsler. "Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition". Int. Conf. on Audio and Video Based Biometric Person Authentication, 1999.
- [LW00] C.Liu and H.Wechsler, "Evolutionary pursuit and its application to Face Recognition", IEEE trans. On PAMI, Vol.22, No. 6, pp570-582, 2000.6
- [LW01] C. Liu and H. Wechsler: "A Shape and Texture Based Enhanced Fisher Classifier for Face Recognition", IEEE Trans. Image Processing, vol. 10, no. 4, pp. 598-608, 2001.
- [LW02] C. Liu and H. Wechsler: "Gabor Feature Based Classification Using the Enhanced

- Fisher Linear Discriminant Model for Face Recognition", IEEE Trans. Image Processing, vol. 11, no. 4, pp. 467-476, 2002.
- [LW03] C. Liu and H. Wechsler: "Independent Component Analysis of Gabor Features for Face Recognition", IEEE Trans. Neural Networks, vol. 14, no. 4, pp. 919-928, 2003.
- [LY96] K.M. Lam and H. Yan, "Locating and Extracting the Eye in Human Face Images", Pattern Recognition, Vol. 29, No. 5, pp. 771-9, 1996
- [LY97] K.M.Lam, H.Yan, An Analytic-to-Holistic Approach for Face Recognition Based on a single Frontal view", IEEE trans. On PAMI, vol.20, no.7, pp. 673-686, July, 1997
- [LY99] Jianhuang Lai, P.C.Yuen, G.C.Feng, Spectroface: A Fourier-based Approach for Human Face Recognition, ICMI, Vol.2, 115-120, 1999
- [LY01] Jianhuang Lai, P.C.Yuen, G.C.Feng, Face recognition using Holistic Fourier Invariant Features", Pattern Recognition, 34(1), 95-109, 2001
- [LY02] S.Z. Li, S.C. Yan, H.J. Zhang, Q.S. Cheng. "Multi-View Face Alignment Using Direct Appearance Models". Proceedings of The 5th International Conference on Automatic Face and Gesture Recognition., pp324-329, Washington, DC, USA. 20-21, May, 2002.
- [LZ04] Stan Z. Li, ZhenQiu Zhang. "FloatBoost Learning and Statistical Face Detection". IEEE Transactions on Pattern Analysis and Machine Intelligence, Accepted, 2004.
- [Ma02] A.M.Martinez. Recognizing Imprecisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 6, pp. 748-763, 2002
- [MA94] Y.Moses, Y.Adini and S.Ullman. Face recognition: The problem of compensating for changes in illumination direction. In Proc. the Third European Conference on Computer Vision, 1994.
- [MB98] A. R. Martinez and R. Benavente. The AR face database. Technical Report 24, Computer Vision Center(CVC) Technical Report, Barcelona, Spain, June 1998.
- [MC92] B.S.Manjunath, R. Chellappa, C.v.d.Malsburg, A Feature based approach to face recognition, in Pro. IEEE Computer Soc. Con. On Computer Vision and Pattern Recognition, 1992, pp373-378
- [MK01] AM Martinez and AC Kak. PCA versus LDA. Trans. PAMI, 23(2), 228-233, 2001
- [MM99] K. Messer, J. Matas, J. Kittler, J. Luetlin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In Second International Conference on Audio and Video-based Biometric Person Authentication, March 1999.
- [MN95] H.Murase, S.Nayar, Visual Learning and recognition of 3D object from appearance, Int. J. Computer Vision, 14:5-24, 1995
- [Mo02] B. Moghaddam, Principal manifolds and probabilistic subspaces for visual recognition, IEEE Trans. PAMI 24(6) 780-788. 2002
- [MP95] B. Moghaddam and A. Pentland, Probabilistic Visual Learning for Object Detection, Proc. Int'l Conf. Computer Vision, pp. 786-793, 1995.
- [MP97] B. Moghaddam and A. Pentland, Probabilistic Visual Learning for Object representation, IEEE trans. On PAMI, vol.20, no.7, pp696-710, July, 1997
- [MP00] Baback Moghaddam, Tony Jebara, Alex Pentland, Bayesian Face Recognition, Pattern Recognition Vol.33(2000), pp1771-1782, 2000
- [MY99] J. Miao, B.C. Yin, K.Q. Wang, et al, A hierarchical multiscale and multiangle system

- for human face detection in a complex background using gravity-center template, Pattern Recognition, 32(7), 1999.
- [Ne99] A.Nefian, “A Hidden Markov Model-Based Approach for Face Detection and Recognition”, Ph.D. Thesis, Georgia Institute of Technology, 1999.
- [NM91] O.Nakamura, S.Mathur, and T.Minami, Identification of human faces based on isodensity maps. Pattern Recognition, 24(3), pp263-272, 1991
- [NM94] S.K.Nayar and H.Murase, “Dimensionality Of Illumination Manifold In Eigenspace”, TR-CUCS-021-94, Columbia University, 1994
- [PA96] P.Penev and J.Atick, “Local Feature Analysis: A General Statistical Theory for Object Representation,” Network: Computation in Neural Systems, vol.7, pp.477-500, 1996
- [彭 97] 彭振云、游素亚、徐光佑, 允许姿态变化的快速人脸特征检测, 《中国图象图形学报》 Vol.2, No.4, pp.225-229, 1997.4.
- [彭 97] 彭辉、张长水、荣钢、边肇祺, 基于 K-L 变换的人脸自动识别方法, 《清华大学学报》(自然科学版), Vol.37, No.3, pp.67 - 70, 1997.3.
- [彭 98] 彭辉, 人脸自动检测与识别的方法研究, 清华大学自动化系工学博士论文, 1998
- [PG03] 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, NISTIR 6965, National Institute of Standards and Technology, 2003, <http://www.frvt.org>.
- [Ph98] P. J. Phillips. Support vector machines applied to face recognition. In Advances in Neural Information Processing Systems, page 803. Editors: M.C.Mozer, M.I.Jordan, and T. Petsche, MIT Press, 1998
- [PL92] S. J. Perantonis and P. J. G. Lisboa, “Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifiers,” IEEE Trans. Neural Networks, vol. 3, pp. 241–251, Mar. 1992.
- [PM94] Alex Pentland, Baback Moghaddam, Thad starner. View-Based and Modular Eigen-spaces for Face Recognition MIT Media Lab perceptual Computing Section Tech IEEE Conference on Computer Vision & Pattern Recognition 1994
- [PM00] P.J.Phillips, H.Moon, etc. “The FERET Evaluation Methodology for Face-Recognition Algorithms”, IEEE Transactions on PAMI, Vol.22, No.10, pp1090-1104, 2000
- [PR00] M.Pantic, L.J.M.Rothkrantz, “Automatic Analysis of Facial Expressions: The State of the Art”, IEEE Transactions on PAMI, Vol.22, No.12, pp1424-1445, 2000.12
- [Ra02] R.Ramamoorthi, Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object, IEEE PAMI, 24(10), pp 1322-1333, Oct 2002
- [RB98] H.A.Rowley, S.Baluja, and T. Kanade, Neural network-based face detection, IEEE-PAMI, 20(1), 23-38, 1998.
- [RB02] Sami Romdhani, Volker Blanz, Thomas Vetter, Face Identification by Fitting a 3D Morphable Model Using Linear Shape and Texture Error Functions, Proceedings of the 7th European Conference on Computer Vision, Vol.4, pp3-19, May. 2002, Editor A.Heyden et al.
- [RL00] Sam Roweis & Lawrence Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, v.290 no.5500, Dec.22, 2000. pp.2323--2326.

- [Sa93] F.Samaria, Face segmentation for identification using Hidden Markov Models. Proceedings of the of 4th British Machine Vision Conference, Springer-Verlag. 1993
- [Sa94] F.Samaria. Face Recognition Using Hidden Markov Models. PhD thesis, Trinity College, University of Cambridge, Cambridge, 1994.
- [SA03] G.Song; H.Ai; G.Xu; L.Zhuang, Automatic Video Based Face Verification and Recognition by Support Vector Machines, Proceedings of SPIE - The International Society for Optical Engineering, v 5286, n 1, 2003, p 127-132
- [SB03] Terence Sim, Simon Baker, and Maan Bsat. The CMU Pose, Illumination and Expression Database. IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 25, No. 12, December 2003, pp. 1615 – 1618
- [SC98] P. Simard, Y. Le Cun, J. Denker, and B. Victorri. Transformation invariance in pattern recognition --- tangent distance and tangent propagation. In Genevieve Orr and Klaus-Robert Muller, editors, Neural networks: tricks of the trade, volume 1524 of Lecture Notes in Computer Science, pages 239--274, Heidelberg, 1998. Springer.
- [SC00] Bernt Schiele, James L.Crowley, Recognition without Correspondence using Multidimensional Receptive Field Histograms, International Journal of Computer Vision, 36(1), pp31-50, 2000
- [SC02] S.Shan, B.Cao, W.Gao, D.Zhao, Extended Fisherface For Face Recognition From A Single Example Image Per Person, IEEE International Symposium on Circuits and Systems, 2002 (IEEE ISCAS2002) , Vol.2, pp81-84, 2002
- [SC04] S.Shan, Y.Chang, W.Gao, B.Cao, Curse Of Mis-Alignment In Face Recognition: Problem And A Novel Mis-Alignment Learning Solution, To appear in Proceeding of The 6th International Conference on Face and Gesture Recognition, Korea, May, 2004
- [SE98] J.Steffens, E.Elagin, and H.Neven. Personspotter – face and robust system for human detection, tracking and recognition, in IEEE International conference on Face & Gesture Recognition, pp516-521, Nara, Japan, 1998.4
- [SG00] S.Shan, W.Gao, Xilin Chen, Jiyong Ma. Novel Face Recognition Based on Individual Eigen-subspaces, Proceedings of ICSP, Vol. III, pp1522-1525, 2000
- [山 01] 山世光, 高文, 陈熙霖 基于纹理分布和变形模板的面部特征提取, 软件学报, Vol.12, No.4, pp570-577, 2001.4
- [SG03] S.Shan, W.Gao, D.Zhao, Face Identification Based On Face-Specific Subspace, International Journal of Image and System Technology, 13(1), pp23-32, 2003
- [SGW03] Shiguang Shan, Wen Gao, Wei Wang, Debin Zhao, Baocai Yin. Enhanced Active Shape Models with Global Texture Constraints for Face Image Analysis, Fourteenth International Symposium On Methodologies For Intelligent Systems, N.Zhong et al. (Eds.): ISMIS 2003, LNAI2871, pp593-597, Springer, Maebashi City, Japan, Oct.2003
- [SG04] S.Shan, W.Gao, Y.Chang, B.Cao, P.Yang, Review the Strength of Gabor features for face recognition from the angle of its robustness to mis-alignment, To appear in Proceeding of International Conference on Pattern Recognition, 2004
- [Sh92] A.Shashua, Geometry and Photometry in 3D Visual Recognition, PhD thesis, Massachusetts Inst. of Technology, Cambridge, 1992.
- [SH94] F. S. Samaria, and A. C. Harter. Parameterisation of a Stochastic Model for Human

- Face Identification. Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, December 1994.
- [Sh97] A.Shashua, On Photometric Issues in 3D visual recognition from a single 2D Image, International Journal of Computer Vision, 21(1/2), 99-122, 1997
- [SH02] M.Sonka, V.Hlavac, R.Boyle, Image Processing, Analysis and Machine Vision. Second Edition, Published by Brooks/Cole, 2002
- [SI92] Ashok Samal, Prasana A. IyenGar. Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey, Pattern Recognition, 25(1), pp65-77, 1992
- [SK87] L.Sirovich and M. Kirby, Low-dimensional Procedure for the characterization of Human Faces. J. Opt. Soc. Am. A, Vol4, No3, 519-524, 1987
- [SP98] K.K. Sung and T. Poggio, Example-based learning for view-based human face detection, IEEE-PAMI, 20(1), 39-50, 1998.
- [SR01] A.Shashua and T.Riklin-Raviv, "The Quotient Image: Class-Based Re-Rendering And Recognition With Varying Illuminations", IEEE Trans. on PAMI, 23(2): pp129-139, 2001.2
- [SR02] R.E. Schapire, M. Rochery, M. Rahim, and N. Gupta. Incorporating prior knowledge into boosting. In Machine Learning: Proceedings of the Nineteenth International Conference, 2002.
- [St00] M.B.Stegmann, Active Appearance Models, Theory, Extension and Cases, Master thesis, Technical University of Denmark, 2000
- [SW96] Daniel L. Swets, John Weng. Using discriminant eigenfeatures for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8):831~836.
- [SY94] F.Samaria, & S.Young, HMM-based architecture for face identification International Vision Conference, 12(8), 537-543. 1994
- [TK01] A.Tefas, C.Kotropoulos, and IPitas, Using Support Vector Machines to Enhance the Performance of Elastic Graph Matching for Frontal Face Authentication, IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 23, NO. 7, pp735-746, JULY 2001
- [TP91] M.Turk, A.Pentland. Eigen-faces for Recognition Journal of cognitive neuroscience, 3(1), pp71-86, 1991
- [TS00] J.B.Tenenbaum, V.de Silva and J.C.Langford , A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science 290 (5500): 2319-2323, 22 December 2000
- [TT03] Y.Tian, T.Tan and YWang, Do Singular Values Contains Adequate Information for Face Recognition? Pattern Recognition, Vol. 36, No.3, pp.649-655, 2003.
- [VA94] D.Valentin, H.Abdi, A.J.O'Toole, Garrison W.Cottrell, Connectionist Models of Face Processing: A survey, Pattern Recognition. Vol.27, pp1209-1230, 1994
- [Ve98] Thomas Vetter, Synthesis of Novel Views from a Single Face Image, IJCV, 28(2), pp103-116, 1998
- [VJ01] Paul Viola, M.Jones, Rapid Object Detection using a Boosted Cascade of Simple, CVPR-2001
- [VJo01] Paul Viola, M.Jones, Fast and Robust Classification using Asymmetric, NIPS-2001,
- [VP97] T.Vetter, Tomaso Poggio, Linear Object Classes and Image Synthesis From a Single

- Example Image, IEEE Trans. On PAMI, Vol.19, No. 7, pp733-742, 1997
- [WA03] Tong Wang, Haizhou Ai, Gaofeng Huang, A Two-Stage Approach to Automatic Face Alignment, Proceedings of SPIE - The International Society for Optical Engineering, v5286, n2, 2003, p558-563
- [王 00] 王蕴红 谭铁牛 朱勇, "基于奇异值分解和数据融合的脸像鉴别" 《计算机学报》, Vol.20, No.3, pp649-653, 2000.
- [WF97] L.Wiskott, J.M.Fellous, N.Kruger, C.v.d.Malsburg, Face Recogniton by Elastic Bunch Graph Matching, IEEE Trans. On PAMI, Vol.19, No. 7, pp775-779, 1997
- [WH01] Philip L. Worthington, Edwin R. Hancock, Object Recognition Using Shape-from-Shading, IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 23, NO. 5, pp535-542, MAY 2001
- [WK01] Toru Wakahara, Yoshimasa Kimura, Akira Tomono, Affine-Invariant Recognition of Gray-Scale Characters Using Global Affine Transformation Correlation, IEEE Trans. on Patten Analysis and Machine Intelligence, Vol.23, No.4, pp384-394, 2001.4
- [WL04] Haitao Wang, Stan Z. Li, and Yangsheng Wang. "Generalized Quotient Image". In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Washington, DC. 27th June - 2nd July, 2004.
- [WS02] Wei Wang, Shiguang Shan, Wen Gao, Bo Cao. An Improved Active Shape Model For Face Alignment, The fourth International Conference on Multi-modal Interface, IEEE ICMI'2002, Pittsburgh, USA, pp523-528, Oct. 2002
- [WT04] X. Wang, and X. Tang, "A Unified Framework for Subspace Face Recognition," To appear in IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2004.
- [Wur97] R.P.Wurtz, Object Recognition Robust Under Translations, Deformations, and Changes in Background, IEEE Trans. On PAMI, Vol.19, No.7, pp769-775, 1997
- [WY01] Yanjiang Wang, Baozong Yuan: A novel approach for human face detection from color images under complex background. Pattern Recognition 34(10): 1983-1992 (2001)
- [刑 00] 刑昕, 汪孔桥, 沈兰荪, 基于器官跟踪的人脸实时跟踪方法, 电子学报, 28(6), 29-31, 2000
- [XS94] X.Xie, R.Sudhakar, H.Zhuang, Improving eye feature-extraction using deformable templates. Pattern Recognition, 27(6), 791-799, 1994
- [YA00] M.H.Yang, N. Ahuja, D. Kriegman Face Recognition Using Kernel Eigenfaces, Int. Conf. on Image Processing, 2000, vol. 1, pp. 37-40
- [Ya02] M.H.Yang. "Kernel Eigenfaces vs Kernel Fisherfaces: Face Recognition Using Kernel Methods". Proc. Int'l. Conf. Automatic Face and Gesture Recognition, pp. 215-210, 2002.
- [杨 02] 杨健, 涂庆华, 杨静宇 "快速 Foley-Sammon 鉴别变换及脸像鉴别" 中国图象图形学报, Vol. 7, No. 1, pp.1-5, 2002.
- [杨 03] 杨琼, 丁晓青, 对称主分量分析及其在人脸识别中的应用, 计算机学报, 2003
- [YH92] A.L.Yuille, P.W.Hallinan, D.S.Cohen, Feature extraction from faces using deformable templates. International Journal of Computer Vision, 8, 99-111. 1992
- [YK02] Ming-Hsuan Yang, David J.Kriegman, Narendra Ahuja, Detecting Faces in Images: A Survey, IEEE PAMI, vol. 24, No.1, pp34-58, Jan.2002

- [YL03] S.C. Yan, C. Liu, S.Z. Li, H.J. Zhang, H. Shum, Q.S. Cheng. "Face Alignment Using Texture-Constrained Active Shape Models". Image and Vision Computing. Pages 69-75. Vol.21, Issue 1, 10 January 2003.
- [游 96] 游素亚, 张永越, 李武军, 徐光佑, 一种基于多视点图像的可变姿态人脸识别系统, 中国图像图形学报, 1 (3), 1996.7
- [YS97] A. Yuille and D. Snow, Shape and Albedo from Multiple Images Using Integrability, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 158-164, 1997.
- [YS04] P.Yang, S.Shan, W.Gao, Stan Li, D.Zhang, Face Recognition Using Ada-Boosted Gabor Features, To appear in Proceeding of the 6th International Conference on Face and Gesture Recognition, Korea, May, 2004
- [Yu91] A.L.Yuille, Deformable templates for face detection, J. Cogn. neurosci. 3, 59-70, 1991.
- [YY00] J.Yang, Y.Yu, W.Kunz. "An Efficient LDA Algorithm for Face Recognition". Proc. Int'l Conf. Control, Automation, Robotics and Vision, 2000.
- [YY01] H. Yu and J. Yang, "A Direct LDA Algorithm for High-dimensional Data - with Application to Face Recognition". Pattern Recognition, Vol. 34, pp 2067-2070, 2001.
- [YY02] J.Yang and J.Yang, From image vector to matrix: a straightforward image projection technique IMPCA vs. PCA Pattern Recognition, 35 (9) (2002) pp. 1997-1999
- [张 03] 张军平, 流形学习及应用, 中科院自动化所博士学位论文, 2003
- [张 96] 张辉, 周洪祥, 何振亚. 基于主元分析神经网络的人脸特征提取及识别研究. 模式识别与人工智能, 1996, 9 (1): 52~58.
- [张 00] 张翠平, 苏光大 人脸识别技术综述, 中国图象图像学报, Vol.5, No.11, pp885-894, Nov.2000
- [张徐 00] 张欣, 徐彦君, 杜利民, 彩色图像中主要人脸特征位置的全自动标定, 中国图像图形学报, 2000, 5 (2), 138-142
- [ZC98] 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, pp.336-341, 1998
- [ZC00] W.Zhao, R.Chellappa, SFS Based View Synthesis for Robust Face Recognition, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp285-292, Grenoble, France, 2000.3
- [ZCR00] W.Zhao, R.Chellappa, A.Rosenfeld and P.J.Phillips. "Face Recognition: A Literature Survey". CS-Tech Report-4167, University of Maryland, 2000.
- [ZG03] Yi Zhou, Lie Gu, Hongjiang Zhang, Bayesian Tangent Shape Model: Estimating Shape and Pose Parameters via Bayesian Inference, Proceeding of The IEEE Conference on Computer Vision and pattern Recognition (CVPR 2003), Wisconsin, USA, June 16-22, 2003
- [Zh99] W.Zhao, Robust Image based 3D face recognition, Ph. D Thesis, 1999, UMD
- [赵 02] 赵嘉莉, 不同光照、表情下的神经网络人脸识别算法研究, 北方交通大学博士学位论文, 2002
- [周 00] 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
- [周 01] 周志华, 皇甫杰, 张宏江, 陈祖翰. 基于神经网络集成的多视角人脸识别. 计算

- 机研究与发展, 38(10): 1204-1210, 2001
- [ZK03] S.Zhou, V. Krueger, and R.Chellappa, "Probabilistic Recognition of Human Faces from Video," Computer Vision and Image Understanding (special issue on Face Recognition), Vol. 91, pp. 224-245, 2003.
- [ZT99] Ruo Zhang, Ping-Sing Tai, James Edwin Cryer, Mubarak Sha, Shape From Shading: A Survey, IEEE Trans. On PAMI, 21(8). pp690-706, 1999
- [ZY97] Jun Zhang, Yong Yan, Martin Lades, "Face Recognition: Eigenface, Elastic Matching and Neural Nets", Proceedings of the IEEE, vol.85, no. 9, pp1422~1435, Sep. 1997
- [ZZ02] Z.Q. Zhang, L. Zhu, S.Z. Li, H.J. Zhang. "Real-Time Multi-view Face Detection". In Proceedings of The 5th International Conference on Automatic Face and Gesture Recognition. Washington, DC, USA. 20-21 May, 2002.

致谢

首先诚挚地感谢我尊敬的导师高文教授。几年来高老师对我悉心关怀，耐心指导，从博士论文的选题，到日常的工作进展讨论，到本论文的撰写，高老师都给予了我大量的指导和帮助，使得我的研究工作得以顺利的开展，并引领我进入了计算机视觉和模式识别研究领域，踏入了奥妙无穷的科学研究殿堂。高老师不仅仅是严师，还是益友，几年来，高老师不仅在学术上对我进行了大量指导，还在思想和生活等各方面帮助我、关心我、鼓励我，使我得以安心从事科学研究工作，他还信任我带领一个课题组集体从事人脸识别的研究，使我得以更全面地发展，在此一并表示感谢！高老师广博的科学知识，精深的学术思想，高瞻远瞩的方向感，严谨求实的科学精神，宽大包容的胸襟，高标准严要求的工作作风，永远是学生学习的榜样，必将使学生受益终身！

真诚感谢联合实验室的赵德斌教授、陈熙霖教授、黄铁军研究员、贺思敏博士、尹保才教授、刘岩副教授、姚鸿勋教授、崔国勤副研究员、黄庆明研究员、宋益波教授以及谢晓东老师。这些老师都是我在联合实验室工作学习的良师益友，我更能够时刻感受到他们对我的工作、学习、生活诸方面的关心、支持和帮助，以及他们对我更快成长地殷切期望。尤其感谢赵德斌老师对我的学术生涯、工作、生活诸方面无微不至的关怀和帮助，以及对本论文撰写的多次督促。特别感谢陈熙霖教授多年来对我的学术研究、工作方法、日常生活诸方面的指导和帮助，陈教授精深的学术思想、广博的知识面、严谨的科学精神、不惧权威、敢于承担挑战任务的工作作风，都让我佩服有加。感谢贺思敏博士在日常科研工作中与我的经常性讨论，贺老师对科学研究和学术精神的深刻见解让我受益匪浅。感谢崔国勤副研究员相关研究工作对我的启迪。

衷心感谢数字化室主任李锦涛研究员，副主任王兆其研究员和林守勋研究员，以及钱跃良研究员。他们敏锐超前的方向把握能力、严谨求实的科学精神、认真负责的工作态度、积极稳健的进取精神创造了一个朝气蓬勃的数字化研究室！

特别感谢我们的合作伙伴上海银晨智能识别科技有限公司的张青董事长、陈军副总经理和曾文斌博士。多年来，银晨科技对联合实验室进行了大力资助，使得我们能够专心从事人脸识别方面的研究，这与张总、陈总和曾总的大力支持和信任是分不开的。尤其感谢张总努力进取、实干肯干的企业家精神和敢为人先、大力支持人脸识别研发工作的事业心和社会责任感。

尤其感谢微软亚洲研究院的李子青研究员、沈向洋院长和张宏江院长。我曾有幸在李子青老师领导的微软人脸识别研究组学习过一段时间，在此期间李老师给予了我大量耐心细致的指导和帮助，并从此建立起了密切的学术交流与合作关系。李老师对我的教诲使我对于人脸识别及其相关研究课题的理解更上层楼。同时感谢沈向洋院长对我的多次鼓励，感谢张宏江院长对我的赏识。作为领域前辈专家，他们的鼓励给了我更多自信。

感谢香港理工大学的张大鹏教授、香港中文大学的唐晓鸥博士、香港浸会大学的阮邦志博士、桑纳福研究院的赵文义博士、自动化所的卢汉青研究员、王蕴红博士、中山大学的赖剑煌教授、清华大学自动化系的张长水教授、周杰教授等本领域内知名专家学者对我的研究工作的指导和帮助，与他们的数次讨论给了我很多启发。

感谢三星综合技术研究院的赵嘉莉博士、陈茂林博士、任海兵博士和王德军硕士；微软亚洲研究院的胡宇晓硕士、张磊博士、顾烈副研究员、博士生颜水成、周怡硕士和张震球硕士；自动化所的刘青山博士等人，他们既是我的同行，也是我的好朋友，与他

们的学术交流和讨论是我很多研究工作的思路来源。

感谢在联合实验室与我朝夕相处多年的学长或同学们：吴枫博士、马继勇博士、晏洁博士、罗森林博士、周德龙博士、陈益强博士、吕岩博士、孙晓艳博士、段立娟博士、王实博士、庞斌博士、廖频博士以及博士生王瑞、苗军、曾炜和高奎等。尤其感谢吴枫研究员多年来对我的特别关照和热情鼓励！

感谢曾经或仍然在和我一起朝夕相处、并肩战斗的曹波、杨澎、唱轶钲、柴秀娟、张文超、卿来云、张晓华、张洪明、王建宇、陈杰、王巍、富小微、阎盛业、邓刚、焦锋、刘心、赵三强、张宝昌、李月敏、杜波、马丙鹏，他们通过辛勤的劳动为联合实验室的发展做出了贡献，与他们的经常性讨论给了我太多的启发。还要感谢我们的软件工程师王乐珩，他的辛勤工作为我们的系统展示做出了贡献。我尤其要表达对曹波、杨澎、唱轶钲和王巍几位同学的诚挚谢意，作为我的紧密合作者，他们在算法实现和测试方面给予了我大力的帮助和支持，没有与他们的合作、交流和讨论就没有本论文的诞生。

感谢联合实验室办公室的张青老师，羌滨健老师，蔡光辉老师，李丹老师，刘敬老师，董强老师，李伟老师，赵巍老师以及联合信源财务总监张亚冬女士。他们的辛勤劳动为联合实验室提供了良好的研究环境，也为我个人的工作、学习和生活提供了极大的便利和许多帮助，在此表示衷心的感谢。

感谢计算所所人教处的靳晓明老师、宋守礼老师、张晓晖老师、周世佳老师、李琳老师、蔡宏志老师和范文宇老师，在我几年的工作学习生活中，他们给予了我太多的热情的鼓励和大力的支持。尤其感谢靳晓明老师几年来在我工作、生活、学习诸方面细致入微的关怀、激励和教诲。

感谢我的父母、兄嫂等所有亲人对我追求学业、从事科研工作的理解和支持，他们的关爱是我前进的动力源泉。感谢多年来支持我、帮助我的周勇林、王闯、王伟平、万芊、秦跃、李小华、张冬艳、李世炜等好朋友们，他们的友谊让我的生活充满了乐趣。

本论文即将完成之际，恰是我与爱人喜结连理之时。自我们相知相爱以来，她一贯的支持我追求学业、从事我所喜爱的科研工作。她对我的爱、理解和支持，使得该论文最终得以完成。故，谨以此文献给她作为我们新婚的礼物！

作者简历

【作者简历】

姓名：山世光 性别：男 出生日期：1975.3.22 籍贯：山东

1999.9 -- 2004.7 中国科学院计算技术研究所计算机应用方向，博士研究生

1997.9 -- 1999.7 哈尔滨工业大学计算机系计算机应用专业，硕士研究生

1993.9 -- 1997.7 哈尔滨工业大学计算机系计算机应用专业，大学本科

【攻读博士学位期间发表的主要论文】

图书章节：

- [1] Wen Gao, Shiguang Shan, Face Verification for Access Control, “Biometrics Solutions for Authentication in an E-World”, Edited by David Zhang, Kluwer Academic Publishers, Chapter 13, pp339-376, 2002

国际期刊：

- [2] Shiguang Shan, Wen Gao, Debin Zhao, Face Identification Based On Face-Specific Subspace, International Journal of Image and System Technology, Special issue on face processing, analysis and synthesis, 13(1), pp23-32, (2003), Publisher: John Wiley & Sons Inc., 605 Third Avenue, New York, NY 10158, USA
- [3] Laiyun Qing, Shiguang Shan, Wen Gao, Face Recognition under generic illumination based on Harmonic relighting, To appear in International Journal of Pattern Recognition and Artificial Intelligence (Accept)
- [4] Wen Gao, Yiqiang Chen, Rui Wang, Shiguang Shan, and Dalong Jiang, Learning and Synthesizing MPEG-4 Compatible 3-D Face Animation From Video Sequence, IEEE Transactions on circuits and systems for video technology, vol. 13, no. 11, pp1119-1128, 2003

国内期刊：

- [5] 柴秀娟，山世光，高文，陈熙霖，基于样例学习的面部特征自动标定算法，软件学报（已录用）
- [6] 山世光，高文，陈熙霖 基于纹理分布和变形模板的面部特征提取，软件学报, Vol.12, No.4, pp570-577, 2001.4
- [7] 陈熙霖，山世光，高文 多姿态人脸识别，中国图像图形学报, Vol.4(A), No. 10, pp818-824, 1999.10
- [8] 焦峰，山世光，崔国勤，高文，李锦涛，基于局部特征分析的人脸识别方法，计算机辅助设计与图形学学报，15（1），pp53-58，2003.1

国际会议：

- [9] Shiguang Shan, Wen Gao, Yizheng Chan, Bo Cao, Peng Yang, Review the Strength of Gabor features for face recognition from the angle of its robustness to mis-alignment, ICPR2004 (Accept)
- [10] Shiguang Shan, Yizheng Chang, Wen Gao, Bo Cao, Curse Of Mis-Alignment In Face Recognition: Problem And A Novel Mis-Alignment Learning Solution, Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp314-320, Korea, May, 2004
- [11] Shiguang Shan, Wen Gao, Bo Cao, Debin Zhao. Illumination Normalization for Robust Face Recognition against Varying Lighting Conditions, IEEE International Workshop on Analysis and Modeling of Faces and Gestures, pp157-164, Nice, France, 2003.10
- [12] Shiguang Shan, Wen Gao, Wei Wang, Debin Zhao, Baocai Yin. Enhanced Active Shape Models with Global Texture Constraints for Face Image Analysis, Fourteenth International Symposium On Methodologies For Intelligent Systems, N. Zhong et al. (Eds.): ISMIS 2003, LNAI2871, pp593-597, Springer, Maebashi City, Japan, Oct.2003
- [13] Shiguang Shan, Wen Gao, Yan Lu, Bo Cao, Xilin Chen Debin Zhao, "Unified Framework For Classifying Facial Images Based On Facial Attribute-Specific Subspaces And Minimum Reconstruction Error", ACCV'2002, vol.2, pp858-863, 23-25 Jan. 2002, Melbourne, Australia
- [14] Shiguang Shan, Wen Gao, Debin Zhao, Face Identification From A Single Example Image Based On Face-Specific Subspace (FSS), IEEE ICASSP2002, Vol.2, pp2125-2128 Florida, USA, 2002.5
- [15] Shiguang Shan, Bo Cao, Wen Gao, Debin Zhao, Extended Fisherface For Face Recognition From A Single Example Image Per Person, IEEE International Symposium on Circuits and Systems, 2002 (IEEE ISCAS2002) , Vol.2, pp81-84, 2002
- [16] Shiguang Shan, Wen Gao, Xilin Chen, Bo Cao, Wenbin Zeng, A Face-Unlock Screen Saver by Using Face Verification Based on Identity-Specific Subspaces, PCM'01, Beijing, China pp1096-1101, 2001
- [17] Shiguang Shan, Wen Gao, JieYan etc. Individual 3d face synthesis based on orthogonal photos and speech-driven facial animation. Proceedings of the International Conference on Image Processing (ICIP'2000), Vol. III, pp238-242, Vancouver, BC, Canada, 2000.9
- [18] Shiguang Shan, Wen Gao, Xilin Chen, Jiyong Ma. Novel Face Recognition Based on Individual Eigen-subspaces, Proceedings of International Conference on Signal Processing, ICSP2000, Vol. III, pp1522-1525, 2000
- [19] Wen Gao, Shiguang Shan, Xiujuan Chai, Xiaowei Fu, Virtual Face Image Generation For Illumination And Pose Insensitive Face Recognition, International Conference on Acoustic, Speech and Signal Processing 2003, Vol.IV, pp776-779, HongKong, 2003.4
- [20] Wenchao Zhang, Shiguang Shan, Wen Gao, Bo Cao, Information Fusion in Face Identification, ICPR2004 (Accept)
- [21] Peng Yang, Shiguang Shan, Wen Gao, Stan Li, Dong Zhang, Face Recognition Using Ada-Boosted Gabor Features, Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp356-361, Korea, May, 2004

- [22] Laiyun Qing, Shiguang Shan, Wen Gao, Eigen-Harmonics Faces: Face Recognition under Generic Lighting, Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp296-301, Korea, May, 2004
- [23] Laiyun Qing, Shiguang Shan, Wen Gao, Face Recognition with Harmonic De-Lighting, Proc. of ACCV2004, vol.2, p824 ~ 829, Jeju, Korea, 2004.1
- [24] Wenchao Zhang, Shiguang Shan, Wen Gao, Jianyu Wang, Debin Zhao, Incremental Face-Specific Subspace for Online-Learning Face Recognition, Proc. of ACCV2004, vol.2, p1080-1084, Jeju, Korea, 2004.1
- [25] Xiujuan Chai, Shiguang Shan, Wen Gao, Pose Normalization for Robust Face Recognition based on Statistical Warping, Fourth IEEE Pacific-Rim Conference On Multimedia, Vol.2, pp1413-1417, December 2003, Singapore.
- [26] Xiujuan Chai, Shiguang Shan, Wen Gao, Bo Cao, Novel Example-Based Shape Learning For Fast Face Alignment, International Conference on Acoustic, Speech and Signal Processing, IEEE ICASSP2003, Vol.III, pp229-232, HongKong, 2003.4
- [27] Bo Cao, Shiguang Shan, Wen Gao, Debin Zhao, "Localizing the Iris Center by Region Growing Search," Proceedings of IEEE International Conference on Multimedia and Expo, Vol. 2, pp129-132, 2002.
- [28] Wei Wang, Shiguang Shan, Wen Gao, Bo Cao. An Improved Active Shape Model For Face Alignment, The fourth International Conference on Multi-modal Interface, IEEE ICMI' 2002, Pittsburgh, USA, pp523-528, Oct. 2002
- [29] Sanqiang Zhao, Wen Gao, Shiguang Shan, Baocai Yin, Enhance the Alignment Accuracy of Active Shape Models Using Elastic Graph Matching, ICBA2004, HK, 2004, (Accept)
- [30] Baochang Zhang, Wen Gao, Shiguang Shan, Peng Yang, Discriminant Gaborfaces and Support Vector Machines Classifier for Face Recognition, Proc. of ACCV2004, vol.2, pp37-41, Jeju, Korea, 2004.1
- [31] Laiyun Qing, Wen Gao, Shiguang Shan, Surface Reflectance Classifying under Natural Illumination, Fourth IEEE Pacific-Rim Conference On Multimedia, vol.3, pp753-757, December 2003, Singapore.
- [32] Baochang Zhang, Wen Gao, Shiguang Shan, Wei Wang, Constraint Shape Model Using Edge Constraint And Gabor Wavelet Based Search, Proceeding of Audio- and Video-Based Biometric Person Authentication, pp52-61, LNCS2688, Josef Kittler (Editors.) Springer, 2003
- [33] Jianyu Wang, Wen Gao, Shiguang Shan, XiaoPeng Hu, Facial Feature Tracking Combining Model-Based And Model-Free Method , Interpretation, International Conference on Acoustic, Speech and Signal Processing, Vol.III, pp205-208, HongKong, 2003.4
- [34] Chunli Wang, Wen Gao, and Shiguang Shan. An Approach Based on Phonemes to Large Vocabulary Chinese Sign Language Recognition. In Proc. Int. Conf. on Automatic Face and Gesture Recognition, pages 411–416, Washington, DC, USA, May 21-22, 2002.
- [35] Feng Jiao, Wen Gao, Shiguang Shan, etc. A Face Recognition Method Based on Local Feature Analysis, ACCV'02, pp.188-192, Melbourne, Australia, Jan. 2002
- [36] Debin Zhao, Wen Gao, Shiguang Shan, Y. K. Chan, LLEC: an Image Coder with Low-complexity and Low-memory Requirement, PCM'01, Beijing, China , 2001

- [37]Wen Gao, Jiyong Ma, Shiguang Shan, Xilin Chen, Wei Zeng, Hongming Zhang, Jie Yan, Jiangqin Wu. HandTalker: A Multimodal Dialog System Using Sign Language and 3-D Virtual Human, ICMIT'2000, Lecture Notes in Computer Science 1948, pp564-571, Beijing, October 14-16, 2000

国内会议：

- [38]Shiguang Shan, Wen Gao, Laiyun Qing, Generating Canonical Form Face Images for Illumination Insensitive Face Recognition, Proc. of Sinobiometrics'03, pp29-35, 2003.
- [39]Laiyun Qing, Shiguang Shan, Wen Gao, Face De-lighting under Natural Illumination, Proc. of Sinobiometrics'03, pp43-48, 2003.
- [40]Wenchao Zhang, Shiguang Shan, Wen Gao, Jianyu Wang and Debin Zhao, Online Updating Face Model Based on Incremental Face-Specific Subspace, Proc. of Sinobiometrics'03, pp89-93, 2003.
- [41]卿来云, 山世光, 高文, 基于商图像的人脸图像光照补偿算法, , Proc. of Sinobiometrics'03, pp106-110, 2003.
- [42]柴秀娟, 山世光, 高文, 用于人脸对准的快速形状学习方法, Proc. of Sinobiometrics'03, pp93-97, 2003.

【攻读博士学位期间参加的主要科研项目】

- [1] 基于生物特征的身份识别研究, 国家自然科学基金委重点项目, 2004-2007, 主要参加人, 编号: 60332010
- [2] 生物特征识别核心技术与关键问题研究, 863计划, 2001-2003, 负责人脸识别部分, 编号: 2001AA114190
- [3] 多媒体掌上电脑与通信掌上电脑, 国家863计划, 2001-2003, 负责移动身份验证子课题, 编号: 2002AA118010
- [4] 面像识别研究, 银晨网讯科技有限公司资助项目, 2000-2006, 负责人脸识别研究
- [5] 多功能感知机研究, 国家863计划, 编号: 2001AA142140, 2001-2002, 负责人脸识别子课题
- [6] 多功能感知机, 国家自然科学基金项目 (No.69789301), 1998-2000, 负责人脸识别、虚拟人面部合成, 结题A
- [7] 多功能感知技术与平台的研究, 国家863项目资助, 合同号: 863-306-ZT03-01-2, 1998-2000, 负责人脸识别子课题

【攻读博士学位期间科研成果与获奖情况】

- [1] 2003年, 上海市科技进步一等奖, 项目: 人脸识别系统及其应用, 排名第四;
- [2] 2003年, 中科院院长奖优秀奖
- [3] 2002年, 中国科学院科研鉴定, 项目: 人脸检测与识别核心技术, 排名第四;
- [4] 2002年, 中科院计算所所长奖学金二等奖;
- [5] 2001年, 中科院计算所所长奖学金二等奖;
- [6] 1999年, 中国科学院科研鉴定, 项目: 手语自动翻译与虚拟人合成系统, 排名第六;