

申请上海交通大学硕士学位论文

实用人脸识别系统初探

系 别： 自动化系

专 业： 控制理论与控制工程

作 者： 戴景文

导 师： 苏剑波 教授

上海交通大学

2009 年 1 月

A Thesis Submitted to Shanghai Jiaotong University
for the Master Degree

FUNDAMENTAL RESEARCH ON PRACTICAL FACE RECOGNITION SYSTEM

Department: Department of Automation

Major: Control Theory and Control Engineering

Author: Jingwen Dai

Advisor: Prof. Jianbo Su

Shanghai Jiaotong University
January, 2009

上海交通大学
学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

上海交通大学 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本授权书。

本学位论文属于

不保密 ☐。

（请在以上方框内打“√”）

学位论文作者签名：

指导教师签名：

日期： 年 月 日

日期： 年 月 日

摘 要

近些年来,构建自动人脸识别系统是计算机视觉和模式识别领域的研究热点之一,陆续出现了一些人脸识别商业系统应用于公众和个人安防等方面。广义的讲,自动人脸识别系统一般需要经过人脸检测,人脸特征点定位,特征选择与提取,比对与识别四个步骤以完成识别任务。人脸检测判定图像或视频序列中是否存在人脸,若存在,确定人脸的大小与位置;人脸特征点定位获得人脸显著特征点(如眼睛,鼻子,嘴巴)的位置;特征选择与提取得到人脸最具有区分性的特征;人脸比对与识别通过与已有数据库的比较,获得待识别人的身份。

本文以人脸识别的实用化,实时化和系统化为目标,对实用人脸识别系统中若干问题进行了初步探讨。论文的主要工作如下:

(1) 提出一种新的基于投影峰分析的眼睛快速定位方法

基于投影峰分析的眼睛快速定位方法,充分利用了人脸五官特征的统计规律,通过简单的运算实现眼睛中心坐标的准确快速的定位,针对额发,鬓角,眉毛,眼镜等干扰,通过投影峰分析,有效地去除了上述干扰,使得定位算法更具鲁棒性。实验结果表明,当人脸姿态,光照,表情和配饰变化时,该方法依然可以完成准确且快速的眼睛定位。由于较低的计算复杂度,该方法完全可以满足实时人脸识别系统的需要。

(2) 改进基于 Gabor 特征的人脸识别方法

针对 Gabor 特征提取中,二维卷积运算复杂度较高,耗时较长的缺点,提出了基于特征选择的 SFFS-Gabor 人脸识别方法。该方法通过 SFFS 方法在训练集上的学习,经过两个步骤,选择出最具分辨能力 Gabor 核的位置、尺度和方向,以此描述人脸特征。实验结果表明,该方法保证较高人脸识别率的同时,降低了算法复杂度,明显缩短了人脸识别的时间消耗,使得基于 Gabor 特征描述的人脸识别方法,更加适用于实用人脸识别系统。

(3) 改进基于 LBP 特征的人脸识别方法

针对一般 LBP 人脸描述维数较高,无法应用于实时识别的不足,提出压缩直方图的思想,有效地降低特征维数,实验结果,表明压缩直方图的方法可缩短识别时间,并保证较高的识别率。同时提出了一种新的基于全变分模型的 LBP 方法,利用全变分模型对人脸边缘、轮廓和角点等显著特征的增强,并结合了 LBP 算子对图像边缘、角点等局部变化特征敏感的特性,提取对于外界因素变化特别是光照变化,较为鲁棒的人脸特征用于人脸识别,明显提高了光照变化情况下的人脸识别率。

(4) 构建开放式人脸识别体系框架

为了给实用人脸识别系统的设计提供参考原型,我们提出一种新的开放式人脸识别体系框架,根据开放式人脸识别体系框架的通用性、开放性和分布式的三大设计目标,具体描述了体系框架的设计思路和系统整体框架;并以面向对象技术、数据库技术和

OpenCV 为技术平台，具体实现开放式人脸识别体系框架。

(5) 实现基于人脸识别的考勤系统

以办公室环境的日常考勤为应用目的，以开放式人脸识别体系框架为基础，采用 SFFS-Gabor 人脸识别方法实现基于人脸识别的考勤系统，以此为实施例，验证开放式人脸识别体系的开放性与通用性，同时也验证了 SFFS-Gabor 人脸识别方法在实际应用中的有效性。

关键词： 人脸识别，眼睛快速定位，投影峰分析，Gabor 小波，特征选择，SFFS，LBP 算子，压缩直方图，全变分模型，开放式人脸识别框架

ABSTRACT

Recent years, building automatic face recognition system has become a hot topic of computer vision and pattern recognition. And some commercial face recognition systems have been developed and applied in public and individual security. In general, an automatic face recognition system is accomplished in four steps, i.e. face detection, facial landmarks localization, facial feature selection and extraction, face recognition. Face detection determines whether or not there are any faces in the image or video sequence and, if present, acquires the location and extent of each face. Facial landmarks localization obtains the location of salient feature points of face, i.e. eyes, nose, mouth etc. Facial feature selection and extraction acquires the most representative facial features. And face recognition identifies or verifies one or more persons in the scene using a stored database of faces.

In this thesis, some key issues are primarily studied, aiming at building real-time practical face recognition system. And the main work of this thesis can be described as follows:

(1) Proposed a novel projection peak analysis method for rapid eye localization

The projection peak analysis method achieves rapid and accurate eye localization by making use of the static rules of human face on the basis of uncomplicated computation. In order to eliminate the interferences (i.e. hair, eyebrow, glasses) around eye region, we improve the general projection method by projection peak analysis. Experimental results show that our method is effective, accurate and rapid in eye localization, especially when the face poses, illuminations, expressions, and accessories varied. Owing to the lower computation cost, our method can satisfy the requirement of real-time face recognition system well.

(2) Improved the face recognition method base on Gabor feature

For the high computational cost and long time consuming shortcoming of two-dimension convolution in the procedure of Gabor feature extraction, we proposed SFFS-Gabor face recognition method based on feature selection. In this method, through the learning on training set by means of sequential floating forward search(SFFS), the most representative Gabor features with different position, scale and orientation could be selected after two-stage procedure. Then the face description under the selected features could be employed in face recognition. Experiment result shows that this method could ensures high face recognition accuracy, and at the same time, it could reduce the complexity of algorithm and shorten recognition time evidently. After feature selection, the

face recognition method on Gabor feature is more suitable for real-time practical face recognition system.

(3) Improved the face recognition method on LBP feature

The dimension of general LBP feature is very high, so it is not suitable for real-time application. To solve this problem, we reduce the feature dimension by taking advantage of compressed histogram. Experimental result shows that compress-histogram method could achieve high recognition rate with short time consuming. In addition, we also proposed a novel LBP method based on total variation model(TVM-LBP). Total variation model could enhance the salient facial feature such as edges, profiles, corners and points, while LBP operator is sensitive to some local features, i.e. edges and corners of image. The TVM-LBP method combines the advantage of TVM and LBP mentioned above, which could improve the face recognition performance, especially when the environmental illumination has been changed.

(4) Initially proposed and implemented an open face recognition framework

In order to provide a reference prototype for the design of practical face recognition system, we proposed an open face recognition framework(OFRF). Aiming at the universal, open and distributive properties, we studied the design pattern and holistic structure of OFRF, and implemented the framework, depending on the support of three technical platform, i.e. object oriented technology, database technology and open computer vision library(OpenCV).

(5) Implemented a face recognition system for checking attendance

For the purpose of daily checking attendance in office environment, we implemented a real-time face recognition system using SFFS-Gabor method based on the open face recognition framework(OFRF). This checking attendance system proved the open and universal properties of open face recognition framework and the effectiveness of SFFS-Gabor face recognition method in real-time practical application.

KEY WORDS: Face Recognition, Rapid Eye Localization, Projection Peak Analysis, Gabor Wavelet, Feature Selection, Sequential Floating Forward Search(SFFS), Local Binary Pattern(LBP), Compressed Histogram, Total Variation Model(TVM), Open Face Recognition Framework(OFRF)

目 录

摘 要	i
ABSTRACT	iii
目录	v
第一章 绪论	1
1.1 引言	1
1.2 人脸识别的研究意义	1
1.2.1 研究意义	1
1.2.2 应用的优势与不足	2
1.3 人脸识别研究概述	3
1.3.1 人脸识别分类	3
1.3.2 特征描述	4
1.3.3 特征选择	5
1.3.4 特征分类	5
1.3.5 特征融合	6
1.4 自动人脸识别系统的一般框架	6
1.5 国内外主要人脸识别商业系统	7
1.6 本文的主要工作	7
1.7 本文的章节结构	9
第二章 人脸检测与特征点的定位	10
2.1 人脸检测	10
2.2 人脸特征点定位	11
2.3 基于投影峰分析的快速眼睛定位方法	12
2.3.1 算法描述	13
2.3.2 实验与分析	16
2.3.3 方法小结	18
2.4 人脸标准化	20
2.4.1 旋转	20
2.4.2 剪切与缩放	21
2.4.3 预处理	21
2.4.4 人脸图像Mask	23

2.5 本章小结	24
第三章 基于 Gabor 的人脸识别	25
3.1 Gabor 小波变换	25
3.2 人脸的 Gabor 表示	28
3.3 PCA+FLD 人脸识别方法	31
3.3.1 图像下采样和归一化	31
3.3.2 PCA(Principle Component Analysis)	31
3.3.3 Fisher-LDA(Fisher Linear Discriminant Analysis)	32
3.3.4 实验结果与分析	34
3.4 基于特征选择的 Gabor 人脸识别方法	35
3.4.1 复杂度分析	36
3.4.2 特征选择	37
3.4.3 实验与结果分析	38
3.4.4 方法小结	41
3.5 本章小结	41
第四章 基于 LBP 的人脸识别	42
4.1 LBP 算子	42
4.2 LBP 在人脸识别中的应用	43
4.2.1 LBP 人脸描述	43
4.2.2 双线性插值的 LBP 人脸描述	45
4.2.3 实验结果与分析	47
4.3 基于全变分模型的 LBP 人脸识别	48
4.3.1 全变分模型	49
4.3.2 TVM+LBP 人脸描述	51
4.3.3 实验	51
4.3.4 结论	52
4.4 本章小结	52
第五章 开放式人脸识别体系框架	53
5.1 背景	53
5.2 设计目标	54
5.2.1 通用性	54
5.2.2 开放性	55
5.2.3 分布式	55
5.3 设计思路	56
5.3.1 系统整体框架	56

5.3.2	模块化	56
5.3.3	层次化	58
5.4	技术平台	58
5.4.1	面向对象技术	58
5.4.2	数据库技术	60
5.4.3	OpenCV	61
5.5	开放式人脸识别体系框架的实现	61
5.5.1	数据库实现	61
5.5.2	类实现	63
5.5.3	GUI 实现	64
5.6	本章小结	65
第六章	基于人脸识别的考勤系统	68
6.1	系统概述	68
6.1.1	编写目的	68
6.1.2	硬件及软件运行环境	68
6.1.3	系统功能与特点	69
6.2	系统实现	70
6.2.1	系统结构	70
6.2.2	附加功能	71
6.2.3	系统界面	72
6.3	本章小结	73
第七章	总结与展望	74
7.1	总结	74
7.2	展望	75
附录 A	流程图	76
附录 B	数据库表图	78
参考文献		81
致谢		90
攻读学位期间发表的论文与成果		91

第一章 绪论

本章提出了本文的研究目标。首先阐述了人脸识别的研究意义；然后介绍了人脸识别的研究现状；接着介绍了自动人脸识别系统的一般框架及国内外主要人脸识别商业系统；最后简要地介绍了本文的主要工作和章节结构。

1.1 引言

人类通过视觉识别周围环境，感知外界信息。在客观世界中，人们 75% 的信息量都来自视觉，因此让计算机或机器人具有视觉，是人工智能的重要环节，也是人类多年以来的梦想。在信号处理理论与高性能计算机出现后，这种希望正在逐步实现。用计算机实现对视觉信息的处理形成了一门新兴的学科——计算机视觉。视频/图像中人脸处理问题是近年来计算机视觉领域内一个相当活跃的研究课题。由于各方面对人脸识别系统的迫切需要，人脸自动识别是一个热门的研究方向 [1]。

1.2 人脸识别的研究意义

人脸识别的研究始于二十世纪六十年代中后期 [2-7]，近 40 年来人脸识别的研究得到了长足的发展，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。

1.2.1 研究意义

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 [8]：

(1) 人脸识别研究可以极大地促进多门相关学科的发展

人脸识别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。

若把人脸识别问题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。其模式种类数目庞大，不同类别模式的差别非常细微，而且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因素的影响使得这一问题更加突出。

若把人脸识别问题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 [9-11]。从计算机图形学的角度，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 [12-14]。

人脸识别同时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，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、安全性和可用性上。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。

1.2.2 应用的优势与不足

与指纹、虹膜、视网膜、声音等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，这主要体现在以下四点：

(1) 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；

(2) 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；

(3) 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）。

(4) 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，而且易于安装，极大的扩展了其应用空间。

然而，人脸作为生物特征识别技术具有上述优势的同时，也存在如下三点固有缺陷：

(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；

正是由于这些缺点，实际应用环境下的实时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下才可以基本满足一般应用的要求 [15, 16]。

1.3 人脸识别研究概述

1.3.1 人脸识别分类

根据系统初始输入模式的不同，人脸识别的研究可以分为不同的种类。根据是否包含运动信息可以将人脸识别分为静态图像识别和序列图像识别两大类。尽管后者可以得到更多关于人脸的信息，但后者存在人脸精确分割问题，因此目前大多数的人脸识别的

研究都是基于静态图像进行的。从输入人脸模式属性不同的角度来看,人脸识别的研究大致可以分为:

(1) 侧面图像或者剪影识别

在人脸识别研究的初期,基于侧面图像或剪影图像的人脸识别得到了很好的发展,现在该方法通常被作为正面和多姿态识别的辅助或者用于融合算法。基于侧面的人脸识别多数采用基于几何特征的识别方法 [17-19]。

(2) 正面或准正面人脸识别

80 年代和 90 年代初期多数人脸识别研究的重点。主要方法有: 基于特征的方法 [4, 20]、基于局部或全局模板匹配 [21, 22]、特征脸 (Eigenface) [22, 23]、主成分分析 (PCA) [24, 25]、奇异值分解 (SVD) [26-28]、人工神经网络 (ANN) [24, 29, 30] 等。

(3) 多姿态人脸识别

90 中后期以来的研究热点,正在得到进一步的研究。多视图模板匹配 [22]、基于 3D 人脸模型的姿态补偿 [31]、弹性图匹配 [32-34] 是解决多姿态问题的主流方法。

(4) 3D 人脸识别

随着激光扫描器 (Laser Scanner) 等特殊设备的发展和成熟,可以得到人脸的三维模型,深度信息的获得可以使识别性能得到增强 [35-37], 因此三维人脸识别渐渐成为研究的热点问题。

目前人脸识别的研究重点是多姿态识别,不同识别算法的融合、实用系统的设计以及多通道信息融合,但就目前的研究成果看来,即使基于正面人脸的识别尚未能够完全满足实用的要求。

人脸识别的综述文献 [38-44] 对人脸识别整个领域做了充分而翔实的概述,下面本文试图从特征描述、特征选择、特征分类和特征融合四个方面来总结人脸识别研究的相关内容。

1.3.2 特征描述

对于人类而言,人脸是皮肤、肌肉和器官的组合。而对于计算机而言,人脸图像是像素点灰度值的矩阵,计算机中用什么样的知识来表达人脸对象,即所说的特征描述,成为人脸识别领域一个重要研究方向。一组好的特征不仅要充分反映不同人脸图像之间的微弱差别,同时又必须保证同一个对象的不同人脸图像特征具有一定的聚类性。人脸识别文献中采用的特征描述的方法主要有如下几种:

(1) 面部几何结构特征 人脸五官的距离、角度、面积等参数 [4, 7, 17, 21]。

(2) 灰度模板 所有像素亮度值串接形成的向量或矩阵 [21]。

(3) 主成分特征 对图像进行 PCA 降维 (去相关) 后的特征 [25, 45]。

(4) 奇异值等代数特征 对图像进行 SVD 分解得到的奇异值 [26-28, 46]。

(5) 可变形模板 抛物线、双曲线等数学模型的参数。

(6) 2D 形状模型 稀疏特征点统计分布模型 (PDM) [47]。

- (7) 光流形状模型 与参考图像的稠密的、像素级的对应关系 [31]。
- (8) 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 [47]。
- (9) 表观模型 形状与纹理通过 PCA 进一步建模 [47]。
- (10) 3D 形变模型 3D 形状 PCA、纹理 PCA 二者的结合 [10,11]。
- (11) 局部特征 能够保留全局拓扑关系的局部结构信息 [48]。
- (12) 弹性属性图 关键特征点的局部结构信息及其几何关系 [32-34]。
- (13) 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 [49-51]。
- (14) 独立元特征 各特征之间具有统计独立性 [51]。
- (15) 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 [52]。
- (16) 商图像 表面反射率之比 [53]。

1.3.3 特征选择

特征选择不但可以降低特征的维数,减少特征的冗余度,同时还可以选择出最具有区分性的特征,因此特征选择方法可以加快识别速度和提高识别精度。传统的特征选择方法有互信息 (mutual information), 信息增益 (information gain), χ^2 统计和相关系数 (correlation criteria) 等方法 [54], 但是这些方法不能应用于像人脸识别这样的高维小样本问题。最近有研究者提出了基于目标函数的特征选择新的方向 [55-60]。有研究者对特征选择方法进行了详细的总结归纳 [55]。基于支持向量机的特征选择方法 [61-64] 是属于这个新的特征选择方向中的一种类型。

1.3.4 特征分类

通常情况下,识别问题最终都归结为输入模式与已知模式在特征空间内精确的或者模糊的匹配问题。匹配有两个核心问题:相似度量度和分类器。最常实用的相似度量包括:欧式距离、街区距离、角度相似性函数、相关量计算、能量函数等。一般认为不同类别的人脸在特征空间内是非线性可分的,因此分类器的设计对于整个人脸识别算法至关重要。文献中 [38-44] 的主要分类算法有:

- (1) 最近邻 /k-近邻分类 这是传统的但经常被采用的分类算法,具有简单有效的特点。多数采用几何构型、主分量、变换系数、代数特征人脸描述方法的算法都可以使用这类分类方法。
- (2) 相关量计算 这是一种与基于模板的人脸描述相对应的方法。用模板和输入模式之间的协相关量作为相似性度量实现匹配。
- (3) Bayes分类器 这是统计模式识别中经常采用的方法,用 Bayes 定理计算输入模式属于各个类别的后验概率实现识别。被广泛应用于需要进行分类融合的系统中。
- (4) 弹性图匹配 该分类匹配方法使用弹性图作为人脸特征的描述方式,定义不同弹性图之间相似度的能量函数,该函数反映它们之间形变的程度,通常有拓扑结构能量

项和空间相似能量项。

(5) 人工神经网络 人工神经网络已经广泛的应用于模式识别领域。自组织映射 (SOM)、自联想神经网络 (SAN)、动态连接结构 (DLA) 等被应用于识别和分类任务, 尤其在表情分析、性别辨别、种族分类、年龄判断等应用中得到了广泛的应用。

(6) 模糊逻辑 该方法主要用作人脸识别算法级的数据融合和多通道信息融合策略。

1.3.5 特征融合

近年来, 信息融合技术在语音、图像等数据处理领域得到了重视。在人脸识别领域, 特征融合的研究主要体现在如下两个方面针对相同数据基提取的相同或不同特征, 可以在不同分类算法之间实现算法层的融合 [65, 66]; 针对不同输入数据基 (如侧脸人脸和正脸人脸 [67]) 的相同或者不同算法之间可以实现结果之间的融合。在个人身份认证领域, 由于来自多种通道的信息 (如人脸, 指纹, 声音或射频卡等) 可以相互支持和补偿, 尤其是在某一个通道受到高噪声干扰的情况下, 多通道信息融合技术可以得到比单通道更好的性能和更高的可靠性。

1.4 自动人脸识别系统的一般框架

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, AFR) 系统。假设我们把照相机、摄像头、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”, 那么 AFR 赋予计算机根据其所“看到”的人脸图片来判断人物身份的能力。广义的讲, 自动人脸识别系统具有如图 1.1 所示的一般框架并完成相应功能的任务。

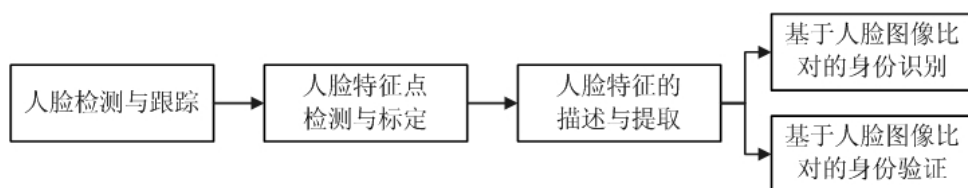


图 1.1 自动人脸识别系统的一般框架

Fig 1.1 General framework of AFR system

(1) 人脸检测与跟踪

人脸检测的任务是判断静态图像或视频流的每帧图像中是否存在人脸。若存在人脸, 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) 人脸特征点检测与标定

通过人脸特征点的检测与标定可以确定人脸图像中显著特征点的位置 (如眼睛、眉毛、鼻子、嘴巴等器官), 同时还可以得到这些器官及其面部轮廓的形状信息的描述。

(3) 人脸特征的描述与提取

根据人脸特征点检测与标定的结果,通过某些运算得到人脸特征的描述(这些特征包括:全局特征和局部特征,显式特征和统计特征等)。

(4) 基于人脸图像比对的身份识别

即人脸识别 (Face Identification) 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

即人脸确认 (Face Verification) 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。

1.5 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 1.1 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 [8]。

1.6 本文的主要工作

本文的以人脸识别的实用化,实时化和系统化为最终目标,研究了人脸面部特征的定位与标定,人脸特征描述与特征提取及人脸特征分类,提出了一种通用的开放式人脸识别框架体系,并以基于人脸识别的考勤系统为具体实施例,验证了该框架体系的有效性与可行性。

本文的主要工作有以下几个部分:

(1) 提出一种新的基于投影峰分析的眼睛快速定位方法

基于投影峰分析的眼睛快速定位方法,充分利用了人脸五官特征的统计规律,通过简单的运算实现眼睛中心坐标的准确快速的定位,针对额发,鬓角,眉毛,眼镜等干扰,提出基于投影峰的定位方法,有效地去除了上述干扰,使得定位算法更具鲁棒性,此外当人脸姿态发生俯仰($\pm 15^\circ$),旋转($\pm 15^\circ$)等变化时,同样可以实现精确定位。

(2) 改进基于 Gabor 特征的人脸识别方法

针对 Gabor 特征提取中,二维卷积运算复杂度较高,耗时较长的缺点,提出了基于特征选择的 SFFS-Gabor 人脸识别方法。该方法通过 SFFS 方法通过训练集的学习,经过两个步骤,选择出最具分辨能力 Gabor 核的位置、尺度和方向,以此描述人脸特征,

表 1.1 主要商业人脸识别系统简表

Table 1.1 Brief introductions of the main commercial face recognition systems

系统名称	开发公司	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。
FaceIt	Identix	基于 Rockfeller 大学开发的 LFA 算法。
PersonSpotter	Neven Vision	基本算法是基于 Gabor 特征的人脸识别方法。
ZN face II	ZN Bochum	基本算法来自美国 USC 和德国 Bochum 大学开发的基于 Gabor 特征的弹性图匹配人脸识别方法。
Face TOOLS	Viisage	基于 MIT 的特征脸技术, 后加入 ICA 方法。
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 方法。
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。

保证较高人脸识别率的同时, 明显缩短了人脸识别的时间消耗, 使得基于 Gabor 特征描述的人脸识别方法, 更加适用于实用人脸识别系统。

(3) 改进基于 LBP 特征的人脸识别方法

针对一般 LBP 人脸描述维数较高, 不适合实时应用的缺点, 提出了压缩直方图的思想, 有效地降低了特征维数, 并保证了较高的人脸识别率。同时提出了一种新的基于全变分模型的 LBP 方法, 利用全变分模型对人脸边缘、轮廓和角点等显著特征的增强, 并结合了 LBP 算子对图像边缘、角点等局部变化特征敏感的特性, 提取对于外界因素变化特别是光照变化, 较为鲁棒的人脸特征用于人脸识别, 明显提高了光照变化情况下的人脸识别率。

(4) 构建开放式人脸识别体系框架

提出一种新的开放式人脸识别体系框架, 根据开放式人脸识别体系框架的通用性、开放性和分布式的三大设计目标, 具体描述了体系框架的设计思路和系统整体框架; 并以面向对象技术、数据库技术和 OpenCV 为技术平台, 具体实现开放式人脸识别体系框架。

(5) 实现基于人脸识别的考勤系统

以办公室环境的日常考勤为应用背景, 以开放式人脸识别体系框架为基础, 采用 SFFS-Gabor 人脸识别方法实现基于人脸识别的考勤系统, 以此为实施例, 验证开放式人脸识别体系的开放性与通用性, 同时也验证了 SFFS-Gabor 人脸识别方法的有效性。

1.7 本文的章节结构

第二章主要介绍自动人脸识别系统中必不可少的两个环节：人脸检测，人脸特征点定位及人脸标准化。首先简要介绍人脸检测的一般原理与常用方法；接着介绍人脸特征点定位的一般方法；然后提出一种基于投影峰分析的眼睛快速定位方法；最后在完成眼睛定位的基础上，实现人脸图像的标准化，从而完成自动人脸识别系统的前期准备工作。

第三章主要以人脸的 Gabor 描述为主线展开。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，是一种良好的人脸特征提取方法。我们在研究人脸识别的过程中发现了 Gabor 小波特征能够对面脸图像进行有效的表达。Gabor 小波特征可以比原始像素特征具有更强的抽象概括能力，所以更适合在特征选择过程中使用。本章首先介绍 Gabor 小波，接着描述基于 FLD 的人脸识别方法，然后分析该方法存在的不足，最后通过特征选择，提出一种新的基于 Gabor 小波的人脸识别方法。

第四章介绍了另一种比较有效的人脸特征描述子：局部二值模式 (Local Binary Pattern, LBP)。首先介绍 LBP 算子，接着描述基于 LBP 的人脸识别方法，然后分析该方法存在的不足，提出了直方图压缩的思想，有效降低 LBP 特征维数，最后提出一种新的基于全变分模型的 LBP 人脸识别方法，以克服光照变化对 LBP 描述的影响。

第五章提出一种开放式人脸识别体系框架。首先先介绍人脸识别系统的背景；接着给出开放式人脸识别体系框架的三大设计目标，即通用性、开放性和分布式；然后给出该体系框架的设计思路和系统整体框架；最后结合三大技术平台(面向对象技术、数据库技术和 OpenCV) 具体实现开放式人脸识别体系框架。

第六章以第五章提出的开放式人脸识别体系框架为基础，以办公室环境下的日常考勤为任务，实现了基于人脸识别的考勤系统。

第七章中总结了全文的工作，对需要进一步研究的问题进行了展望。

第二章 人脸检测与特征点的定位

本章主要介绍自动人脸识别系统中必不可少的两个环节：人脸检测，人脸特征点定位及人脸标准化。首先简要介绍人脸检测的一般原理与常用方法；接着介绍人脸特征点定位的一般方法；然后提出一种基于投影峰分析的眼睛快速定位方法；最后在完成眼睛定位的基础上，实现人脸图像的标准化，从而完成自动人脸识别系统的前期准备工作。

2.1 人脸检测

人脸检测 (Face Detection) 不是本文研究的重点，但人脸检测是人脸特征点定位的基础，故在此节中简要介绍一下人脸检测的一般原理与方法。人脸检测问题一般可以描述为：给定静止或动态图像，判断其中是否有人脸；若有，将所有人脸从背景中分割出来，并确定每个人脸在图像中的位置和大小 [68,69]。人脸检测的基本思想是利用知识或统计的方法对人脸建模，比较所有可能的待检测区域与人脸模型的匹配度，从而得到可能存在人脸的区域。人脸检测问题所包含的内容十分广泛，是一个独立的研究方向。归纳起来，根据图像来源，可以将人脸检测方法分为基于静止图像的方法和基于动态图像的方法两类。其中前者是基础与核心，后者一般与人脸跟踪 (Face Tracking) 问题交织在一起。根据利用特征的色彩属性，可将人脸检测方法分为基于肤色特征的方法和基于灰度特征的方法两类，其中前者适用于构造快速的人脸检测算法，后者利用人脸区别于其它物体的最为本质的特征，是人脸检测领域研究的重点。根据利用人脸知识的不同，可以将人脸检测方法分为基于特征 (Feature-based) 的方法和基于图像 (Image-based) 的方法两大类，其中前者主要利用人脸的明显特征，如几何特征、肤色、纹理等，将人脸图像视为一个高维向量，从而将人脸检测问题转化为高维空间中分布信号的检测问题；后者则把人脸检测问题视为一个广义的模式识别问题，通过训练过程将样本分为人脸和非人脸两类。

由于人脸图像的复杂性，显示地描述人脸特征具有一定地困难，因此基于图像的人脸检测方法是该领域研究的重点。目前基于图像的人脸检测方法一般能够获得较好的检测效果 [70]，这类方法不需要提取具体的人脸特征，而是从大量典型的数据中获得统计特征，从而更好的反映人脸图像和非人脸图像之间的差别。因此，基于图像的人脸检测方法具有较强的适应能力和鲁棒性。虽然由于这类方法需要对所有可能的检测窗口进行穷举搜索导致计算的复杂度偏高，但从实际应用的效果来看，完全可以满足实时应用的要求。

Boosting [71,72] 方法是目前最常使用的统计集成学习的方法之一，该方法是可以将弱分类器组合成为强分类器的统计学习方法。其基本思想为对训练失败的训练样本赋以

较大的权值,让学习算法在后续的学习中集中对比较难的训练样本进行学习,最后将由算法挑选出来的若干个弱分类器加权相加组成强分类器。P. Viola 等 [73]在 2001 年提出了一种基于 Boosting 方法的实时人脸检测系统。其首先提出了一种称为“积分图像(IntegralImage)”的图像表示方法,这种表示方法提高了对特征的计算速度。第二,他们采用了 AdaBoost 方法对由积分图像表示的特征进行选择,选择出少量但是关键的特征来组合成强分类器。最后,他们使用“级联(Cascade)”的策略,每级的特征由少至多,在开始粗检的几级就排除了大量的非脸区域,因此该系统的检测速度可以达到每秒 15 帧。很多研究者以此为基础对 Adaboost 人脸检测方法做了很多改进 [74,75]。

鉴于 Adaboost 人脸检测方法提供了优异的检测速度,完全符合实时系统的要求,故本文采用 Adaboost 方法训练出来的人脸分类器作人脸检测。

2.2 人脸特征点定位

脸部特征点定位 [76],即在图象或图象序列的给定区域内搜索部分或所有人脸特征(如眼、鼻、嘴、耳等)的位置、关键点或轮廓线。这 3 种信息都有十分广泛的应用,例如,脸部特征的位置信息既可以在人脸检测中用于定位人脸、验证人脸检测的结果以及精确指明人脸位置,也可在人脸识别中用于人脸对齐,或作为识别的依据之一。在姿态识别中,它更是识别的重要依据之一;关键点信息可用于头部(Talking Head)图象的压缩和重构、脸部动画等领域,轮廓线信息则可用于图象的分割等。

近年来,国内外学者们已提出了许多种脸部特征定位方法。根据定位所依据的基本信息的类型,可以将现有方法分为基于先验规则、基于几何形状信息、基于色彩信息、基于外观信息和基于关联信息等 5 类:

(1) 基于先验规则

根据脸部特征的一般特点总结出一些经验规则,搜索前,先对输入图象作变换,使目标特征得到强化,而后根据上述规则从图中筛选出候选点或区域。

(2) 基于几何形状信息

根据脸部特征的形状特点构造一个带可变参数的几何模型,并设定一个评价函数度量被检测区域与模型的匹配度。搜索时不断调整参数使评价函数最小化,从而使模型收敛于待定位的脸部特征。

(3) 基于色彩信息

使用统计方法建立起脸部特征的色彩模型,搜索时遍历候选区域,根据被测点的色彩与模型的匹配度筛选出候选点。

(4) 基于外观信息

将脸部特征附近一定区域(窗口)内的子图象作为一个整体,映射为高维空间中的一个点,这样,同类脸部特征就可以用高维空间中的点集来描述,并可以使用统计方法得到其分布模型。在搜索中,通过计算待测区域与模型的匹配度即可判定其是否包含目标脸部特征。

(5) 基于关联信息

在局部信息的基础上, 引入脸部特征之间的相对位置信息, 以缩小候选点范围。

表 2.1 从运算量、准确率与鲁棒性 (包括图象质量需求和姿态表情光照等外界因素影响) 3 个方面对各类方法的性能作了粗略的比较。

表 2.1 脸部特征定位方法分类

Table 2.1 Methods of facial landmark localization

算法分类	运算量	准确率	鲁棒性	
			图象质量需求	外因影响
先验规则	镶嵌图	大	高	大
	几何投影	小		
	二值化定位	小		
	广义对称	大		
几何信息	Snakes	大	高	大
	可变形模板	大		
	ASM	较大		
色彩信息		小	低	较高
外观信息	神经网络			
	PCA	大	高	低
	SVM			较小
相关信息	概率网络			小
	DLA	较大	高	较低
	GWN			较小

2.3 基于投影峰分析的快速眼睛定位方法

很多研究人员在做识别实验时都假定面部关键特征点是已经给定的, 一些商业系统也往往在条件允许的情况下要求用户进行一定程度的交互 (如手动标定眼睛的位置)。目前的主流人脸识别算法, 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, 从而进行比对识别。而且人脸特征的误配准, 对识别率有很大的影响。Fisherface 是比较重要的人脸识别算法之一, 文献 [77] 指出, 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。因此精确而快速的定位眼睛是实用人脸识别系统的关键环节之一, 同时, 眼睛的定位也是其他特征点 (如鼻子, 嘴巴等) 定位的基础, 所以研究精确定位眼睛的算法至关重要。

在 2.2 节所述人脸特征点定位的方法中, 灰度积分投影是一类经典算法, 它利用脸部特征灰度值较皮肤低的特点, 先统计出 X 或 Y 方向上灰度值的和或者灰度函数值的和, 找出特定的变化点, 然后根据先验知识将不同方向上的变化点位置相结合, 即得到

脸部特征的位置，Kanade 等首先将这一方法应用于人脸识别 [4]。文献 [78] 中提出了 Generalized Projection Function，文献 [79] 根据人脸图像的面部比例特征，提出了一种基于候选窗口的灰度积分投影的方法来定位眼睛。灰度积分投影法计算量较低，但在大的姿态变化，光照变化或人脸有配饰（如眼镜）时容易失效。本节将在灰度积分投影的基础上，提出了一种基于投影峰的眼睛定位算法，在人脸姿态和光照变化及佩戴眼镜的情况下，可以准确快速地实现眼睛定位。

2.3.1 算法描述

本节所提出的算法是基于二值化后的人脸图像的，因此需要首先把窗口内的眼区图像截取出来，然后对眼区图像通过阈值分割实现二值化。下面将对算法的各个步骤分别进行描述：

(1) 候选窗口的选取

人脸检测的方法有很多，如2.1提到的 AdaBoost [80]等。在得到人脸的大致的位置后，可将人脸图截取出来（如图 2.1(a) 所示），然后在截取出来的图像上定位眼睛的位置（如图 2.1(b) 所示）。

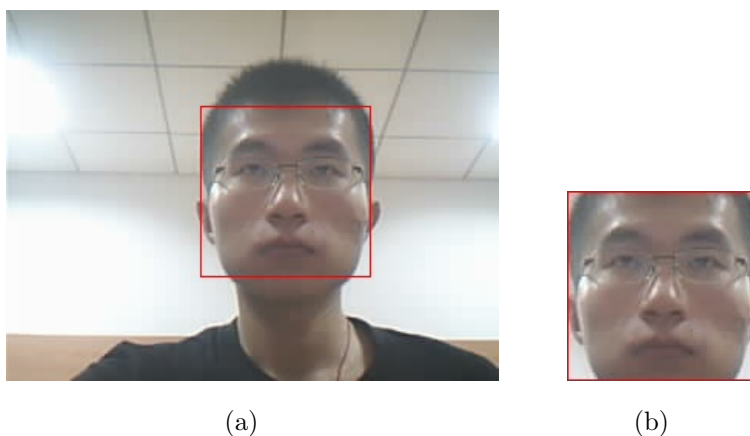


图 2.1 (a) AdaBoost 人脸检测 (b) 截取的人脸图像

Fig 2.1 (a) Face detection with Adaboost (b) Segmented face image

人脸的五官位置具有一定的统计规律，如双眼一定位于人脸的上半部分，眉毛一定位于眼睛上面，双眼一定对称分布于正面人脸的中轴线两侧等等。这些先验知识可以帮助缩小眼睛的搜索区域范围，这样不但可以去除一些干扰，而且还可以提高搜索速度。

确定的眼睛区域需要对人脸的姿态发生偏转或旋转变化的时候有一定的鲁棒性，图 2.2 显示了眼睛区域过小的情况。但如果窗口取得过大，将会引入过多的干扰，如眉毛，鬓角，刘海，眼镜框等，这些干扰的灰度值也相对较低，对通过使用灰度投影法来定位眼睛有很大的影响。

图 2.3 所示的候选窗口保证了人脸姿态发生变化的情况下，眼睛依然在候选窗口内。但我们需要把引入的少量的干扰（如眉毛，鬓角，刘海，眼睛框等）尽量去除。当候选窗口选定后，就可以把眼区截取出来，如图 2.4(a) 所示。

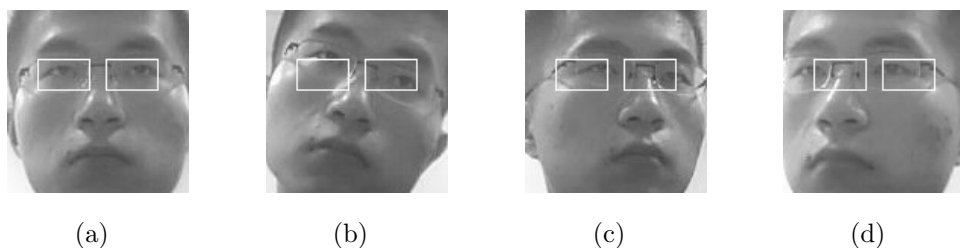


图 2.2 (a) 理想状态下的候选窗口, (b - d) 当人脸姿态发生变化时, 眼睛不在候选窗口内
Fig 2.2 (a) Candidate windows under ideal conditions (b - d) Eyes out of candidate windows when face pose changed

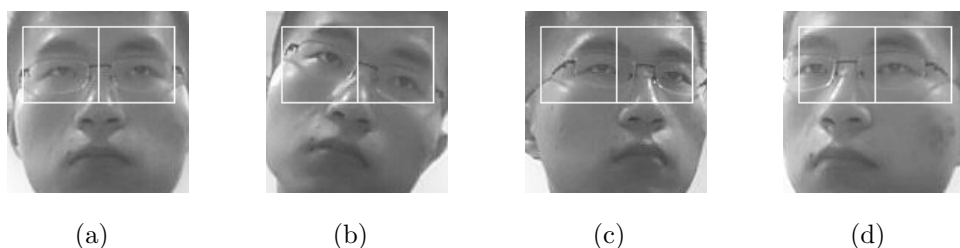


图 2.3 (a) 正脸情况下的候选窗口, (b - d) 当人脸姿态发生变化时, 眼睛仍在候选窗口内
Fig 2.3 (a) Candidate windows of frontal face (b - d) Eyes still in the candidate windows when face pose changed

(2) 阈值分割

图像阈值分割是一种广泛使用的图像分割技术 [81], 其关键在于阈值的选取, 常用的阈值选取方法有直方图阈值分割法、类间方差阈值分割法、二维最大熵值分割法等等。为了把灰度值较低的眼睛分割出来, 本文采用直方图分析的方法来选取阈值, 令灰度值最低的 $p\%$ 个像素点的值为 255, 其他像素点的值为 0。分割结果如图 2.4(b) 所示。

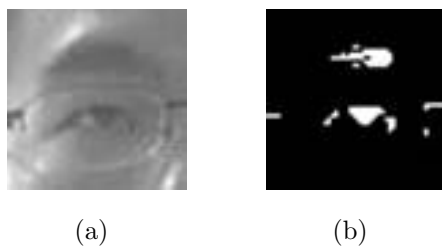


图 2.4 (a) 原始眼区图像, (b) 直方图阈值分割结果
Fig 2.4 (a) Original image of eye region (b) Image after gray-histograms threshold segmentation

(3) 灰度积分投影

理想情况下, 眼睛的中心的横坐标和纵坐标分别为垂直灰度投影曲线和水平灰度投影曲线峰值点对应坐标, 但过大的候选窗口会引入干扰, 如眉毛, 刘海, 鬓角, 眼镜边框等, 它们的灰度值可能甚至比眼睛瞳孔的灰度值还要低; 另外由于光照的原因, 使得眼

区内会出现阴影,这同样会对灰度投影产生影响。所以不能简单地认为水平和垂直灰度投影的峰值点为眼睛的坐标值。

(4) 投影积分峰分析

为解决上述问题,首先需要分析可能存在的干扰。水平方向从上到下眼区内可能存在灰度值较低的点依次为刘海,眉毛,眼镜上框,眼睛瞳孔,眼镜下框;垂直方向从左到右眼区内可能存在灰度值较低的点依次为鬓角,眼镜外侧边框,眼睛瞳孔,眼睛内侧边框。当然,在阈值分割之后,这些干扰点有时会同时出现,而有时会个别出现,有时甚至全部不会出现。其次,眉毛和瞳孔在眼区内的灰度相对比较低,而且和上面所说的干扰比起来,所占的面积也比较大,所以通过选取合适的阈值,总可以把它们分割出来(如果眉毛比较稀疏或颜色比较淡,可能分割不出来,但瞳孔总是可以分割出来的)。

以由阈值分割后的二值化图像得到的水平灰度投影曲线 $P(y)$ 为例。通过分析可以发现,投影曲线上会呈现若干个峰,记作 $[y_{11}, y_{12}], [y_{21}, y_{22}], \dots, [y_{n1}, y_{n2}]$, 其中 $y_{11}, y_{21}, \dots, y_{n1}$ 为投影峰的上升沿, $y_{12}, y_{22}, \dots, y_{n2}$ 为投影峰的下降沿。眼睛、眉毛和其他的一些干扰,分别与这些峰相对应,而且眼睛对应的峰宽度较宽,面积较大且偏离图像中心的位置较小。如果做垂直灰度投影,可以得到同样的结果。

根据这一特性,我们提出一种新的基于投影峰的快速眼睛定位算法。对于每一个投影峰,我们定义评价值 U_n 见公式(2.1),

$$U_n = \alpha W_n + \beta S_n + \gamma D_n, \quad (2.1)$$

$$W_n = y_{n2} - y_{n1}, \quad (2.2)$$

$$S_n = \sum_{y=y_{n1}}^{y_{n2}} P(y), \quad (2.3)$$

$$D_n = \left| \frac{y_{n2} - y_{n1}}{2} - Y_c \right|, \quad (2.4)$$

其中 W_n, S_n, D_n 分别为投影峰的宽度、面积和与图像中心的距离,见公式 (2.2-2.4), α, β, γ 为权重。公式 (2.4) 中 Y_c 是图像沿 y 轴方向上的中心。

对所有的投影峰求 U 值,并排序。 U 值最大的峰可认定是眼睛所在区域对应的投影峰;在此峰内求最大值点,即可得到眼睛瞳孔中心位置的横坐标或纵坐标。

算法步骤如下(以求眼睛中心纵坐标为例):

1. 对阈值分割出来的图像做水平灰度积分投影;
2. 消除积分曲线上的毛刺;
3. 搜索曲线上的投影峰;
4. 求取所有投影峰的 U 值,排序, U 值最大的投影峰对应眼睛所在区域;
5. 在确定的眼睛所处区间内,求灰度积分值最大的点,则该点对应眼睛瞳孔中心的纵坐标。瞳孔中心的横坐标的求取与上述步骤类似。

投影结果见图 2.5。

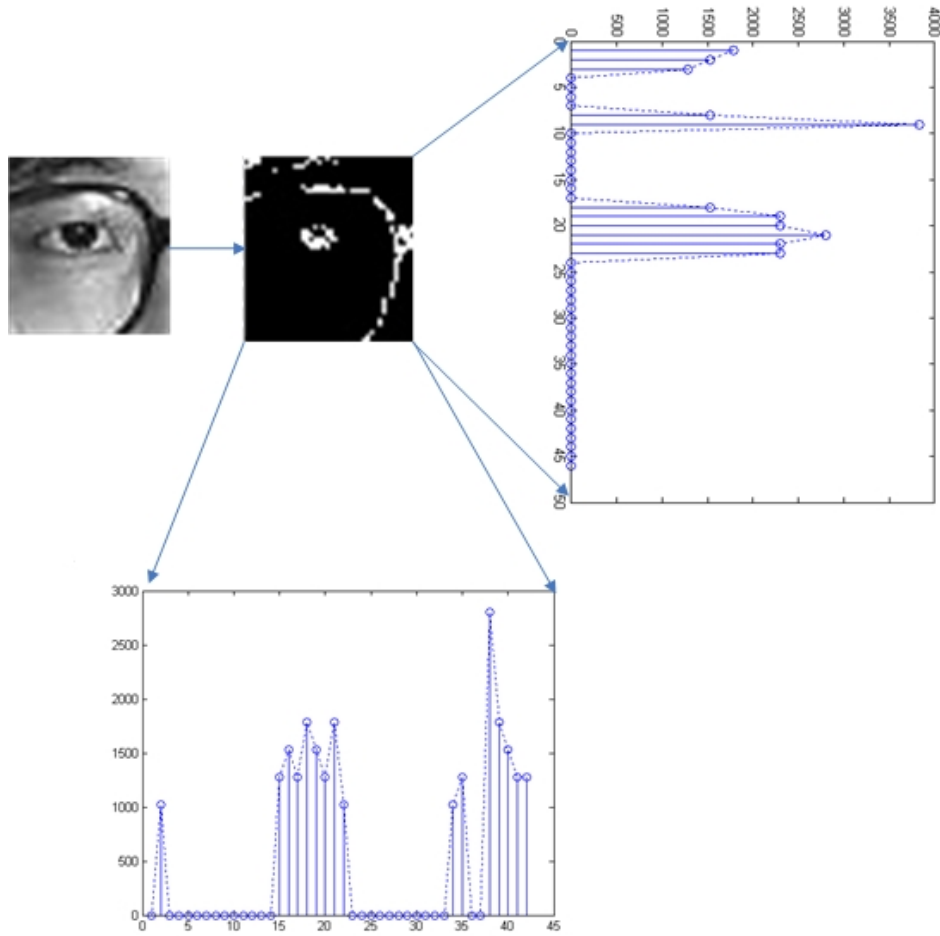


图 2.5 水平和垂直投影峰

Fig 2.5 Horizontal and vertical projection peaks

2.3.2 实验与分析

为了测试算法的有效性，用C++编程实现了上述算法，进行实验。实验中采用文献 [82]给出的眼睛定位精度评价标准，设 C_l 和 C_r 为手工标定的左右眼睛位置， C'_l 和 C'_r 为自动定位的左右眼睛位置， d_l 为 C_l 和 C'_l 间的欧式距离， d_r 为 C_r 和 C'_r 间的欧式距离，则眼睛定位的相对误差如公式 (2.5) 所定义

$$err = \frac{\max(d_l, d_r)}{\|C_l - C_r\|}. \quad (2.5)$$

如果 $err < 0.25$ ，则认为定位是准确的。所以对于包含N张人脸图片的数据库，定位的准确率为：

$$rate = \sum_{\substack{i=0 \\ err_i < 0.25}}^N \frac{1}{N} \times 100\%, \quad (2.6)$$

其中 err_i 为第 i 张人脸图片的相对误差。

(1) 标准人脸库测试

FERET [83] 人脸库中共包含 14051 张人头肩部的图像, 其中手工标定双眼坐标的图像有 3816 张, 在这 3816 张人脸图像中把人脸区域截取出来作为眼睛定位算法的测试集。BioID 人脸库 [84] 中包含 1521 张手工标定双眼坐标的人脸图像, 也被用作算法的测试集。JAFPE 人脸库 [85] 中包括 213 张正脸图像, 这些人脸图像有比较明显的表情变化, 该人脸库用于测试本算法在不同表情下的定位效果。

设人脸图像的宽度为 $width$, 高度为 $height$ 。实验中我们按照如下规则选择双眼候选窗口: 左眼窗口原点为 $(width/12, height/12)$, 右眼窗口原点为 $(width/2, height/12)$, 双眼窗口大小均为 $(5 \times width/12, 5 \times height/12)$ 。

用于测试的部分人脸图像如图 2.6 所示, 样本包括了各种姿态、光照、表情和佩戴眼镜的人脸图像。

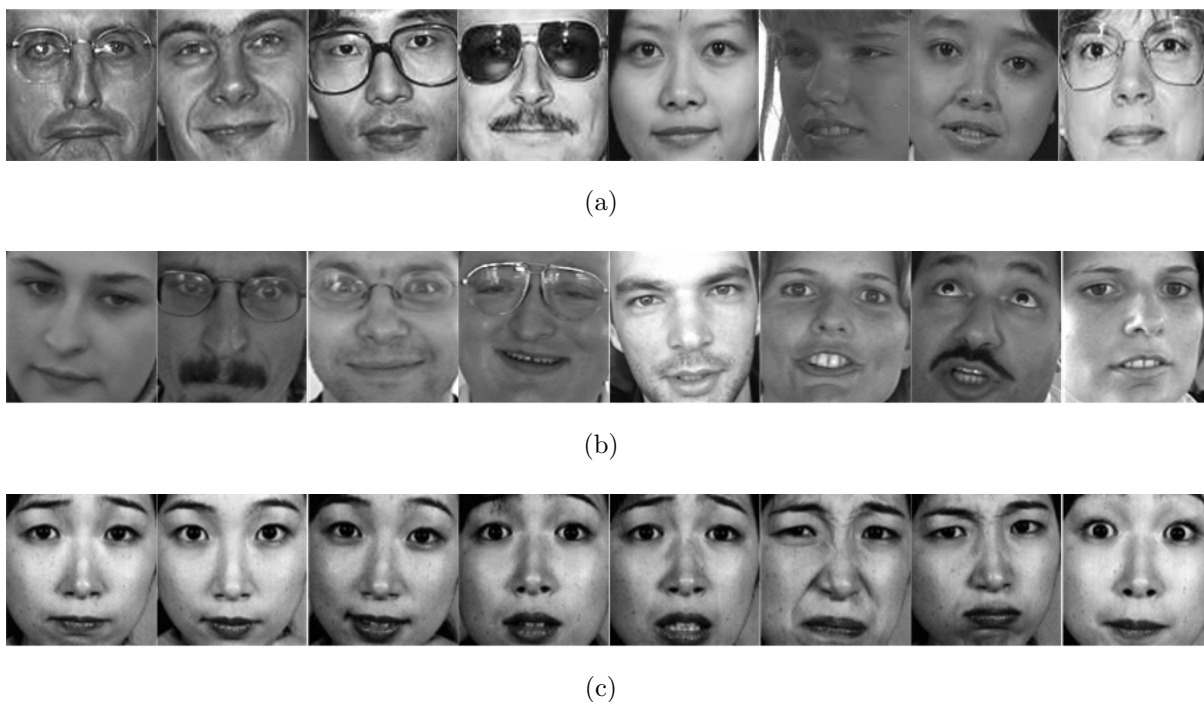


图 2.6 用于测试的部分人脸图像样本: (a) FERET 中的部分样本, (b) BioID 中的部分样本, (c) JAFPE 中的部分样本

Fig 2.6 Some images from experimental database: (a) Sample images from FERET (b) Sample images from BioID (c) Sample images from JAFPE

测试结果如表 2.2 所示, 表中 p 为直方图阈值分割的阈值。实验表明当 $p = 5$ 时, 本文提出的方法对三个标准人脸库的人眼定位准确率相对最高, 下表中用黑体字标出。

通过本文方法定位出双眼位置的部分人脸图像如图 2.7 所示。其中图 2.7(a)(d)(f) 分别为 FERET, BioID 和 JAFPE 人脸库中部分正脸样本的定位的结果, 实验结果表明, 当光照、人脸表情发生变化时, 本文的方法仍可以准确定位出眼睛的位置; 图 2.7(c) 为

表 2.2 标准人脸库上算法定位的准确率

Table 2.2 Rates of localization accuracy on standard face database

p(%)	FERET(%)	BioID(%)	JAFFE(%)
1	79.18	71.79	84.62
3	93.19	85.78	90.36
5	98.92	95.87	99.26
8	96.19	92.63	97.10
10	95.71	86.06	96.51

FERET 人脸库中部分侧脸样本的定位结果, 当人脸姿态发生变化时, 本文提出的方法依然有效; 图 2.7(b)(e) 分别为 FERET 和 BioID 人脸库中部分佩戴眼睛的人脸样本的定位结果, 结果表明, 我们的方法可以有效地去除眼镜的干扰, 实现眼睛位置的精确定位; 图 2.7(g) 为部分错误定位的样本, 通过对这些样本的分析, 发现当人脸受到强光照射眼镜发生强烈反光, 可能导致眼睛瞳孔不完全可见, 从而出现错误定位。

(2) 与其他方法的比较

表 2.3 与其他方法的定位准确率及平均耗时的比较

Table 2.3 Comparison with other methods

	Our Method	Adaboost	HPF
FERET(%)	98.92	98.93	—
BioID(%)	95.87	96.03	94.81
JAFFE(%)	99.26	99.47	97.18
Time(ms)	0.56	10.36	0.49

文献 [80] 中提出的 Adaboost 是一种通用且有效的物体检测的方法, 我们为了得到用于检测眼睛的 Adaboost 分类器, 选取 4532 张 20×20 的眼睛图片为正样本, 每层取 2236 张不含眼睛的图片为负样本, 训练出 16 层的分类器。在上文设置的候选窗口内使用这个分类器作眼睛定位, 并在上述的三个标准人脸库上测试, 得到的实验结果见表 2.3。文献 [78] 给出的在 BioID 和 JAFFE 人脸库上测试结果和本文的测试结果也列于表 2.3 中。比较三种方法的实验结果可以发现, 我们的方法在保证较高定位准确率的同时平均耗时也较少, 此外我们的方法不需要训练和学习过程, 便于实现。

(3) 实时系统测试

本节提出的眼睛定位算法, 在 Intel P4 1.8G CPU, 512M 内存配置的 PC 机上, 对于 128×128 大小的人脸图像, 定位双眼中心坐标耗时约 0.56ms。本节所述方法已应用于 RCIR@SJTU 自主开发的人脸识别系统上 (图 2.8)。准确的人眼定位是人脸识别系统具有高识别率的重要保障。

2.3.3 方法小结

基于投影峰分析的眼睛快速定位方法, 充分利用了人脸五官特征的统计规律, 通过



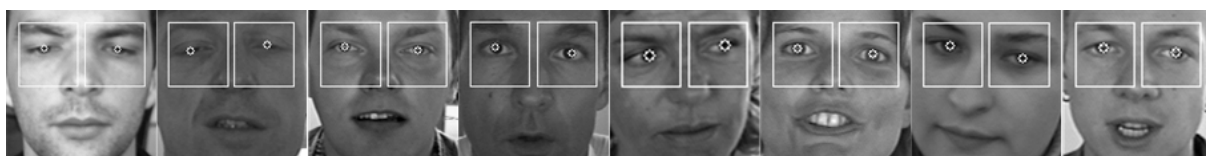
(a)



(b)



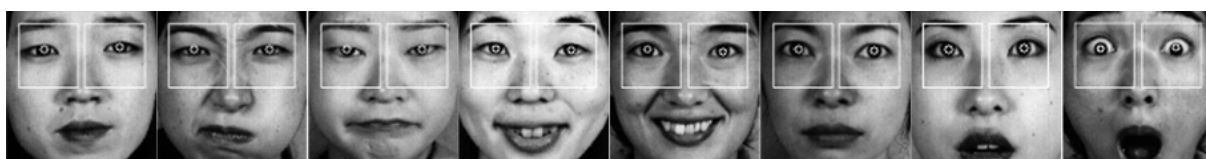
(c)



(d)



(e)



(f)



(g)

图 2.7 部分眼睛定位的结果: (a) FERET库内正脸样本, (b) FERET库内戴眼镜的样本, (c) FERET库内侧脸样本, (d) BioID库内正脸样本, (e) BioID库内戴眼镜的样本, (f) JAFFE库内的样本, (g) 部分错误定位的样本



图 2.8 实时环境下的眼睛定位

Fig 2.8 Real-time eye localization

简单的运算实现眼睛中心坐标的准确快速的定位，针对额发，鬓角，眉毛，眼镜等干扰，提出基于投影峰的定位方法，有效地去除了上述干扰，使得定位算法更具鲁棒性，此外当人脸姿态发生俯仰 ($\pm 15^\circ$)，旋转 ($\pm 15^\circ$) 等变化时，同样可以实现精确定位。通过研究误定位的样本发现，当人脸图像受到强光照眼镜发生强烈反光或眼睛完全闭合时，由于瞳孔不可见，导致算法失效，对于这种情况，需要进一步研究解决。

2.4 人脸标准化

为了保证原型集 (Gallery) 和测试集 (Probe) 中人脸图像尽可能的统一，需要对人脸图像进行标准化。当完成人脸特征点定位之后，可根据各个特征点的位置，把人脸图像从整帧视频图像中截取出来，然后通过旋转、剪切与缩放、预处理和Mask的步骤，以完成人脸图像的标准化，最终使得所有人脸图像具有相同的大小，相同的双眼位置，经过相同的预处理过程和Mask。本节将详细描述以上过程。

2.4.1 旋转

当人脸图像中，双眼位置不在同一水平线上时，需要在图像平面内对人脸图像进行旋转变换，以使在旋转后的图像中，双眼位于同一水平线上。

设左眼中心坐标为 (x_1, y_1) ，右眼中心坐标为 (x_2, y_2) ，则双眼中心连线与水平方向的夹角为：

$$\theta = \arctan \frac{y_2 - y_1}{x_2 - x_1}. \quad (2.7)$$

当 $\theta \neq 0$ 时, 需对图像进行如下旋转变换:

$$x_1 = x \cos \theta + y \sin \theta - a \cos \theta - b \sin \theta + a, \quad (2.8)$$

$$y_1 = -x \sin \theta + y \cos \theta + a \sin \theta - b \cos \theta + b, \quad (2.9)$$

$$P(x_1, y_1) = P(x, y), \quad (2.10)$$

其中, (x_1, y_1) 为旋转后图像像素点的坐标, (x, y) 原图像像素点的坐标, (a, b) 为未旋转前图像的中心坐标, $P(x_1, y_1)$ 和 $P(x, y)$ 分别为旋转前后图像像素点的值。假设旋转前后图像的高度和宽度不变, 图像以图像的中心为原点进行旋转, 旋转后的图像中心坐标仍为 (a, b) 。图 2.9(b) 显示了图 2.9(a) 经过旋转变换后的结果。

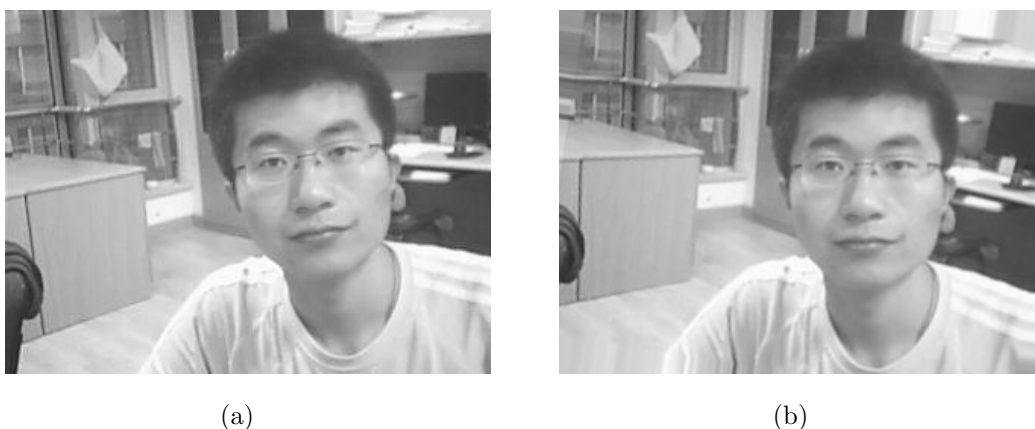


图 2.9 (a) 原始图像, (b) 旋转变换后图像

Fig 2.9 (a) Original image (b) Image after rotation transform

2.4.2 剪切与缩放

经过旋转变换, 确保双眼在同一水平线上后, 可对视频图像进行剪切, 把人脸图像截取出来。本文使用的人脸剪切规则如图 2.10 所示。其中 w 和 h 为人脸图像的宽度和高度, 设定 $w = h$, 双眼间距为 $w/2$, 双眼与图像左右边界的距离为 $w/4$, 眼睛与图像上边界和下边界的距离分别为 $h/4$ 和 $3h/4$ 。

完成剪切后, 把人脸图像缩放至统一的尺寸。本文使用的人脸图像尺寸为 128×128 。

图 2.9(b) 的视频图像经剪切与缩放变换后的人脸图像见图 2.11 所示。

2.4.3 预处理

直方图均衡 [81] 是一种常用的人脸图像预处理的方法, 其利用直方图的统计数据进行直方图的修改以通过某种对应关系改变图像中各点灰度值来达到图像增强的目的, 是一种基于直方图的调整方法。为使图像清晰, 合乎逻辑的想法是把灰度级分布拉开, 使灰度层次分明, 这相当于增大了图像对比度。理想的直方图均衡处理的目的是使处理后

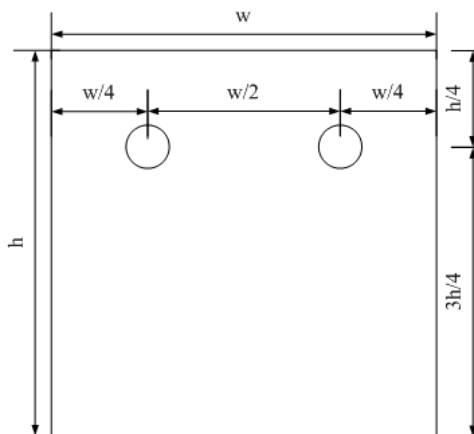


图 2.10 人脸图像剪切规则

Fig 2.10 Rule of face image segmentation



图 2.11 剪切缩放后的人脸图像

Fig 2.11 Face image after segmentation and resize

的图像的直方图是平直的，即各灰度级具有相同的出现频数，由此各灰度级具有均匀的概率分布，使图像看起来就更清晰。变换后的新灰度图像应该具有平直的直方图。

为保持变换后的灰度仍然具有从暗到明的单一变化顺序，并考虑到灰度变换不影响像素位置分布，也不会增减像素的数目，因此有公式 (2.11)。

$$\int_{r_j}^{r_j+\Delta r} \Phi_r(r) dr = \int_{s_j}^{s_j+\Delta s} \Phi_r(s) ds, \quad (2.11)$$

其中 r 和 s 分别表示变换前后的灰度值， $\Phi_r(r)$ 和 $\Phi_s(s)$ 分别表示变换前后的直方图密度函数。通过矩法近似求积，最终可得到公式 (2.12)。

$$T(r) = \int_0^r \Phi_r(r) dr, \quad (2.12)$$

这是计算直方图均衡的灰度变换式。并且由此可得它的离散灰度形式，见公式 (2.13)。

$$S_k = T(r_k) = \sum_{j=0}^k P_r(r_j) = \sum_{j=0}^k n_j/n, \quad (2.13)$$

其中，图象的象素总数 n ，分 L 个灰度等级， n_k 代表第 k 个灰度级 r_k 出现的频数，于是第 k 个灰度级出现的概率为，

$$P_r(r_j) = n_k/n, \quad (2.14)$$

因此,可以根据原图象的直方图统计量,求得均衡后各象素的灰度级变换值。在离散灰度值下,直方图只能近似于密度函数,所以很少能在均衡后获得完全平直的直方图。

图 2.11 的人脸图像经直方图均衡后的图像见图 2.12 所示。



图 2.12 直方图均衡后的人脸图像

Fig 2.12 Face image after histogram equalization

2.4.4 人脸图像Mask

经过几何归一化和直方图均衡以后的人脸图像,还要经过 Mask 来去除图像边缘的背景干扰。图 2.13 显示了图 2.12 Mask后的结果。



图 2.13 Mask后的人脸图像

Fig 2.13 Face image after mask

图 2.14 显示了 FERET 人脸数据库的几张原始的人脸样例图像。图 2.15 显示了图 2.14 中的几张原始人脸图像经过人脸检测,人脸特征点定位,几何归一化,直方图均衡和 Mask 以后的归一化人脸图像。



图 2.14 FERET人脸数据库样例图像

Fig 2.14 Samples of original face images in FERET

本文后续几章中对人脸识别相关算法的研究均以图 2.15 所示归一化的标准人脸图像为基础。



图 2.15 人脸标准化后图像

Fig 2.15 Samples of face images after face normalization

2.5 本章小结

本章叙述了自动人脸识别系统中必不可少的两个环节：人脸检测，人脸特征点定位及人脸标准化。首先简要介绍人脸检测的一般原理与和常用方法，并总结了人脸特征点定位的一般方法；然后提出一种基于投影峰分析的眼睛快速定位方法；最后在完成眼睛定位的基础上，实现人脸图像的标准化，以此为基础研究自动人脸识别系统的相关算法。

第三章 基于 Gabor 的人脸识别

良好的人脸表示是鲁棒高效的人脸识别算法的关键因素，所以人脸图像的特征提取对于人脸识别非常重要。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，是一种良好的人脸特征提取方法。传统的弹性图匹配方法 (EBGM) 和 Gabor 特征 Fisher 判别分类方法就是应用 Gabor 特征的典型人脸识别算法。我们在研究人脸识别的过程中也对人脸图像的特征提取方法进行了研究，发现了 Gabor 小波特征能够对人脸图像进行有效的表达。Gabor 小波特征可以比原始像素特征具有更强的抽象概括能力，所以更适合在特征选择过程中使用。本章首先介绍 Gabor 小波，接着描述基于 FLD 的人脸识别方法，然后分析该方法存在的不足，最后通过特征选择，提出一种新的基于 Gabor 小波的人脸识别方法。

3.1 Gabor 小波变换

Gabor 小波变换 (Gabor Wavelet Transform, GWT) 是多分辨率图像表达里面一个非常重要的方法。Gabor 小波是一种二进有向小波。二进有向小波是计算机视觉领域里面经常使用的一种小波。纹理的综合和辨识经常使用这种小波。

1985年，Daugman 将 Gabor 函数扩展成二维形式 [86]，并在此基础上构造了二维 Gabor 滤波器。人们惊奇地发现，二维 Gabor 滤波器不仅可以同时获取时间和频率域的最小不确定性，而且与哺乳动物视网膜神经细胞的接收场模型相吻合。Gabor 滤波器及其在图像处理、理解、识别等方面的应用研究得到了广泛关注。同样，利用 Gabor 函数推导得到的 Gabor 小波，在特征提取和跟踪方面也取得了较多的应用。Gabor 小波进行边缘检测 [87]，脸部图像跟踪 [88, 89]，取得了很好的效果。

传统 Gabor 变换的缺陷是窗口尺寸一旦确定后，在信号分析过程中就无法更改，也就是说它只能在一个分辨率上进行分析。为了解决这一问题人们将 Gabor 变换与小波理论相结合提出了 Gabor 小波。因此，Gabor 小波具有小波变换的多分辨率特性，同时具有 Gabor 函数本身所具有的局域性和方向性。

Gabor 小波特征描述方法是一种基于灰度和基于特征的综合方法。这种方法同时具有基于灰度和基于特征的两种方法的优点 [90]。

1. 因为描述是基于特征的，对于几何变形和亮度的变化不敏感；
2. 与一般的基于特征的方法相比，除了丢失了图像的灰度均值信息以外，小波特征可以重建原始图像。因此，没有损失有用的图像数据；
3. 可以根据需要使用特征数量选取，实现可变精度的跟踪。鉴于上述 Gabor 小波特征的优点，使用 Gabor 小波提取样本的特征。

鉴于上述 Gabor 小波特征的优点, 使用 Gabor 小波提取样本的特征。

二维二进小波变换用几个通常具有不同空间方向的母小波 $\{\psi^k\}_{1 \leq k \leq K}$ 来计算。对于 $x = (x_1, x_2)$ 有

$$\psi_{2^j}^k(x_1, x_2) = \frac{1}{2^j} \psi^k\left(\frac{x_1}{2^j}, \frac{x_2}{2^j}\right), \quad (3.1)$$

和

$$\overline{\psi_{2^j}^k}(x) = \psi_{2^j}^k(-x), \quad (3.2)$$

在位置 $u = (u_1, u_2)$ 和尺度 2^j , $f \in L^2(R^2)$ 在方向 k 的小波变换定义为

$$W^k f(u, 2^j) = \langle f(x), \psi_{2^j}^k(x - u) \rangle = f * \overline{\psi_{2^j}^k}(u), \quad (3.3)$$

如果存在 $A > 0$ 和 B 使得

$$\forall \omega = (\omega_1, \omega_2) \in R^2 - \{(0, 0)\}, A \leq \sum_{k=1}^K \sum_{j=-\infty}^{\infty} |\hat{\psi}^k(2^j \omega)|^2 \leq B. \quad (3.4)$$

所以, 二维二进小波变换是完全和稳定的信号表示。并且存在重构小波 $\{\tilde{\psi}^k\}_{1 \leq k \leq K}$, 它的傅立叶变换满足

$$\sum_{j=-\infty}^{+\infty} \frac{1}{2^{2j}} \sum_{k=1}^K \tilde{\psi}^k(2^j \omega) \hat{\psi}^{k*}(2^j \omega) = 1, \quad (3.5)$$

可以得出

$$f(x) = \sum_{j=-\infty}^{+\infty} \frac{1}{2^{2j}} \sum_{k=1}^K W^k f(\cdot, 2^j) * \tilde{\psi}_{2^j}^k(x), \quad (3.6)$$

满足公式 3.4 小波就是二进小波。

沿任何角度 α 的有向小波可以设计为 K 个母小波的线性展开式。例如, 在方向 α 的小波可以定义为窗口函数 $\theta(x)$ 沿方向向量 $\vec{n} = (\cos \alpha, \sin \alpha)$ 的 p 阶偏导数

$$\psi^\alpha(x) = \frac{\partial^p \theta(x)}{\partial \vec{n}^p} = \left(\cos \alpha \frac{\partial}{\partial x_1} + \sin \alpha \frac{\partial}{\partial x_2} \right)^p \theta(x), \quad (3.7)$$

这个偏导数是 $K = p + 1$ 个母小波的线性展开式

$$\psi^\alpha(x) = \sum_{k=0}^p \psi^k(x), \quad (3.8)$$

其中

$$\psi^k(x) = \frac{\partial^p \theta(x)}{\partial x_1^k \partial x_2^{p-k}}, 0 \leq k \leq p \quad (3.9)$$

对于合适的窗口 θ , 这个 $p + 1$ 个偏导数定义一族二进小波

$$W^\alpha f(u, 2^j) = f * \bar{\psi}_{2^j}^\alpha(u), \quad (3.10)$$

可以用展开式3.8中 $p + 1$ 个分量

$$W^k f(u, 2^j) = f * \bar{\psi}_{2^j}^\alpha(u), \quad (3.11)$$

计算得到。

Gabor 小波特征表达被证明从空间和频率域二维联合不确定性的最小化意义上讲是最优的。Gabor 小波滤波器可以被认为是方向和尺度可以调节的边缘和线条检测器。在图像给定区域的 Gabor 小波特征的统计特性可以用来刻画该图像区域的纹理信息 [51, 91, 92]。人脸识别领域经典方法之一的弹性图匹配 (Elastic Bunch Graph Matching, EBGM) 方法被提出来 [93]。在弹性图匹配方法中 Gabor 小波被非常成功地应用于人脸识别中。因此我们使用 Gabor 小波作为人脸识别的主要特征提取方法。Gabor 小波核具有位置, 方向和频率选择特性, 类似于大脑皮层的简单细胞的响应域。

基于 Gabor 小波变换, 一族 Gabor 核可以被定义为

$$\psi_j(\vec{x}) = \frac{\vec{k}_j^2}{\sigma^2} \exp\left(-\frac{\vec{k}_j^2 \vec{x}^2}{2\sigma^2}\right) \left[\exp(i\vec{k}_j \vec{x}) - \exp\left(-\frac{\sigma^2}{2}\right) \right]. \quad (3.12)$$

在平面波形向量 $\vec{k}_j$ 的形状里面, 是被相对宽度为 $\sigma = 2\pi$ 的高斯封套函数所限制。这里 $\vec{x} = (x, y)$ 是一个在空间域里面的变量。 $\vec{k}_j$ 是一个频率向量, 它决定了 Gabor 核的尺度和方向。

$$\vec{k}_j = k_s e^{i\phi_d}, \quad (3.13)$$

其中 $k_s = \frac{k_{max}}{f^s}$, $k_{max} = \frac{\pi}{2}$, $f = 2$, $s = 0, 1, 2, 3, 4$, and $\phi_d = \frac{\pi d}{8}$, for $d = 0, 1, 2, 3, 4, 5, 6, 7$. 我们使用了一个离散的集合, 其中包括五个不同的空间频率和八个不同的方向。公式 (3.12) 的最后一项使 Gabor 核 DC-free, 即积分 $\int \psi_j(\vec{x}) d^2 \vec{x}$ 被去除了, 这样保证了 Gabor 核对光照的不敏感性。小波变换是一族相互类似的小波核, 它们都是由同一个母小波通过膨胀和旋转得到的。给定一幅图像 $I(\vec{x})$, 它在特定位置 $\vec{x}_0$ 上面的 Gabor 变换可以通过和 Gabor 核卷积来得到

$$(\psi_j * I)(\vec{x}_0) = \int \psi_j(\vec{x}_0 - \vec{x}) I(\vec{x}) d^2(\vec{x}). \quad (3.14)$$

Gabor 滤波器在三维中的显示如图 3.1, 图 3.1(a) 代表实部 (偶函数), 图 3.1 代表虚部 (奇函数)。

Gabor 滤波器中有两个参数 σ 和 ω , 它们分别决定滤波器的尺度和方向。在实际应用中, 为了得到一组合理的滤波器族, 就需要对参数 σ 和 ω 的取值进行合理的控制。一个直观的想法就是保证在不同尺度和不同方向上的采样尽量均匀, 让 Gauss 窗口的面积以成倍的方式递减, ω 在区间 $[0, \pi)$ 上等间隔取值。定义形式如下,

$$k_v = 2^{\frac{v+2}{2}} \pi, \quad v = 0, \dots, M-1 \quad (3.15)$$

$$\varphi_\mu = \mu \frac{\pi}{N}, \quad \mu = 0, \dots, N-1 \quad (3.16)$$

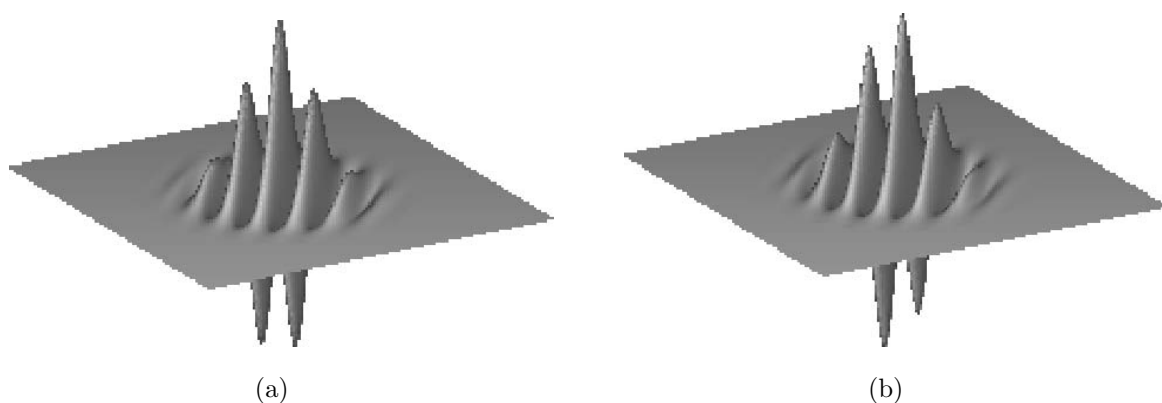


图 3.1 (a)Gabor 滤波器实部, (b)Gabor 滤波器虚部

Fig 3.1 (a) Real parts of Gabor filters. (b) Imaginary parts of Gabor filters.

在公式 (3.15) 和 (3.16) 中, 选取不同的下标 v , 可以描述不同的 Gauss 窗的波长, 从而控制采样的尺度, 或者说是频率; 而选取不同的下标 μ , 则可以描述振荡函数不同的振荡方向, 从而控制采样的方向。这样就可以得到 $M \times N$ 个 Gabor 滤波器; φ 只需要在 $[0, \pi)$ 区间上进行采样, 是因为此区间可以描述所有的方向。即, $\varphi_\mu + \pi$ 方向上的滤波结果完全由 φ_μ 方向上的结果确定。本文中, 选取 $M = 5, N = 8$ 就可以得到 40 个滤波器, 如图 3.2 所示。

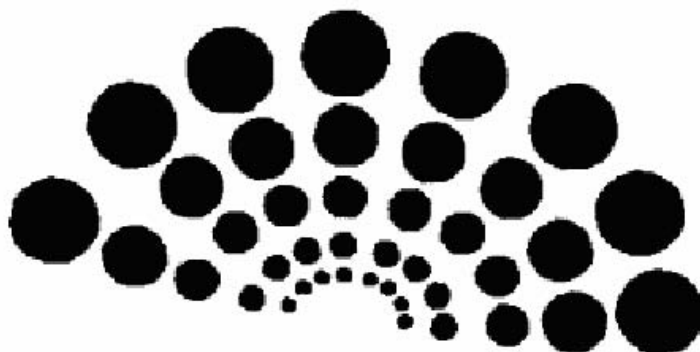


图 3.2 Gabor 滤波器族

Fig 3.2 Gabor filter family

3.2 人脸的 Gabor 表示

图像的 Gabor 小波表示可以通过图像与公式 (3.12) 定义的 Gabor 基簇的卷积运算得到。在本文中, 公式 (3.12) 中各个参数取为: $\sigma = 2\pi, k_{max} = \pi/2, f = \sqrt{2}$ 。按照上述参数, 得到的 5 个尺度 8 个方向共 40 个 Gabor 基的实部如图 3.3 所示。

设 $I(z), z = (x, y)$ 表示一幅灰度人脸图像, $\psi_{\mu,v}(z)$ 为对应尺度 v 和方向 μ 的 Gabor 基, 则有公式 (3.17)

$$O_{\mu,v}(z) = I(z) * \psi_{\mu,v}(z), \quad (3.17)$$

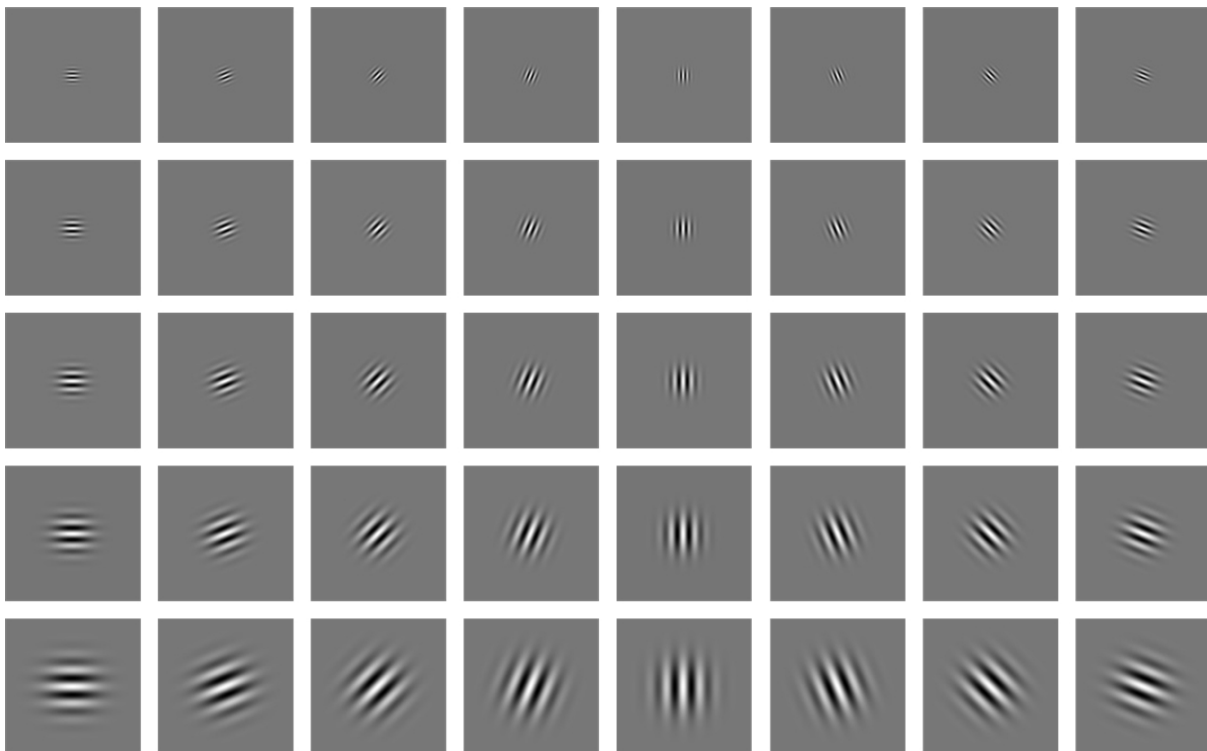


图 3.3 Gabor 基实部

Fig 3.3 The real part of the Gabor kernels

其中 $*$ 表示卷积运算, 则 $O_{\mu,v}(z)$ 表示图像在尺度为 v , 方向为 μ 时的 Gabor 描述。集合 $S = \{O_{\mu,v}(z) : \mu \in \{0, \dots, 7\}, v \in \{0, \dots, 4\}\}$ 构成了人脸图像 $I(z)$ 的 Gabor 描述。

根据卷积定理, 可以利用快速傅立叶变换(Fast Fourier Transform, FFT), 计算 $O_{\mu,v}(z)$,

$$\Im\{O_{\mu,v}(z)\} = \Im\{I(z)\} \Im\{\psi_{\mu,v}(z)\} \quad (3.18)$$

$$O_{\mu,v}(z) = \Im^{-1}\{\Im\{I(z)\} \Im\{\psi_{\mu,v}(z)\}\} \quad (3.19)$$

其中 $\Im$ 和 $\Im^{-1}$ 分别为傅立叶变换和傅立叶反变换。图 3.4 所示人脸样本与图 3.3 中 40 个 Gabor 基的实部卷积的结果见图 3.5, 再与 Gabor 基虚部做卷积运算后, 可以求模做为人脸 Gabor 的最终表示, 见图 3.6。



图 3.4 人脸样本例图

Fig 3.4 A sample of facial image

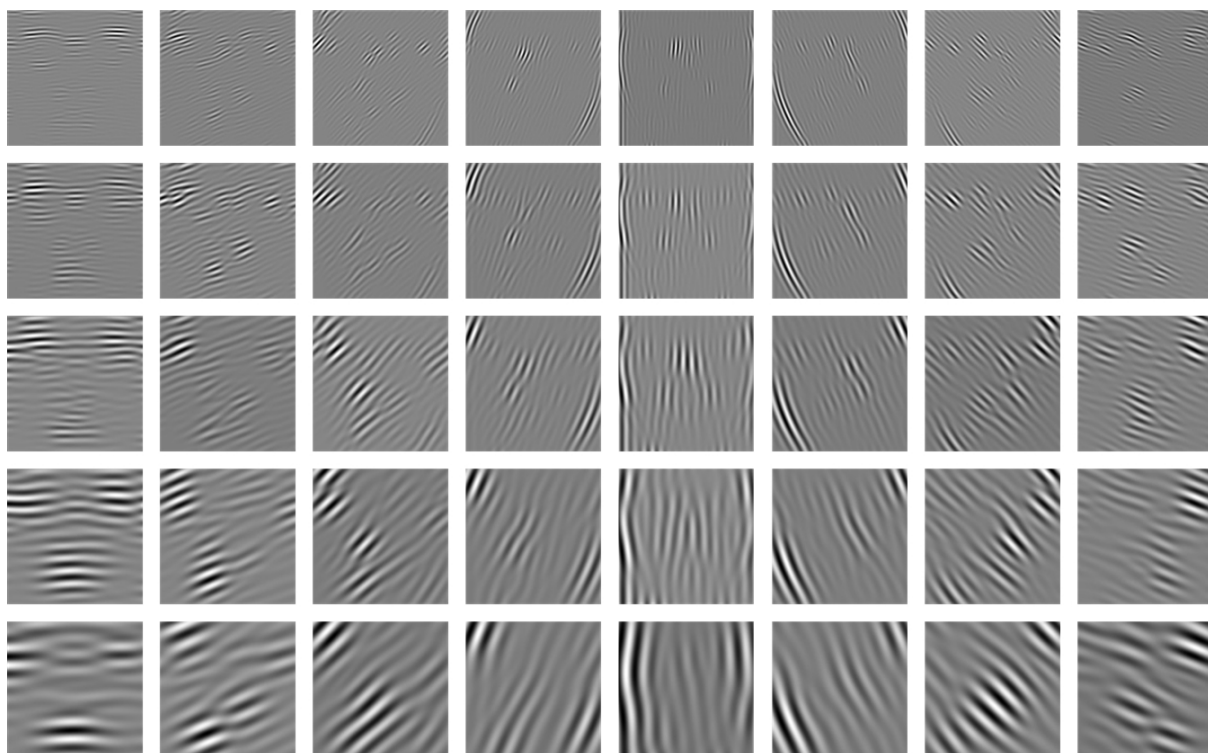


图 3.5 人脸 Gabor 表示的实部

Fig 3.5 The real part of Gabor wavelet representation of a sample image

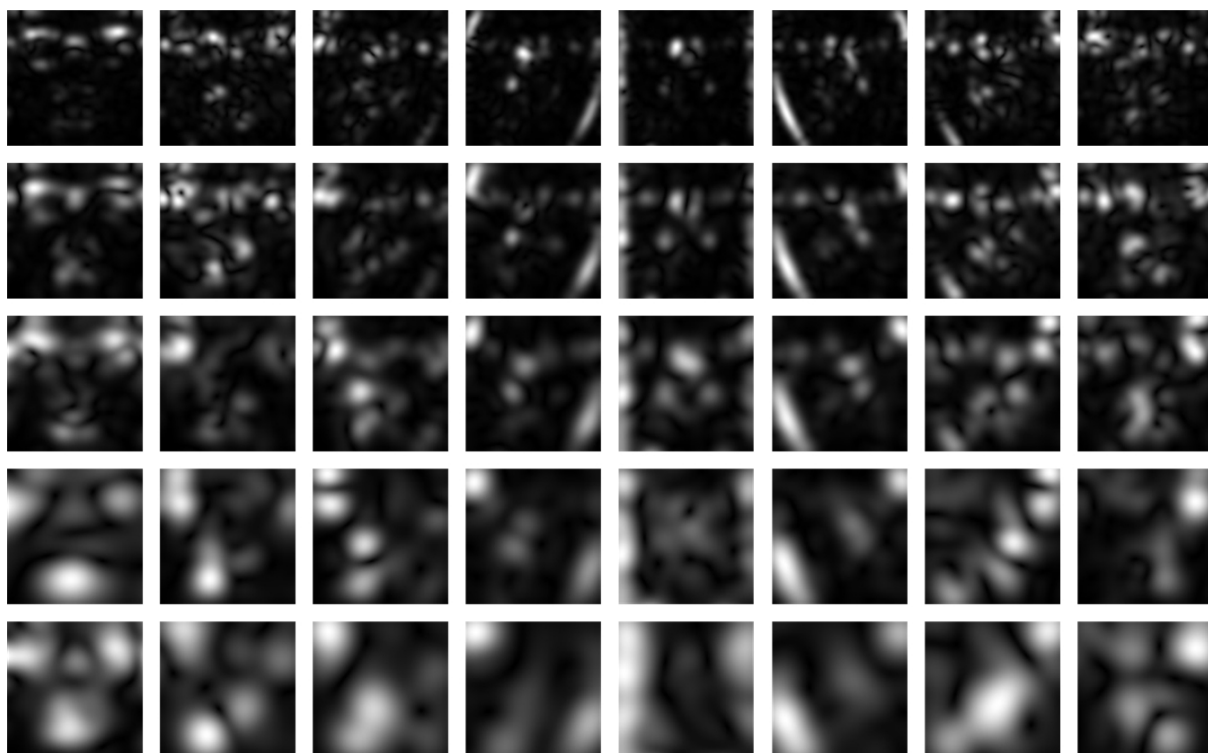


图 3.6 人脸 Gabor 表示的模

Fig 3.6 The magnitude of Gabor wavelet representation of a sample image

维数分析

本文中使用的脸样本图像大小为 128×128 ，若把 40 个 Gabor 模表示连接起来，此时人脸特征的维数为 $655360 (= 40 \times 128 \times 128)$ ，如此高的特征维数，不利于分类，而且无法满足实时应用的要求。因此，我们需要对当前的人脸 Gabor 模表示进行降维处理。

3.3 PCA+FLD 人脸识别方法

3.3.1 图像下采样和归一化

图像下采样是指从高分辨率的图像获得低分辨率的图像的过程。为了降低 Gabor 模表示的人脸特征的维数，首先对每幅 Gabor 模表示的人脸图像 $O_{\mu,v}(z)$ 进行均匀下采样。设均匀下采样的系数为 ρ ，沿 x 轴方向和 y 轴方向每间隔 n 个像素点取一个像素点，则 $\rho = 4^n$ ；然后再做均值为 0 方差为 1 的归一化处理；将经下采样和归一化处理的 Gabor 模图像中的像素点按行（或按列）连接起来，得到一维向量 $O_{\mu,v}^{(\rho)}$ ；最后将上述 40 一维向量连接起来得到最终的 Gabor 人脸表示记为：

$$\chi^{(\rho)} = (O_{0,0}^{(\rho)t} O_{0,1}^{(\rho)t} \dots O_{4,7}^{(\rho)t})^t, \quad (3.20)$$

其中 t 为转置运算符。

若取 $\rho = 64$ ，则 $\chi^{(\rho)}$ 的维数为 $10240 (= 40 \times 16 \times 16)$ 。

3.3.2 PCA(Principle Component Analysis)

PCA [94] 是 Principal component analysis 的缩写，中文译名为主元分析。它是一种对数据进行分析的技术，最重要的应用是对原有数据进行简化。正如它的名字：主元分析，这种方法可以有效的找出数据中最“主要”的元素和结构，去除噪音和冗余，将原有的复杂数据降维，揭示隐藏在复杂数据背后的简单结构。它的优点是简单，而且无参数限制，可以方便的应用与各个场合。因此应用极其广泛，被誉为应用线性代数最价值的结果之一。

1991 年，MIT 的 M. Turk 和 A. Pentland 根据 PCA 原理提出的 Eigenface [23, 95] 是人脸识别领域里程碑式的成果，接下来近 20 年的时间中，人脸识别成为模式识别领域最热门的方向之一。

本文采用 PCA 方法对经下采样后的人脸 Gabor 描述 $\chi^{(\rho)}$ 进行降维处理。具体步骤如下：

步骤 1：

选取 M 个人脸 Gabor 描述 $\chi_1, \chi_2, \dots, \chi_M$ 做为训练样本， χ_i 的维数为 $N \times 1$ 。

步骤 2：

计算平均脸向量 Ψ ：

$$\Psi = \frac{1}{M} \sum_{i=1}^M \chi_i \quad (3.21)$$

步骤 3:

每个训练样本 χ_i 减去平均脸 Ψ :

$$\Phi_i = \chi_i - \Psi \quad (3.22)$$

步骤 4:

计算协方差矩阵 C :

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T \quad (3.23)$$

其中 $A = [\Phi_1 \Phi_2 \dots \Phi_M]$

步骤 5:

计算 AA^T 的特征向量 u_i 。一般情况下训练样本的个数远远小于样本特征的维数, 因此矩阵 $AA^T (N \times N)$ 的维数特别高, 从计算复杂度的角度考虑, 直接求解矩阵 AA^T 的特征向量不太可行。我们可以先求解 $A^T A (M \times M)$ 的特征值, 然后再得到 AA^T 的特征值。

求矩阵 $A^T A$ 的特征向量 v_i :

$$A^T A v_i = \lambda_i v_i \quad (3.24)$$

通过以下推导, 可以由 $A^T A$ 的特征向量求得 AA^T 的特征向量。

$$A^T A v_i = \lambda_i v_i \Rightarrow (AA^T)(Av_i) = \lambda_i (Av_i) \quad (3.25)$$

则 AA^T 和 $A^T A$ 具有相同的特征值, AA^T 的特征向量 $u_i = Av_i$ 。标准化 u_i , 使得 $\|u_i\| = 1$ 。

步骤 6:

选取最大的 K 个特征值对应的特征向量 u 构成 PCA 的映射矩阵 $P (M \times K)$, 降维后的人脸 Gabor 表示为:

$$\Gamma = P^T \chi \quad (3.26)$$

此时 Γ 的维数为 $K \times 1$, 最终实现降维。

由于 PCA 是一种无监督的分类机制, 导致了不同类间距离变大的同时, 相同类内距离的距离也变大了, 这非常不利于分类。PCA 这一不足可以通过使用 Fisher 线性判据 (Fisher Linear Discriminant, FLD) 来补偿。

3.3.3 Fisher-LDA(Fisher Linear Discriminant Analysis)

为了达到最好的分类效果, FLD 的中心思想是使类间距离与类内距离比值最大, 最大化下列目标函数:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (3.27)$$

其中 S_B 和 S_W 分别为类间分离矩阵和类内分离矩阵。

令 $\omega_1, \omega_2, \dots, \omega_L$ 表示 L 个类, $N_1, N_2, \dots, N_L$ 表示每个类中样本的个数, $M_1, M_2, \dots, M_L$ 表示每一类的样本均值, M 为所有样本均值。则有

$$M_k = \frac{1}{N_k} \sum_{j=1}^{N_k} \Gamma_j^{(k)}, \quad k = 1, 2, \dots, L \quad (3.28)$$

$$M = \sum_{k=1}^L P(\omega_k) M_k \quad (3.29)$$

其中 $\Gamma_j^{(k)}, j = 1, 2, \dots, N_k$ 表示 ω_k 类中 N_k 个经过 PCA 降维的人脸 Gabor 表示。 $P(\omega_k)$ 为类 ω_k 的先验概率, 一般的 $P(\omega_k) = \frac{N_k}{N}$, N 为样本总数。

类间分离矩阵 S_B 和类内分离矩阵 S_W 可表示为:

$$S_W = \sum_{k=1}^L P(\omega_k) \left\{ \frac{1}{N_k} \sum_{j=1}^{N_k} (\Gamma_j^{(k)} - M_k)(\Gamma_j^{(k)} - M_k)^t \right\} \quad (3.30)$$

$$S_B = \sum_{k=1}^L P(\omega_k) (M_k - M)(M_k - M)^t \quad (3.31)$$

理想情况下, 使式 (3.27) 取最大, 需 $w^T S_W w = 0$, 但由于 $w^T S_W w$ 是分母不能为零。我们把式 (3.27) 近似的转化为下列优化问题:

$$\begin{aligned} \min_w \quad & -\frac{1}{2} w^T S_B w \\ \text{s.t.} \quad & w^T S_W w = I \end{aligned} \quad (3.32)$$

上式的 Lagrange 方程可写成:

$$L_P = -\frac{1}{2} w^T S_B w + \frac{1}{2} \lambda (w^T S_W w - I), \quad (3.33)$$

令 $\frac{\partial L_P}{\partial w} = 0$ 得式 (3.34),

$$S_B w = \lambda S_W w \Rightarrow S_W^{-1} S_B w = \lambda w. \quad (3.34)$$

对矩阵 S_B 进行正交分解有, $S_B = U \Lambda U^T$, 则 $S_B^{\frac{1}{2}} = U \Lambda^{\frac{1}{2}} U^T$ 。令 $v = S_B^{\frac{1}{2}} w$, 则式 (3.34) 可化为:

$$S_B^{\frac{1}{2}} S_W^{-1} S_B^{\frac{1}{2}} v = \lambda v \quad (3.35)$$

通过 SVD 分解求得矩阵 $S_B^{\frac{1}{2}} S_W^{-1} S_B^{\frac{1}{2}}$ 的特征值与特征向量为 λ_k 和 v_k , 则 $w_k = S_B^{-\frac{1}{2}} v_k$ 。令 $w = [w_1 \ w_2 \ \dots \ w_n]$, 最终的 FLD 变换见公式 (3.37),

$$Y_i = w^T \Gamma_i \quad (3.36)$$

最终经 PCA+LDA 变换之后的人脸 Gabor 描述为:

$$Y = w^T \Gamma = w^T P^T \chi \quad (3.37)$$

设 $F_k^0, k = 1, 2, \dots, L$ 为类 w_k 经 PCA+LDA 变换后的训练样本均值。使用最近邻分类器 (Nearest Neighbor Classifier) 做人脸分类:

$$\delta(Y, F_k^0) = \min_j \delta(Y, F_j^0) \longrightarrow Y \in w_k \quad (3.38)$$

其中 δ 为距离判据, 常用的距离判据有:

$$\delta_{L_1}(X, Y) = \sum_i |X_i - Y_i| \quad (3.39)$$

$$\delta_{L_2}(X, Y) = (X - Y)^t (X - Y) \quad (3.40)$$

$$\delta_{Md}(X, Y) = (X - Y)^t \Sigma^{-1} (X - Y) \quad (3.41)$$

$$\delta_{\cos}(X, Y) = \frac{-X^t Y}{\|X\| \|Y\|} \quad (3.42)$$

式 (3.39 - 3.42) 分别为 L_1 距离, L_2 距离, 马氏距离和 cos 距离。

3.3.4 实验结果与分析

为了测试算法的有效性, 用 C++ 编程实现了本节所述算法, 并在 FERET 人脸库和 RCIR 人脸库上进行实验。

(1) FERET 上的测试

实验准备

在 FERET 人脸库中选取 300 个人, 每人 5 张样本, 其中任意选取 4 张作为训练样本, 1 张作为测试样本。把所有样本经 1.4 节所述的人脸标准化后的图片作为实验用样本。部分样本图片如图 3.7 所示。

实验结果

PCA 降维时采用不同的维数, 得到不同的实验结果, 见表 3.1。当主元数取 180 时, 识别率最高为 96.60%。

表 3.1 FERET 人脸库实验结果

Table 3.1 Experimental result on FERET

实验	PC 数	识别率(%)
1	10	80.23
2	30	84.80
3	50	89.73
4	80	91.63
5	100	93.50
6	120	93.50
7	150	95.14
8	180	96.60

(2) RCIR 人脸库上的测试



图 3.7 FERET 部分实验样本

Fig 3.7 Some experimental Samples from FERET

实验准备

为了更好的测试人脸识别算法，我们在实验室环境下采集了 184 个人的人脸图像，这些人脸图像的采集时间不同，采集时环境光照不同，采集所使用的摄像头不同，人脸的配饰不同 (如佩戴不同的眼镜)，人脸表情不同，人脸的姿态也略有不同。这些变化，保证了测试样本的多样性，更能反映人脸识别算法在实际环境下的识别效果。

对于上述 184 个人，每人 5 张样本，其中任意选取 4 张作为训练样本，1 张作为测试样本。把所有样本经 1.4 节所述的人脸标准化后的图片作为实验用样本。部分样本图片如图 3.8 所示。

实验结果

同样的，PCA 降维时采用不同的维数，得到不同的实验结果，见表 3.2。当主元数取 180 时，识别率最高为 93.76%。

3.4 基于特征选择的 Gabor 人脸识别方法

本章 3.3 小节中，所述的 Gabor+PCA+FLD 人脸识别方法在 FERET 标准人脸库和 RCIR 人脸库上均达到令人满意的识别效果。当我们把该方法在 Intel P4 1.8G CPU，512M 内存配置的 PC 机上进行实时的人脸识别时，发现识别一幅 128×128 大小的标准人脸图像时，耗时约 $2400ms$ 。这样的时间消耗，无法完全满足实时人脸识别的要求。本小节将分析耗时过大的原因，并提出基于特征选择的 Gabor 人脸识别方法来解决上述问题。



图 3.8 RCIR 人脸库部分样本

Fig 3.8 Some experimental Samples from RCIR Face Database

表 3.2 RCIR 人脸库实验结果

Table 3.2 Experimental result on RCIR face database

实验	PC 数	识别率(%)
1	10	75.50
2	30	80.12
3	50	88.65
4	80	90.32
5	100	91.04
6	120	91.27
7	150	93.14
8	180	93.76

3.4.1 复杂度分析

本章 3.2 节中, Gabor 的人脸表示是通过原始人脸图像与 Gabor 核的实部与虚部分别卷积, 最后求模得到的。注意到一维卷积的算法复杂度为 $\mathcal{O}(n^2)$, 通过 FFT 和 IFFT 计算一维卷积的算法复杂度为 $\mathcal{O}(n \log n)$, 则二维图像间卷积的算法复杂度为 $\mathcal{O}((n \log n)^2)$ 。因此 Gabor 特征提取时所做的 80 次二维卷积是耗时的主要原因, 根据我们的测算, 对 128×128 大小的图像每进行一次二维卷积大约需要 $30ms$ 。

Gabor 特征描述的特点在于通过原始人脸图像与不同 scale 和不同 orientation 的 Gabor 核卷积, 得到多分辨率下的人脸描述。然而, 40 个 Gabor 核对人脸特征的描述能力是否一致呢? 是否存在某些 Gabor 核更加能够反映人脸的特征呢?

此外, 本章 3.3.1 小节通过均匀下采样的方法降低过高的 Gabor 人脸特征描述, 均

匀采样的方法即默认人脸不同位置对于人脸识别的作用是一样的,而我们人类往往是通过人脸特殊部位的某些显著特征(如大眼睛,高鼻梁,宽额头等)来识别一个人。因此,我们尝试使用机器学习和数据挖掘中某些特征选择的方法[96,97]来选择最具有区分性的 Gabor 特征。

3.4.2 特征选择

特征选择问题可定义为:给定包含 d 个特征的特征集合,选择出大小为 m 的子集,可达到最小的分类误差。基于特征选择的分类方法的一般模型见图 3.9。

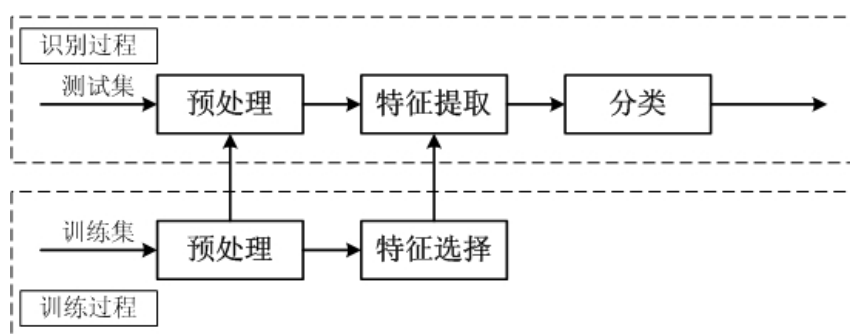


图 3.9 基于特征选择的分类方法的一般模型

Fig 3.9 Model of classifier based on feature selection

由于人脸的 Gabor 特征描述,既包含了位置信息,又包含了 Gabor 核的参数信息(如尺度 v 和方向 μ)。因此我们采用两步特征子集的选择机制:第一步,通过 SFFS [98] 方法找出最具有区分性的人脸区域;第二步,通过 SFFS 方法学习得到描述能力最强的 Gabor 核参数。当然上述两个特征选择步骤的先后顺序不是特别重要,但首先针对整个人脸区域选择 Gabor 核的参数,不是非常合理,因为不同的人脸区域对应不同的 Gabor 核参数。

(1) Gabor 核位置选择

在根据 Gabor 核的位置构成待选特征集合时,我们需要进行采样,一般有如下几种采样模式:(a) 稠密采样,即每个像素点都进行卷积运算,这种方法极大的提高了待选特征集合的维数,用 SFFS 方法选择特征不太可行;(b) 基于人脸特征点采样(见图 3.10(a)),即对人脸的显著特征区域(如眉毛,眼睛,鼻子,嘴巴等)进行卷积,从而构成待选特征集合,这种方法需要手工或自动标定人脸特征点的辅助工作;(c) 基于网格的采样(见图 3.10(b)),在人脸区域覆以 $N \times N$ 的矩形网格,在网格的节点出作卷积,构成待选特征集合,这种方法,简单易行,同时保证了待选特征集合的维数适中,本文采用这种方法进行采样。

网格的每一个节点,分别与 $M = v \times \mu$ 个 Gabor 核作卷积, M 个卷积结果的实部与虚部分别求模,然后连接起来,构成一维特征向量,该特征向量作为待选特征空间的一维。故待选特征空间共 $N \times N$ 维。我们设定特征子集维数为 d ,并以此作为 SFFS 方

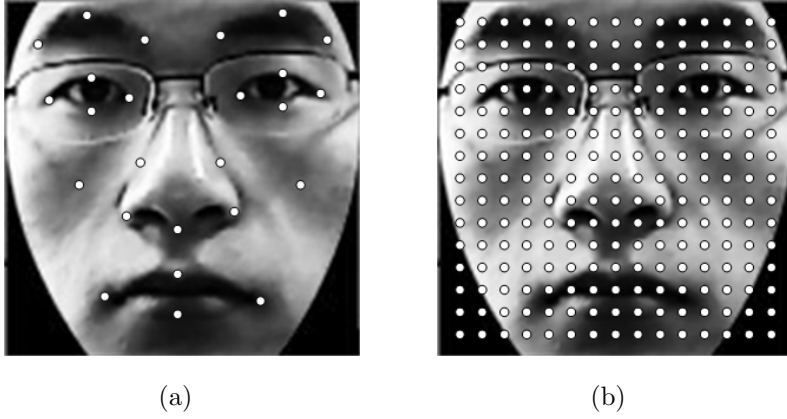


图 3.10 (a) 基于特征点的采样, (b) 基于网格的采样

Fig 3.10 (a) Landmark-based sampling (b) Lattice-based sampling

法的终止条件。

(2) Gabor 核参数选择

当选定 Gabor 核的位置之后, 我们选择在这些位置处, 最具有分辨能力的 Gabor 核的尺度与方向。特征选择的第一步得到了一个维数为 $d \times M$ 的特征子集 $\mathcal{X}_{loc}$ 。在第二步 Gabor 核尺度与方向参数的选择中, 我们将得到集合 $\mathcal{X}_{loc}$ 的一个子集 $\mathcal{X}_{v,\mu}$, 且 $|\mathcal{X}_{v,\mu}| \ll |\mathcal{X}_{loc}|$ 。注意到 $\mathcal{X}_{v,\mu}$ 特征子集中每一维与一个 (v, μ) 参数组对应。此阶段 SFFS 特征选择的终止准则为验证集的分类准确性。

3.4.3 实验与结果分析

(1) 特征选择过程

我们使用 FERET [83] 人脸库的一个子集, 该子集中含有 150 个人, 每人有 3 张标准化成 128×128 的人脸图像, 其中 2 张构成训练集用于选择特征, 另外 1 张用于判定终止的验证集。待选位置为 7×7 的网格节点 (见图 3.11(a)), Gabor 核的大小为 16×16 , 包含 5 个尺度和 8 个方向。在每个网格节点 i , 我们通过 Gabor 核的卷积运算, 得到局部的特征向量 $\mathcal{V}_i = \{m_{0,0}, m_{0,1}, m_{0,2}, \dots, m_{v,\mu}\}, v = 0, \dots, 4; \mu = 0, \dots, 7$, 其中 $m_{i,j}$ 表示 Gabor 卷积的模。 $\mathcal{V}_i$ 维数 $|\mathcal{V}_i| = 40$ 。把所有局部特征向量连接起来, 得到一个全局的特征向量 $\Phi = \{\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k\}, k = 49$, 则 Φ 的维数 $|\Phi| = 49 \times 40$ 。设 Φ_{LOC} 为选择的维数为 d 的位置特征子集, $\Phi_{LOC} = \{\mathcal{V}_i : i \in 1, \dots, k\}$, 本实验中, k 取为 15。

经 SFFS 方法选择出来的 15 个位置, 如图 3.11(b) 所示。通过对结果的分析可以发现, 选择出来的特征点大部分都在人脸上部, 特别是眼眉和鼻子区域, 脸颊部分的区分度不大, 因此位于脸颊部的特征点较少, 同时由于人脸的表情变化对嘴周围区域有较大影响, 该部分的特征不是很稳定, 所以位于嘴附近区域的特征点也较少。

得到最具区分性的位置特征子集 Φ_{LOC} 后, 我们继续在 Φ_{LOC} 选择描述能力较强的 (v, μ) 构成最终的特征子集 $\Phi_{v,\mu}$ 。注意到 Φ_{LOC} 的维数 $|\Phi_{LOC}| = 15 \times 40 = 600$ 。 $\Phi_{v,\mu}$ 通

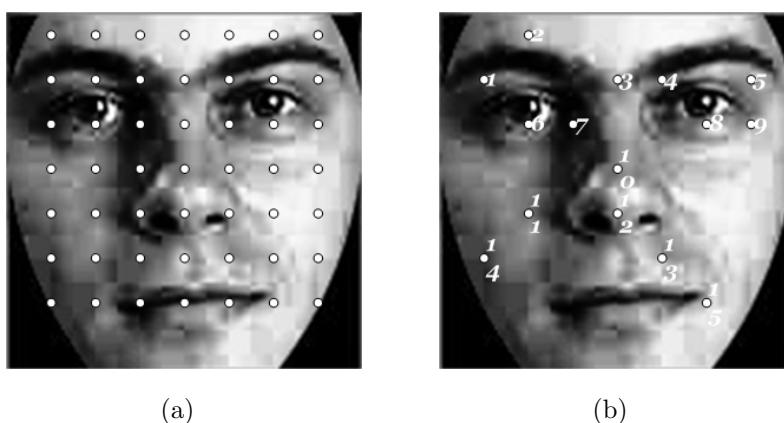


图 3.11 (a) 网格节点的位置, (b) 选出的位置

Fig 3.11 (a) Position of lattice node (b) Position after selection

过 SFFS 方法选择得到, 终止条件通过观察验证集的认识率来确定, 当验证集的认识效果在一定的选择轮次内无明显改善时, 选择过程终止。 $\Phi_{v,\mu}$ 最终表示以表格 3.3 给出。

表 3.3 选择出的尺度和方向特征

Table 3.3 Feature of scale and orientation after selection

位置	(v, μ)
1#	(0, 0) (1, 0) (1, 1) (2, 1) (1, 3) (2, 3) (0, 4) (1, 4) (2, 4) (2, 5) (0, 6) (1, 7) (2, 7)
2#	(2, 0) (1, 1) (2, 1) (0, 3) (1, 4) (3, 4) (4, 4) (0, 5) (3, 5) (4, 5) (1, 6) (2, 7)
3#	(0, 0) (3, 0) (4, 0) (3, 1) (4, 1) (2, 2) (4, 3) (3, 4) (4, 4) (4, 5) (2, 7) (3, 7)
4#	(2, 1) (3, 1) (4, 1) (3, 2) (2, 3) (1, 4) (2, 4) (3, 4) (1, 5) (2, 5) (2, 7)
5#	(2, 1) (4, 1) (0, 3) (1, 3) (4, 3) (0, 4) (1, 4) (2, 4) (3, 4) (4, 4) (1, 5)
6#	(2, 0) (3, 0) (4, 0) (0, 1) (2, 1) (3, 1) (4, 1) (4, 3) (2, 4) (2, 4) (0, 5) (1, 5) (3, 5)
7#	(2, 0) (3, 0) (1, 2) (3, 2) (0, 3) (1, 3) (2, 4) (0, 5) (1, 5) (4, 5) (1, 6) (0, 7) (2, 7)
8#	(0, 0) (0, 1) (3, 0) (4, 0) (2, 1) (3, 2) (0, 4) (3, 4) (4, 4) (2, 5) (3, 5) (2, 7)
9#	(0, 0) (2, 0) (4, 0) (4, 2) (2, 3) (3, 3) (0, 4) (2, 4) (3, 4) (4, 4) (4, 6)
10#	(0, 0) (3, 0) (1, 3) (2, 3) (3, 3) (1, 4) (2, 4) (3, 4) (2, 5) (3, 5) (4, 5) (0, 7) (1, 7)
11#	(2, 0) (3, 0) (1, 1) (2, 1) (3, 1) (0, 4) (1, 4) (1, 5) (1, 6) (0, 7)
12#	(0, 0) (1, 1) (2, 3) (1, 4) (0, 5) (2, 5) (3, 5) (4, 6) (0, 7)
13#	(1, 1) (1, 4) (3, 5) (3, 6)
14#	(1, 2) (4, 3) (3, 4) (3, 5)
15#	(1, 1) (3, 1) (1, 2) (0, 3) (2, 3) (1, 5) (2, 5) (4, 6)
总计	155

通过分析上表中通过 SFFS 方法选择出的尺度和方向特征, 可得到如下结论: (a) 不同尺度的 Gabor 特征对于识别的贡献有较大的不同, 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。这一点与直观的感觉相吻

合。(b) 不同方向的 Gabor 特征也具有不同的贡献, 竖直方向 ($\mu = 4$) 的 Gabor 特征描述能力最强, 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。总体而言, 45 度方向 ($\mu = 2, 6$) 的 Gabor 特征的分辨能力最弱。

经上述两步的特征选择过程, 我们最终得到 155 个不同位置、尺度和方向的 Gabor 特征描述, 把这些向量连接起来所构成的向量, 作为人脸特征向量用于识别。分类时采用最近邻分类器 (NN) 即可。

此时再分析运算复杂度, 我们提出的基于特征选择的 Gabor 人脸识别方法在特征提取时, 只需要进行 155 次 16×16 的二维卷积, 明显低于 40 次 128×128 的二维卷积的运算复杂度, 识别效率的提高在下文给出的 FERET 和 RCIR 人脸库上的测试结果得到明显体现。

(2) FERET 上的识别测试

为了验证方法的有效性, 我们在 3.3.4(1) 小节使用的含有 300 个人的 FERET 人脸库上进行测试, 每人 5 张样本中, 任意选取 3 张作为 Gallery 样本, 剩余的 2 张作为 Probe 样本, 实验结果见表 3.4。

表 3.4 在 FERET 上与 FLD-Gabor 识别率及平均耗时的比较

Table 3.4 Comparison with LDA-Gabor on FERET

	识别率(%)	平均耗时(ms)
FLD-Gabor	96.60	2400
SFFS-Gabor	95.43	50

实验表明, SFFS-Gabor 方法在保证较高识别率的前提下, 把平均耗时从 2400ms 缩短至 50ms, 极大地提高了识别效率。

(3) RCIR 上的识别测试

由于特征选择的训练集为 FERET 人脸库的子集, 为了验证 SFFS-Gabor 具有一定的泛化性, 我们在 3.3.4(2) 小节所述的含有 184 个人的 RCIR 人脸库上进行测试, 每人 5 张样本中, 任意选取 3 张作为 Gallery 样本, 剩余的 2 张作为 Probe 样本, 实验结果见表 3.5。

表 3.5 在 RCIR 上与 FLD-Gabor 识别率及平均耗时的比较

Table 3.5 Comparison with LDA-Gabor on RCIR

	识别率(%)	平均耗时(ms)
FLD-Gabor	93.76	2400
SFFS-Gabor	87.28	50

虽然 SFFS-Gabor 方法的识别率有所下降, 但仍表现了一定的泛化能力, 若以 RCIR 人脸库作为特征选择的训练集, 识别率会得到相应提高。

3.4.4 方法小结

SFFS-Gabor 方法通过特征选择, 选择出最具有分辨能力的 Gabor 特征描述, 极大降低了 Gabor 特征提取的运算复杂度, 因而缩短整个识别过程的平均耗时, 使得基于 Gabor 的人脸识别方法更适用于实时的人脸识别系统。由于 SFFS 选择出来的特征子集为次优解, 如何通过其他特征选择的方法, 选择出更加优化的特征子集, 以得到更高的人脸识别率, 仍需要继续研究。

3.5 本章小结

本章工作基于 Gabor 小波特征对人脸图像的有效表达, 以 Gabor 小波作为特征提取方法展开。首先介绍了 Gabor 小波, 给出了 Gabor 的人脸描述方法, 接着描述基于 FLD-Gabor 的人脸识别方法, 然后分析该方法存在的不足, 最后通过特征选择, 提出一种新的 SFFS-Gabor 小波的人脸识别方法, 实验结果表明, 该方法保证较高人脸识别率的同时, 缩短了人脸识别过程消耗的时间, 使得基于 Gabor 特征描述的人脸识别方法, 更加适用于实用人脸识别系统。

第四章 基于 LBP 的人脸识别

人类视觉系统主要根据图像的特征来识别图像。所谓图像特征是指图像的原始特性或属性，如区域的亮度、边缘的轮廓、纹理或色彩等等。其中，纹理特征反映了物体表面颜色和灰度的某种变化，而这些变化又与物体本身的属性相关。比如同一个树种的木材有相同的或相近似的纹理，人们通过识别木纹的种类来识别木材的树种和材质。纹理在图像处理中起着非常重要的作用，现已广泛应用于气象云图分析，卫星遥感图像分析、生物组织和细胞的显微镜照片分析等领域。

构成纹理特征的有两个要素，一是纹理基元，另一个是纹理基元的排列组合。纹理基元是一种或多种图像基元的组合。纹理基元有一定的形状和大小。纹理基元的排列组合就是反映基元排列的疏密、同期性和方向性的不同，这些不同使图像的外观产生极大的改变。

纹理特征提取指的是通过一定的图像处理技术抽取出纹理特征，从而获得纹理的定量或定性描述的处理过程。因此，纹理特征提取应包括两方面的内容：检测纹理基元和获得有关纹理基元排列分布方式的信息。

局部二值模式 (Local Binary Pattern, LBP) 是一种比较有效的纹理描述子，在纹理分析中得到广泛的引用。继 Gabor 描述之后，基于 LBP 的人脸识别近年来成为人脸识别领域的研究热点之一 [99, 100]。

本章首先介绍 LBP 算子，接着描述基于 LBP 的人脸识别方法，然后分析该方法存在的不足，最后提出一种新的基于全变分模型的 LBP 人脸识别方法。

4.1 LBP 算子

1996 年 T. Ojala 等人引入了局域二值模式 (Local Binary Pattern, LBP) 作为纹理描述子来分析图像纹理特征 [101]。它利用结构法思想分析固定窗口特征，然后通过统计分析作整体的特征提取。LBP 描述子一般定义为 3×3 的窗口中，以窗口中心点的灰度值为阈值对窗口内其 8 邻域的像素作二值化处理，然后根据像素不同位置进行加权求和得到该窗口的 LBP 值。详细描述见图 4.1 所示。

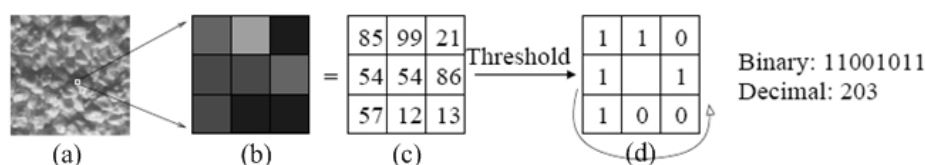


图 4.1 LBP 描述子

Fig 4.1 LBP descriptor

从图 4.1(a) 的图像中提取一个 3×3 的窗口 (图 4.1(b))，其灰度值如图 4.1(c) 所示。中心像素点的灰度值为 54，以其为阈值周围 8 个像素点灰度值若大于 54 则置为 1，反之置为 0 (图 4.1(d))。然后可以得到一个 8 位的二进制序列， $Binary = 11001011$ ，其十进制表示 $Decimal = 203$ 做为中心像素点的最终像素值。

原始的 LBP 描述子由于仅仅限于 3×3 的窗口，所以描述能力受到了一定限制。Ojala et al. 对 LBP 描述子的定义进行了扩展。

定义：T 表示由相邻 $P+1$ 个像素点灰度值之间的关系。

$$T = t(g_c, g_0, \dots, g_{p-1}) \quad (4.1)$$

其中， g_c 表示中间像素点的灰度值， $g_p (p = 0, 1, \dots, p-1)$ 表示围绕 g_c ，以 R 为半径的像素点的灰度值。 R 和 P 取不同的值，表示不同的 LBP 描述子。

假设不计信息损失， g_c 能被 g_p 减，上式可写为：

$$T = t(g_c, g_0 - g_c, \dots, g_{p-1} - g_c) \quad (4.2)$$

考虑到该数值对 g_c 的依赖性不大，上式可近似的写为：

$$T \approx t(g_c) t(g_0 - g_c, \dots, g_{p-1} - g_c) \quad (4.3)$$

由于 $t(g_c)$ 是描述图像的全部灰度值，它与局部图像的纹理无关，所以可以忽略不计。

$$T \approx t(g_0 - g_c, \dots, g_{p-1} - g_c) \quad (4.4)$$

尽管 $g_p - g_c (p = 1, 2, \dots, p-1)$ 之间在数值上存在差异，但它们之间大小关系的比较，足以描述局部区域纹理的变化，所以式 (4.4) 可继续化简为：

$$T \approx t(s(g_0 - g_c), \dots, s(g_{p-1} - g_c)) \quad (4.5)$$

其中 $s(x)$ 定义为：

$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4.6)$$

给每个 $s(g_p - g_c)$ 值赋 2^p 的权重，这样就得到了一个与邻域像素点相关的唯一 LBP 代码，见公式 (4.6)。

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (4.7)$$

4.2 LBP 在人脸识别中的应用

4.2.1 LBP 人脸描述

为了得到人脸的 LBP 描述，需要进行如下步骤：

(1) LBP变换

人脸原始图像经过公式 (4.6) 的变换, 可以得到人脸的 LBP 图像。图 4.2 为 RCIR 人脸库中部分人脸样本图像的原始图像和 $LBP_{1,8}$ 图像。

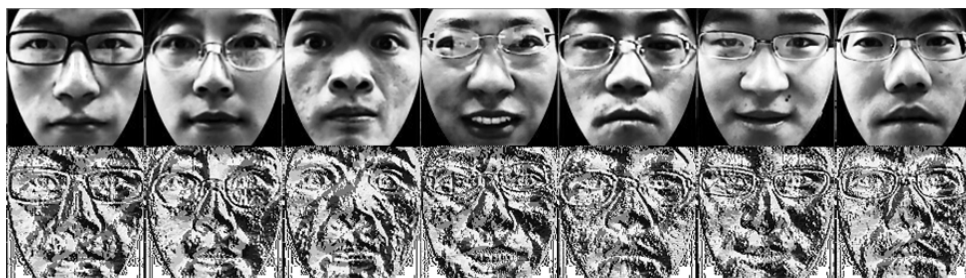


图 4.2 LBP 表示的人脸图像

Fig 4.2 Face image under LBP

(2) 分块

局部二值化模式中的“局部”包括两层含义: 1、每个像素点和它周围像素点之间的“局部”关系; 2、图像区域与整幅图像的“局部”关系。因此, 需要对步骤 (1) 中的得到的 LBP 人脸图像分成 N 个子区域。图 4.3 中, 分别显示了 4×4 、 8×8 、 16×16 三种平均的非重叠的分块结果。

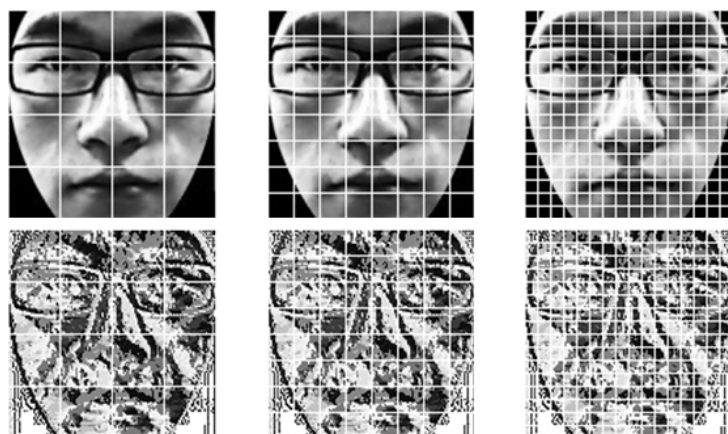


图 4.3 人脸图像的分块

Fig 4.3 Sub-region of face image

(3) 直方图统计

分别在由步骤 (2) 得到每一个 LBP 子区域内进行直方图分析 (见图 4.4), 把得到的每个直方图序列连接起来, 最终得到的直方图序列作为 LBP 人脸描述。

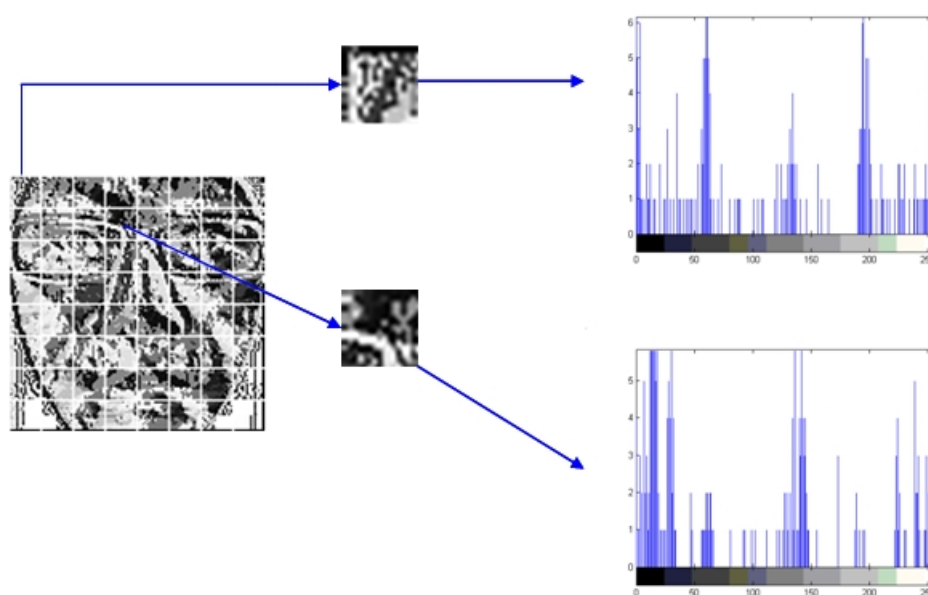


图 4.4 LBP 直方图序列

Fig 4.4 Histogram sequence of LBP

4.2.2 双线性插值的 LBP 人脸描述

经分析比较可以发现, 图 4.2 中的 *LBP* 人脸图像中, 皮肤部分呈现灰度值不均匀的现象, 这与人脸皮肤相对平滑的事实相悖, 因此若这样的信息用于人脸识别, 将会对识别效果产生很大影响。

图 4.5 表示了 (R, P) 为 $(1, 8)$, $(2, 16)$, $(2, 8)$ 的三种邻域的分布情况, 从图中可以看出, 很多邻域点并没有出现在像素点的中心, 若取像素点的灰度值作为邻域点的灰度值, 就会出现相对误差, 从而导致上面所描述的问题。为了更加准确的描述邻域点的灰度值, 我们采用双线性插值 (Bilinear Interpolation) 的方法, 通过邻域点周围的四个像素点的灰度值求出该邻域点的灰度值。

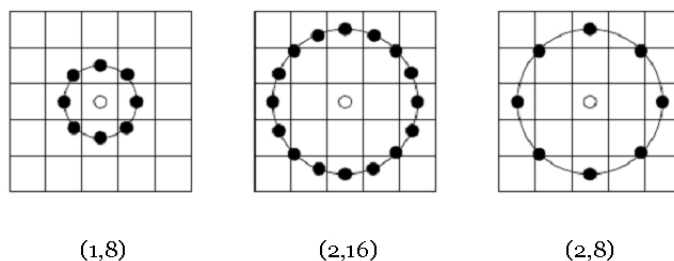


图 4.5 邻域分布

Fig 4.5 The distribution of neighborhood

图 4.6 表示了双线性插值的具体过程, 首先根据公式 (4.8) 和 (4.9) 求出插值的比例系数, 然后根据公式 (4.10) 和 (4.11) 线性插值出 E 和 F 点的值, 最后由公式 (4.12) 根据

E 和 F 点的值线性插值出最终结果，从而完成双线性插值过程。

$$\alpha = u_0 - u, \quad (4.8)$$

$$\beta = v_0 - v, \quad (4.9)$$

$$f(u_0, v) = f(u, v) + \alpha[f(u+1, v) - f(u, v)], \quad (4.10)$$

$$f(u_0, v+1) = f(u, v+1) + \alpha[f(u+1, v+1) - f(u, v+1)], \quad (4.11)$$

$$f(u_0, v_0) = f(u_0, v) + \beta[f(u_0, v+1) - f(u_0, v)]. \quad (4.12)$$

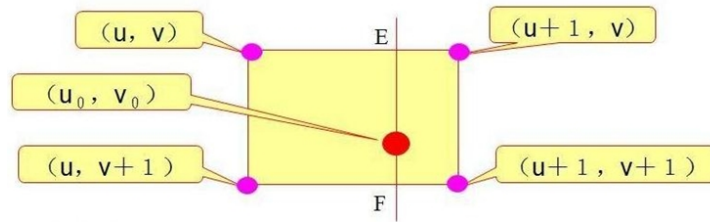


图 4.6 双线性插值

Fig 4.6 Bilinear interpolation

图 4.7 中第一行是未作双线性插值的 LBP 人脸图像，第二至第五行分别是 $(R, P) = (1, 8), (2, 8), (2, 16), (3, 8)$ 经过双线性插值后的 LBP 人脸表示。比较发现，插值后不但边缘轮廓和角点等特征更加显著，皮肤等非显著特征变得平滑。

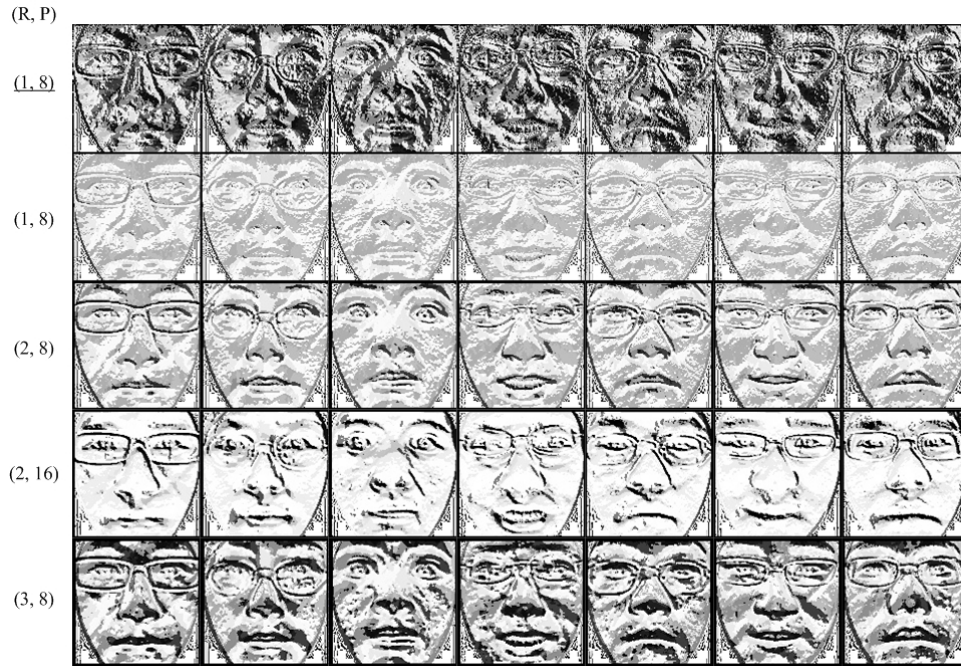


图 4.7 双线性差值的 LBP 人脸图像

Fig 4.7 Bilinear interpolation LBP face image

人脸的 LBP 描述最终可写为:

$$\Psi_{LBP} = (HS_1 HS_2 \dots HS_N) \quad (4.13)$$

其中 $HS_1, HS_2, \dots, HS_N$ 为 N 个子图的直方图序列。

人脸识别使用 NN(Nearest-Neighbor) 分类器, 常用的距离判据有, 直方图相交:

$$D(S, M) = \sum_i \min(S_i, M_i), \quad (4.14)$$

Log 概率统计:

$$L(S, M) = - \sum_i S_i \log M_i, \quad (4.15)$$

χ^2 概率统计:

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i}. \quad (4.16)$$

4.2.3 实验结果与分析

为了测试本文所提方法的性能, 我们在 FERET 人脸库进行实验。它提供了指定的原型集 (gallery) 和测试集 (probe), 原型集是由 1196 人的每人一幅正面图像构成。4 个测试集分别是: fb, fc, Dup.I 和 Dup.II。其中, fb 含有 1195 幅与原型集同时采集, 并且与原型集图像光照相同的表情变化图像; fc 中含有 194 幅与原型集图像采集光照条件不同的人脸图像; 在 Dup.I 和 Dup.II 中, 分别有 722 幅和 234 幅图像, 采集时间距原型集分别为一个月和一年左右。人脸图像样本示例如图 4.8 所示。

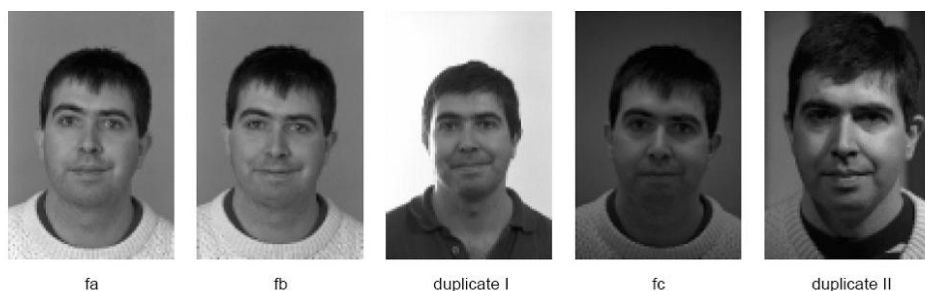


图 4.8 FERET 人脸库例图

Fig 4.8 Some sample images from FERET

把上述所有人脸头肩部图像经 1.4 节所述的人脸标准化后的图像作为实验用样本。其中图像大小为 128×128 , 分为 16×16 个子图, 则每个子图的大小为 8×8 。

实验结果见表 4.2:

实验结果表明, 上述方法在测试集 fb 上可以达到 91.3% 的识别率, 可以满足一般应用的需要。

表 4.1 双线性插值 LBP 人脸识别实验结果

Table 4.1 Experimental result of BI-LBP face recognition

	fb(%)	fc(%)	Dup.I(%)	Dup.II(%)
识别率	91.3	50.0	50.0	31.1

我们再来分析一下特征维数。特征维数可以通过下列公式求取，

$$Dimension = SubRegNum \times HistBin, \quad (4.17)$$

其中 $SubRegNum$ 和 $HistBin$ 分别为子区域的个数和直方图的维数。就上文所述实验而言， $SubRegNum = 16 \times 16 = 256$, $HistBin = 256$ ，所以特征维数为 65536。如此的高维特征，对于一个实用人脸识别系统而言非常不现实。主要存在以下两点问题：

1. 高维特征需要耗费大量的存储空间；
2. 高维特征的存储与读取以及比对需要耗费大量时间。

上文所述实验的实验平台为 PC 机器，其主要配置为 CPU:Inter Pentium D 2.8G，内存 DDR2 800 2G，数据存储 Microsoft SQL Server 2000 数据库中，经过实验测算得知，从数据库读取一维数据的时间约为 $0.67\mu s$ ，比对一维特征的时间约为 $0.034\mu s$ 。对于处理 65536 维特征总共耗时约为 $65536 \times (0.67 + 0.034) = 46ms$ 。

如果 Gallery 中有 100 个人，每人 3 张训练样本，这样识别一个人需要耗时 $46ms \times 300 = 13.8s$ ，根本无法满足实时应用。

通常意义上，灰度直方图的 $bin = 256$ 。由于子区域的面积比较小，若子区域的大小为 8×8 ，一个区域的像素点只有 64，这样分布在 bin 为 256 的直方图会非常稀松，将出现很多零值，如图 4.9 所示。这些不具有区分性的零值完全可以视为冗余信息，我们使用压缩直方图的方法来消除这些冗余的零值，实现 LBP 特征描述的维数。

我们分别取 $HistBin = 8, 16, 32, 64$ ，在 FERET 测试集 fb 上进行测试。实验结果见表 4.2。

当 $HistBin$ 取 64，子图大小取 8×8 时，fb 上的识别率为 91.23%，但特征维数和识别时间只有原来的 $1/4$ 。从实验结果，我们可以看出压缩直方图的方法有效降低特征维数，减少识别耗时的同时保证了与原始 LBP 相当的识别率。同时我们也得到这样的结论：在特征维数相等的情况下，子图相对较小时，识别率相对较高。说明稠密的分块比高的直方图维数对于人脸特征更具有描述性。

4.3 基于全变分模型的 LBP 人脸识别

LBP 方法本质上提取的是图像边缘、角点等局部变化特征，它们是区分不同人脸的重要特征。当外界光照发生变化时，人脸图像上会出现许多阴影和模糊的边缘，特别受到侧光照射时，人脸图像上甚至会出现明暗分明的区域，这些变化严重干扰了 LBP 方

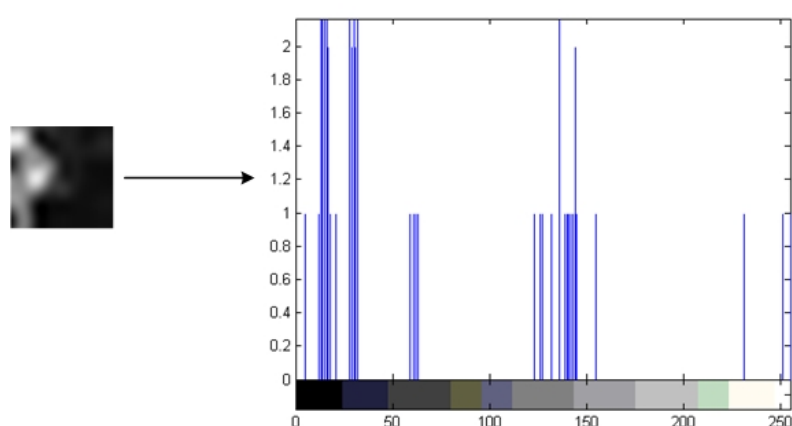


图 4.9 子图的直方图

Fig 4.9 Histogram of a sub-region image

表 4.2 压缩直方图 LBP 在 fb 上的识别实验结果

Table 4.2 Experimental result of compressed-histogram LBP on fb

直方图维数	子图大小	特征维数	耗时(ms)	fb 识别率(%)
256	8×8	65536	46	91.30
8	8×8	2048	1.45	84.47
16	8×8	4096	2.89	87.52
32	8×8	8192	5.78	89.74
64	8×8	16384	11.56	91.23
8	16×16	512	0.34	78.10
16	16×16	1024	0.71	80.23
32	16×16	2048	1.45	83.27
64	16×16	4096	2.89	86.36

法对图像边缘和角点特征的提取，从而导致人脸识别率的下降。4.2 节给出 FERET 人脸库 fc 测试集识别率仅为 50.0% 的实验结果反映了上述问题。很多研究者提出了改进 LBP 人脸识别的方法 [102,103]，本节将提出一种新的基于全变分模型的 LBP 人脸识别方法，结合了 LBP 和 TVM 在人脸特征描述上的优点，以 TVM 为人脸图像预处理的手段，保证了 LBP 方法可以提取稳定的边缘和角点特征，在光照发生变化的情况下，仍可得到较高的人脸识别率。

4.3.1 全变分模型

全变分模型 (Total Variation Model, TVM) 是一种有效光照补偿的方法，该模型可以修正因光照影响而下降的图像质量，恢复原本模糊不清的面部轮廓。区别于其他传统的轮廓提取、轮廓增强方法 [81]，该模型有效保留了利于识别的纹理信息。图 4.10(a) 为

FERET 库中同一个人的几张不同人脸样本图像，其采集的时间及采集的光照情况都有所不同。图 4.10(b) 为经 TVM 模型处理后的样本图像，通过比较可以看到，TVM 保留并强化了面部边缘和角点等关键特征，淡化了受光照影响较大且对识别作用不大的均匀纹理部位（如面颊，额头等）。

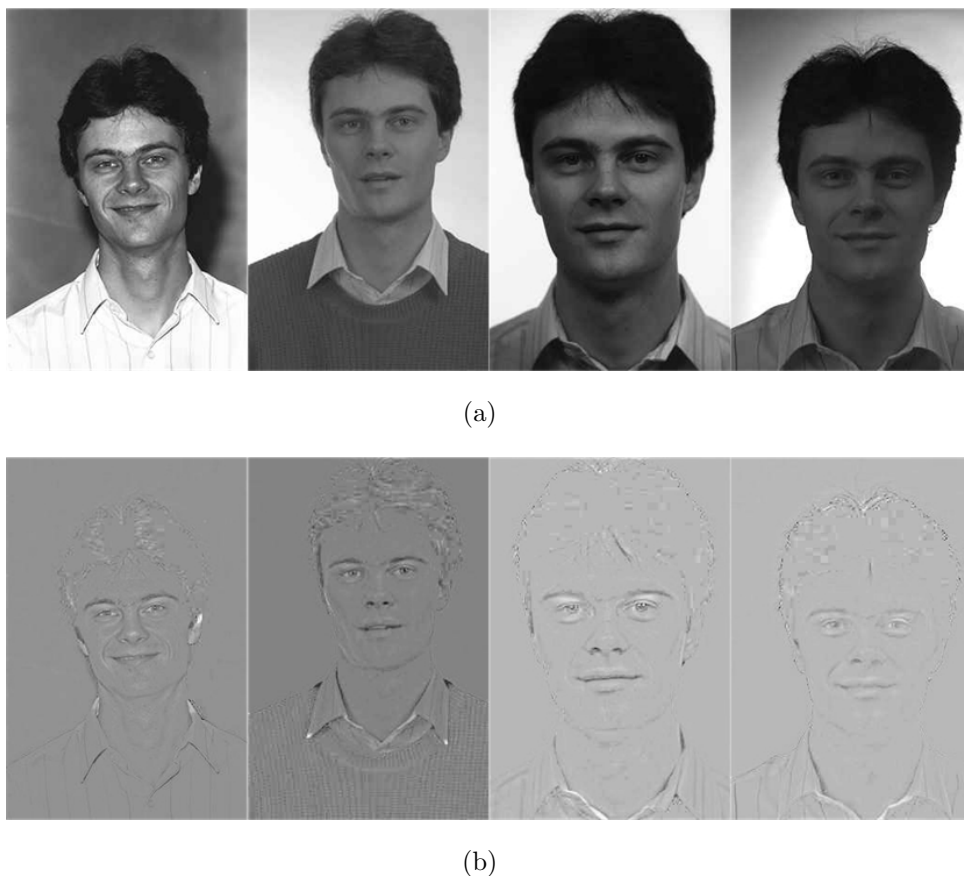


图 4.10 (a)FERET 中部分样本, (b)TVM 变换的结果

Fig 4.10 (a) Sample images from FERET (b) Images after TVM transform

简单说来，全变分模型的基本思想是把图像的纹理分为两部分：肤色背景纹理，及精细的轮廓纹理。当外界光照情况发生变化时，前者所受影响较大，变化模式较多，且不是人脸的显著特征，所以成为人脸识别中的干扰因素，应当尽量去除；后者包含丰富的面部特征，如边缘、轮廓和角点，是人脸最具有区分性的特征，且受外界因素变化影响较小，应当保留用于人脸识别。全变分模型的基本运算见公式 (4.18-4.20)。

$$I(x, y) = \rho(x, y) \cdot S = (\rho/\rho_l) \cdot (S_{\rho_l}) = \rho' \cdot S', \quad (4.18)$$

$$u = \arg \min \int_{\Omega} |\nabla u| + \lambda |\log(I) - u| dx, \quad (4.19)$$

$$v = \log(I) - u, \quad (4.20)$$

其中， I 是原图纹理，被分成肤色纹理 S' 和精细轮廓纹理 ρ' 。受光照影响较大的纹理信息 u 可通过最小化公式 (4.19) 中的模型目标函数得到。公式 (4.19) 中， $\int_{\Omega} |\nabla u|$ 为对 u

作全变分变换。 λ 为权重。公式 (4.20) 通过对数运算把原纹理两部分的乘积关系分离成和关系, 并通过减法运算, 剔除易受光照影响的纹理, 保留核心的轮廓信息 v , 作为全变分模型最终的输出结果。

4.3.2 TVM+LBP 人脸描述

鉴于 TVM 可以很好地保留并增强边缘和角点等对 LBP 敏感的特征, 故将 TVM 作为人脸图像预处理的方法, 与 LBP 方法结合, 以提高 LBP 在人脸识别上的性能。整个特征提取的过程, 见图 4.11 所示。TVM-LBPHS 为人脸特征的最终表示, 用于人脸识别。

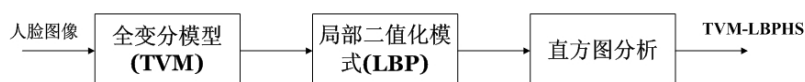


图 4.11 TVM-LBP 特征提取流程

Fig 4.11 Feature extraction process of TVM-LBP

4.3.3 实验

实验所采用的人脸库与 4.2.3 实验相同。

(1) 实验准备

首先, 在 FERET 人脸库的 3368 张头肩部图像中, 根据手工标定的双眼坐标, 把人脸图像截取出来, 并通过旋转平移等几何变换, 使得所有人脸图像的双眼在同一水平线上, 再通过缩放, 获得大小为 128×128 归一化的人脸样本。经 TVM 模型变换, 最终得到实验样本。部分实验样本如图 4.12 所示。



图 4.12 部分实验样本

Fig 4.12 Some experimental images

为了与 4.2.3 的结果对比, 实验中, LBP 图被分为 $16 \times 16 = 32$ 个子块, TVM-LBP 直方图序列采用公式 (4.16) 的 χ^2 距离进行比对。

(2) 结果与分析

本节提出 TVM-LBP 方法的实验结果列于表 4.3, 4.2.3 的实验结果同样也列于表 4.3。

比较实验结果可以发现, TVM-LBP 方法在光照变化的 fc 测试上得到的实验结果明显优于原始 LBP 方法, 同时 TVM-LBP 方法保证了在 fb、Dup I 和 Dup II 测试集上,

表 4.3 TVM-LBP 人脸识别实验结果

Table 4.3 Experimental result of TVM-LBP face recognition				
	fb(%)	fc(%)	Dup.I(%)	Dup.II(%)
TVM-LBP	92.6	78.5	52.2	30.7
Original-LBP	91.3	50.0	50.0	31.1

获得跟原始 LBP 方法相当的识别率。实验证明，当光照发生变化时，TVM-LBP 方法可显著改善人脸识别系统的性能。

4.3.4 结论

本节提出了 TVM-LBP 方法，利用全变分模型对人脸边缘、轮廓和角点等显著特征的增强，并结合了 LBP 算子对图像边缘、角点等局部变化特征敏感的特性，提取对于外界因素变化特别是光照变化，较为鲁棒的人脸特征用于人脸识别，明显提高了光照变化情况下的人脸识别率。值得注意的是，由于 TVM 模型的运算复杂度较高，运算比较费时，不能完全满足实时人脸识别系统的要求，对于这种情况，需要进一步研究解决。

4.4 本章小结

本章首先介绍 LBP 算子，接着描述基于 LBP 的人脸识别方法，针对一般 LBP 人脸描述维数较高，不适合实时应用的缺点，提出了压缩直方图的思想，有效地降低了特征维数，并保证了较高的人脸识别率。同时提出了一种新的基于全变分模型的 LBP 方法，利用全变分模型对人脸边缘、轮廓和角点等显著特征的增强，并结合了 LBP 算子对图像边缘、角点等局部变化特征敏感的特性，提取对于外界因素变化特别是光照变化，较为鲁棒的人脸特征用于人脸识别，明显提高了光照变化情况下的人脸识别率。

第五章 开放式人脸识别体系框架

本章将提出一种开放式人脸识别体系框架。首先先介绍人脸识别系统的背景；接着给出开放式人脸识别体系框架的三大设计目标，即通用性、开放性和分布式；然后给出该体系框架的设计思路和系统整体框架；最后结合三大技术平台（面向对象技术、数据库技术和 OpenCV）具体实现开放式人脸识别体系框架。

5.1 背景

人脸识别研究经过 40 年左右的发展，在人脸检测与人脸跟踪，人脸特征点的检测与标定，人脸特征描述与特征提取，人脸特征分类等各个领域，研究人员都展开了广泛而深入的研究，各种理论，各种算法层出不穷，同时实用系统的开发技术上也达到了一定的成熟度，因此从二十世纪九十年代后期开始，陆续出现了一些人脸识别商业系统。对于一个成熟的人脸识别商业系统而言，不但需要良好硬件平台的支持，软件系统和核心算法更为重要。同时自动人脸识别是一个复杂的过程，从视频流的采集直到最终的得到识别结果，中间需经历许多步骤的处理，因此各个步骤间数据的表示，运算及传递显得尤为关键。

与此同时，推出商用系统之前，人脸识别系统需要进行各种深入的测试，为了保证核心识别算法的有效性，人脸识别算法的研究者通过测试和比较来评估算法显得尤为重要，特别是在考虑到环境光照、人脸姿态、人脸配饰、人脸表情等因素变化的情况下，快速高效的评估人脸识别算法在静态下或实时应用下效果是非常困难的。

经对现有技术文献的检索发现，对于人脸识别算法的测试，目前大多数研究者都是通过自己搭建测试平台来实现的。现存的通用开放式人脸识别算法测试系统很少。美国 Colorado State University 曾开发出 Face Identification Evaluation System，该系统提供了四种基本人脸识别算法 (Eigenfaces, Combination PCA and LDA, Bayesian Intrapersonal/Extrapersonal Image Difference Classifier, Elastic Bunch Graph Matching)，并可以得到标准人脸库 FERET 在上述四种算法上的测试结果，但该系统采用文本方式保存实验数据，读写效率比较低，而且该系统扩展性很差，用户很难将自己的人脸识别算法嵌入进行测试。由美国 NIST(National Institute of Standards and Technology) 主持开发的 FRGC(Face Recognition Grand Challenge) 是学术界普遍使用一套人脸识别算法评估系统，该系统主要针对高分辨人脸图像、三维人脸和人脸图像预处理下人脸识别算法的测试，用户可以方便快速的把自己的人脸识别算法嵌入评估系统进行测试，实验数据以 XML 文件格式存储，加快了读写效率，但该系统只支持 FRGC 人脸库，不支持其他标准人脸数据库，而且上述系统只能对 image-based 人脸识别算法进行测试，无法进行 video-based 人脸识别算法的测试。

因此本章提出一种通用的开放式人脸识别体系框架,该框架克服现有技术中的不足,提供一个同时满足前期测试和实际应用的人脸识别平台。我们充分利用数据库在管理海量数据上的优势,采用模块化和层次化的设计,实现人脸样本的采集,人脸识别数据的管理,人脸分类器的训练,标准人脸库的导入,静态图片人脸识别测试,动态实时人脸识别测试,生成测试报告等一系列功能,各个模块可独立工作或共同工作以完成测试和实时识别等任务。此外,在此体系框架下,经过简单扩展,即可实现性别识别,表情识别,年龄估计等算法的测试与实用。

5.2 设计目标

本章旨在提出一种通用的开放式人脸识别体系框架,从而为实时人脸识别系统提供模型参照。为实现支持不同的人脸识别方法,能够同时满足人脸识别算法的测试与实时应用,该体系框架必须符合以下特性:

5.2.1 通用性

分析人脸识别任务,我们认为,通用性包括五层含义:

一、该体系框架应该能够满足人脸识别算法的测试与实际应用的双重需要,若某种人脸识别算法在测试阶段时可以达到很好的效果,则不需要改动或经过很小的改动,该算法便可以直接用于实际应用,这样可以缩短测试到实际应用的开发周期;

二、该体系框架应该尽可能的兼容现有较流行的人脸识别算法,使之不需改造或只经过功能封装就可嵌入这套体系结构中,即把不同人脸识别算法的共同属性抽象出来,通过统一的模块或接口来实现;

三、由于人脸识别方法法主要包括基于图像的人脸识别方法(如照片相似度的比对等)和基于视频的实时人脸识别方法(如实时的视频监控等)两个方面,因此该体系框架应该同时适用于上述两个方面的应用,即提供不同的输入模式(静态图片和实时的视频流);

四、现今有几种不同的视频采集方式(USB 数字摄像头,PCI 视频采集卡+模拟摄像头,USB 视频采集卡+模拟摄像头,PC104 视频采集卡+模拟摄像头),该体系框架应该同时兼容以上主流的视频采集方式;

五、为了使得人脸识别算法的测试有统一的评判标准,从人脸识别的研究伊始,许多研究者或研究机构就采集了不同特点和规模的标准人脸数据库,因此,本文提出的体系框架应该适用于在不同的标准人脸数据库上测试人脸识别算法(如 FERET, CAS, PIE 等);

实际上,现有各种人脸识别算法或实用系统程序的编制千差万别,提出一种通用的开放式人脸识别体系框架有着很大的实际意义:

(1) 通用的体系框架模型可以大大提高人脸识别程序的可重用性和代码效率;

(2) 以一种通用的开放体系模型为参照, 研发人脸识别任务中不同的功能模块, 可以使之更加模块化、平台化, 使各组件的智能升级换代和相互替换成为可能。

5.2.2 开放性

通用的开放式人脸识别体系框架的开放性设计应满足以下含义:

(1) 开放性的模型

一个具有良好开放性的人脸识别体系框架不仅应兼容现有识别方法, 还应具有一定的预见性, 使未来对人脸识别实用软件系统以该体系结构作为参照, 并在其中找到相应的功能原型。因此必须在广泛分析已有人脸识别方法的基础上, 对人脸识别各个必要过程的各要素进行抽象; 按照各个模块的功能定义其抽象原语, 而不规定各模块的实现细节。

(2) 体系框架的实现必须开放

这里的“实现”是指将体系框架以一定的计算机语言编写成可二次开发的代码库。通用的开放式人脸识别体系框架的实现选择纯面向对象技术并采用 C++ 语言作为具体实现平台。在体系框架的实现过程中, 首先必须保证封装体系中各原语的对象具有开放的设计, 使之成为具有通用意义的基类; 其次各对象的接口须合理定义, 使之能够进行接口扩充, 增加相应功能。例如, 可以通过继承接口的方法扩展行为对象的输入输出, 使之能够被动态修改其参数。

5.2.3 分布式

通用的开放式人脸识别体系框架的分布式设计应满足以下两层含义:

(1) 该体系框架的各功能模块可以独立分布于同一硬件平台上 (例如同一台 PC 机上或服务器上), 完成各自相对独立的功能, 并可以通过协同工作, 完成整个人脸识别任务;

(2) 该体系框架的各功能模块可以独立分布于不同的硬件平台上 (如 LAN 中不同 PC 机或嵌入式平台, 或 Internet 上不同的主机与终端等), 以满足不同的应用需求。例如, 以人脸识别为门禁的酒店系统中, 所有住客的信息存储在中央服务器上, 当住客入住酒店登记时, 前台的终端可采集住客的人脸图像信息, 并发送到中央服务器上, 中央服务器然后进行特征提取, 同时将该住客的特征信息及其他信息记录于后台数据库; 当住客进入自己的客房时, 安装于客房门上的嵌入式采集设备采集住客的人脸图像发送给中央服务器, 中央服务器完成识别后, 若识别结果与相应客房登记信息吻合时, 发送开门指令驱动客房的电子锁完成开门动作, 这样住客就可以通过自己的人脸进入客房; 当住客在酒店餐厅或其他场所消费结帐时, 服务员通过手持设备采集住客的人脸图像发送给中央服务器, 完成识别得到住客身份时, 便可从住客的账户上扣除相应的金额, 这样住客可通过自己的人脸完成付帐。该酒店的中央服务器甚至可以于公安部门的人脸数据库连接, 当疑犯入住酒店, 便可及时向公安部门报警。

5.3 设计思路

针对以上目标,本文研究在分析人脸识别测试与实际应用的任务需要的基础上,提出了以下的设计思路:

5.3.1 系统整体框架

开放式人脸识别体系框架的整体框架图见 5.1。为了体现上一小节提出的通用性,开放性和分布式的设计目标,系统整体框架采用分布式的模块化设计,各个模块以数据库为桥梁,通过读写数据库,实现数据交互,共同完成人脸识别任务。各个模块功能将在 5.3.2 小节中给出详细描述。

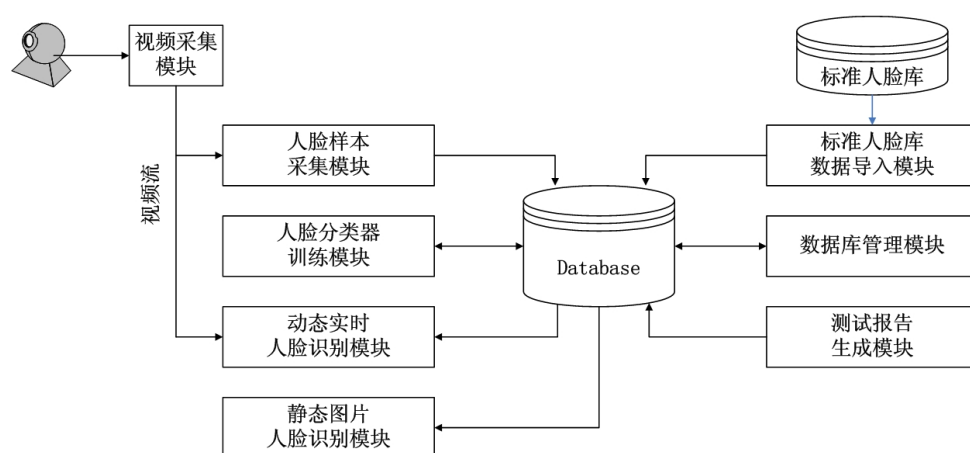


图 5.1 开放式人脸识别体系框架结构

Fig 5.1 The structure of OFFR

5.3.2 模块化

本系统的设计采用模块化思想,采用模块化设计具有如下优点:

- (1) 各模块功能相对独立,可实现分布式的设计思想;
- (2) 各模块可独立于整个系统框架之外,实现相应的功能,当单独的模块发生改变或功能升级时,不会对其他模块及整个系统造成影响;
- (3) 各模块可独立开发,不但便于调试,降低开发难度,同时可以协同开发,加快开发进度。

下文将对系统整体框架图 5.1 中各模块的实现的功能进行详细介绍。

(1) 视频采集模块

该模块通过 VFW(Video For Windows) 与摄像头或视频采集卡通信,实现视频流的实时采集。该模块可根据不同的摄像头或视频采集卡使用不同的视频采集方式。

(2) 人脸样本采集模块

人脸识别分类器的训练需要一定数量的人脸样本图片,该模块能够自动采集适用于人脸识别的人脸样本图片。该模块同时具有人脸检测功能,可在每帧视频图像中检测出

是否存在人脸,以及存在的人脸数量,位置和大小。当在视频图像中检测到人脸时,该模块可以自动定位人脸的主要特征点(如眼睛,眉毛,鼻子,嘴巴等),并根据这些特征点的位置,完成人脸图像的标准化(如旋转,平移,剪切,缩放,预处理,蒙板等),最终得到用于训练人脸分类器的人脸样本,并将这些样本数据存入后台数据库。

(3) 数据库管理模块

后台数据库的海量数据需要进行管理,通过数据库管理模块系统管理员可管理后台数据库中各种涉及到人脸识别数据。该模块可进行添加,删除,修改,备份等操作,实现对数据库的人员,人脸样本图片、人脸特征、人脸分类器等数据的维护。

(4) 人脸分类器训练模块

人脸识别的核心算法体现在分类器的训练,人脸分类器训练模块可根据不同的算法训练出不同人脸分类器。该模块可高效的组织训练样本(包括人工手动选择和自动随机选择),并可以根据样本图片的属性设置相应过滤器,选择出不同属性(如表情,年龄,光照,姿态,性别,配饰等)的人脸样本图片来训练不同的人脸分类器,以满足人脸识别算法在表情,年龄,光照,姿态,配饰等外界因素变化时的应用。通过该模块训练出的人脸分类器数据自动存入后台数据库。

(5) 标准人脸库导入模块

标准人脸数据库中有成千上万张人脸图像,通过手工的方式处理这些人脸图片显然不现实,因此标准人脸库导入模块必不可少。该模块可将用于测试人脸识别算法性能的通用标准数据库导入后台数据库。一般的标准人脸数据库包含不同的测试集合和训练集合,每个集合中一般有数千张或数万张人脸图片,该模块根据人脸样本的不同属性,自动的设置样本图片的属性,从而实现样本的分类,并将人脸样本图片存入后台数据库。

(6) 静态图片人脸识别模块

有些人脸识别任务,要求根据给定的静态图片判断照片中人像的身份,静态图片人脸识别模块提供静态图片的输入接口,并读取后台数据库中已有的人脸样本图片和人脸分类器,完成识别任务,并给出识别结果。

(7) 动态实时人脸识别模块

动态实时人脸识别模块根据摄像头采集到人脸图像,动态实时完成人脸识别。该模块在外接的摄像头采集到的视频流中,检测到人脸,完成人脸图像的标准化得到适于测试的人脸图像样本,供人脸分类器识别,最终给出识别结果。

(8) 测试报告生成模块

人脸识别算法需要经大量人脸样本的测试,测试报告对算法研究者或实用系统的开发人员改进算法具有指导意义。测试报告生成模块可根据测试结果生成测试报告。该模块可以读取测试结果,统计出人脸识别算法的性能指标(如识别率,错误拒绝率,错误接受率等),并可生成相应曲线图表,以便不同人脸识别算法的比较,从而得到相对性能更优越的识别算法,在实际系统中应用。

5.3.3 层次化

开放式人脸识别体系框架的层次见图 5.2。整个体系大致分为如下四层：

(1) 预处理

人脸检测，人脸特征点定位与定位及人脸图像的标准化，都归为预处理层，视频流中每帧图像或静态人像图像经过此层后，输出为适于人脸识别的标准人脸图像；

(2) 分类器训练

根据识别算法性质不同，分类器训练层的数据流分为两部分，一部分为非泛化的算法 (以子空间方法为例)，另一部分为泛化较好的算法。训练数据经过此层后生成相应的人脸分类器；

(3) 识别与测试

该层实现识别与测试任务，对于子空间方法，需要经过子空间的映射，其他泛化性较好的方法，直接通过特征的比对完成识别；

(4) 结果分析与输出

该层完成人脸识别任务的最后一步，对上一层识别与测试的结果进行分析，生成识别结果或测试报告等。

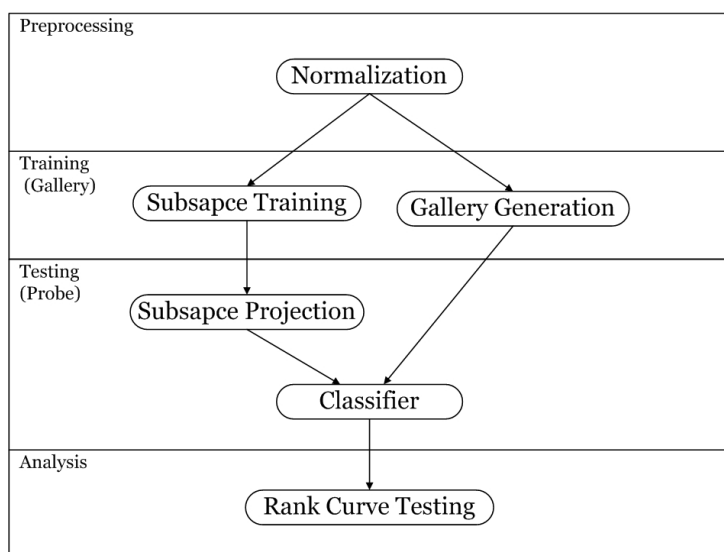


图 5.2 开放式人脸识别体系框架层次结构

Fig 5.2 Layer structure of OFFR

每帧视频流的图像经过上述四层中或其中的几层，最终得到人脸识别的结果。

5.4 技术平台

5.4.1 面向对象技术

面向对象技术 [104,105] 不仅仅是一种编程方法，随着它的发展和完善，面向对象技术正日益成为一种系统建模和系统构建的方法。面向对象技术最初是作为提供代码可重

用性的编程方法提出的。随着 20 世纪 90 年代面向对象技术的广泛应用,人们发现仅仅在代码阶段使用面向对象的概念是不够的,需要在分析和设计阶段就使用面向对象技术,于是面向对象分析和面向对象设计技术迅速发展起来,并得到了工业界和研究学者们的一致认可。

面向对象方法都支持三种基本的活动:对象和类,描述对象和类之间的关系,以及通过描述每个类的功能定义对象的行为。对象是指既包含数据域,又包涵数据域操作方法的编程单元。应用面向对象技术可以快速、准确、高效的建立系统编程模型,其编写的代码具有良好的可重用性,开放的接口,高效的代码效率。

在开放式人脸识别系统中,通用、规范的人脸描述及附件设计是系统设计的核心内容。如何设计一套规范化人脸特征描述,将各种人脸特征的描述规范约束起来,并保持各自的功能特性和开放性,是本课题的最大难点。

传统上,不同系统间的接口协议多是以面向过程的应用函数调用 (API) 方式实现的,即以 API 方式提供实现一定接口功能的函数库,库中函数接受一些参数,用以完成指定功能。用户通过组织调用这些函数实现相应功能的接口。面向过程的 API 方式最大的缺陷是可重用性差,随着接口复杂程度的增加,API 函数库变得越来越庞大,用户调用越来越困难。

随着面向对象的系统建模和实现技术的普及,面向对象的 API 接口方式已成为接口协议设计的发展趋势。相对于面向过程的 API 接口,面向对象的接口提供给用户的是系列可继承的对象和相应的方法。用户只需实例化相应对象或实现继承类,就可实现接口或对原接口进行扩展。与面向过程的 API 接口相比,设计良好的面向对象的接口有以下优点:

(1) 高度模块化

模块化是软件开发过程中处理复杂问题的重要手段。在面向对象的接口中,各部分功能被分解成不同组件,每个组件的内部信息被封装起来,组件间所有的交互都通过各个组件可被调用的方法和属性进行,避免了许多内部参数的传递;

(2) 复用性好

对象继承允许对象实现复用。通过继承,新的组件可以复用旧的组件的功能,并扩展出新的功能,而不必重复编制旧的功能模块。这可以使接口协议的升级变得更加简单,并保持很好的兼容性,且在此基础上,整个接口的拓扑结构可变,从而实现通用灵活的接口协议;

(3) 可维护性好

因为面向对象的接口中,各个组件相互独立,整个复杂的接口被划分为相对简单、独立的组件。一个组件的改动不会影响其他组件,使维护工作更加简单和可靠。

因此本文选择面向对象技术建立开放式人脸识别系统的编程模型和代码实现。

5.4.2 数据库技术

数据库技术是 20 世纪 60 年代后期产生和发展起来的一项计算机数据管理技术，是现代计算机信息系统和计算机应用系统的基础和核心。它的出现使计算机应用到人类社会的广阔领域。目前数据库的建设规模和性能、数据库信息量的大小和使用频度已经成为衡量一个国家信息化程度的重要标志。数据库技术也成为计算机科学与技术学科的一个重要分支。

在开放式人脸识别系统中，

数据库技术与人工管理和文件系统相比，主要有如下特点：

(1)数据高度结构化

数据结构化是数据库系统与文件管理系统的根本区别。从整体上讲，文件系统的据是没有结构的。比如文本文件，其中的数据就没有结构，即文件的记录之间没有联系。以人脸识别为例，可以把一个用户的用户信息放在文本文件的一个段中，也可以把用户的人脸图片信息放在一段中，尽管该文本文件可以显示用户和人脸图片的信息，但没有相应的记录结构，处理数据较为麻烦。如果查找一个用户的信息，必须浏览整个文件。当数据量比较少时，问题不大，但如果数据量很大，浏览成千上万条用户信息将是一件非常困难的工作。

数据库系统不仅要考虑数据项之间的联系，还要考虑数据类型之间的联系。例如一个人脸识别系统中，如果采用多种特征融合，如果采用文件系统管理相应数据，数据都必须存放在相互分割的文件中。显然同一个用户的不同特征之间是有客观联系的，而使用文件管理系统就人为地将它们割裂开来，使得相互联系只能通过应用程序才能体现出来。

数据库系统实现了整体数据的结构化，这是数据库的主要特征。在数据库系统中，数据不再针对某一项应用，而是面对所论及问题的整体。

(2)数据的共享性高、冗余度低、易于扩充

数据库中的数据是高度共享的，也就是说：同一个用户可以因不同的应用目的访问同一数据，不同用户可以同时访问同一数据，即所谓的“并发访问”。

数据共享程度直接关系到数据的冗余程度。数据库系统是从整体架构来描述数据的，数据不再面向某个特定的应用程序而是面向整个系统，因而可以大大减少数据冗余，节省存储空间。

(3)数据独立性高

用户只需关注数据库的名称、数据文件的名称和文件中的属性名称等逻辑概念，而不用过多考虑数据的实际物理存储，也就是不需要关心实际数据究竟存储在磁盘的什么位置。这样以来，数据与应用程序之间就具有了独立性，数据的定义和描述就可以从应用程序分离出来；而且由于具有数据独立性，就有可能开发出专门用于数据管理的系统软件，即数据库管理系统，通过这个系统具体处理数据的存取路径等技术细节，从而简化了应用程序的编写，减少了应用程序维护和修改的开销。

(4)数据库管理系统

数据库是一个多级系统结构,需要一组软件提供相应的工具进行数据的管理和控制,以达到保证数据的安全性和一致性的基本要求。这样一组软件就是数据库管理系统(Database Management System, DBMS)。DBMS 的功能随着系统的不同而有所差异,但一般都具有一下功能:1、数据的并发控制;2、数据的安全性保护;3、数据的完整性检查;4、数据库故障恢复等。

现代数据库中除了提供 DBMS 之外,还提供各种各样的应用开发工具和客户端数据库工具,例如 ORACLE 的 SQL Plus、Informix 的 DBAccess、SQL Server 的查询分析器等。

5.4.3 OpenCV

OpenCV [106] 是 Open Source Computer Vision Library 的缩写,它是由 Intel 公司位于俄罗斯的研究实验室所开发的一套可免费获得的由一些 C 函数和 C++ 类所组成的库,用来实现一些常用的图像处理及计算机视觉算法。OpenCV 与 Intel 公司所开发的另一图像处理库 IPL(Intel Image Processing Library)兼容, IPL 用于实现对数字图像的一些低级处理,而 OpenCV 则主要用于对图像进行一些高级处理,比如说特征检测与跟踪、运动分析、目标分割与识别以及 3D 重建等。由于 OpenCV 的源代码是完全开放的,而且源代码的编写简洁而又高效,特别是其中大部分的函数都已经过汇编最优化, OpenCV 的代码执行效率是非常高的,所以近年来在国外的图像处理相关领域中被广泛地使用,成为一种流行的图像处理软件。

总结起来, OpenCV 作为开放的数字图像处理和计算机视觉软件平台,有以下特点:

- 1、开放的 C/C++ 源码;
- 2、基于 Intel 处理器指令集开发的优化代码;
- 3、统一的结构和功能定义;
- 4、强大的图像和矩阵运算能力;
- 5、方便灵活的用户接口
- 6、同时支持 MS Windows 和 Linux 平台

因此,本文选择 OpenCV 作为二次开发的工具,辅助实现开放式人脸识别系统。

5.5 开放式人脸识别体系框架的实现

5.5.1 数据库实现

后台数据库是整个开放式人脸识别体系框架的中心,不但所有与人脸识别任务相关的数据都存储于后台数据库中,更重要的是各功能模块之间的通信与数据交换,均通过对数据库读写来完成,因此,后台数据库担当了“桥梁”的角色,把各功能模块连成一个有机的整体,协同工作以完成人脸识别任务。

Microsoft SQL Server 是普遍应用的企业级数据库，同时提供了功能强大的 DBMS，本文采用 Microsoft SQL Server 2000 做为后台服务器，搭建整个开放式人脸识别体系框架。开发实际应用系统时，从便于移植和小数据量的角度考虑，可选用 Microsoft Access 为后台数据库，在下一章中，我们设计的基于人脸识别的考勤系统，即由 Access 支持。此外，若在 Linux 平台下，可采用 MySQL 这一著名开源的数据库为后台支持数据库。

使用关系型数据库为基础搭建数据库，关系数据库的整体结构如图 5.3 所示。下面将对数据库中各表进行简要说明：

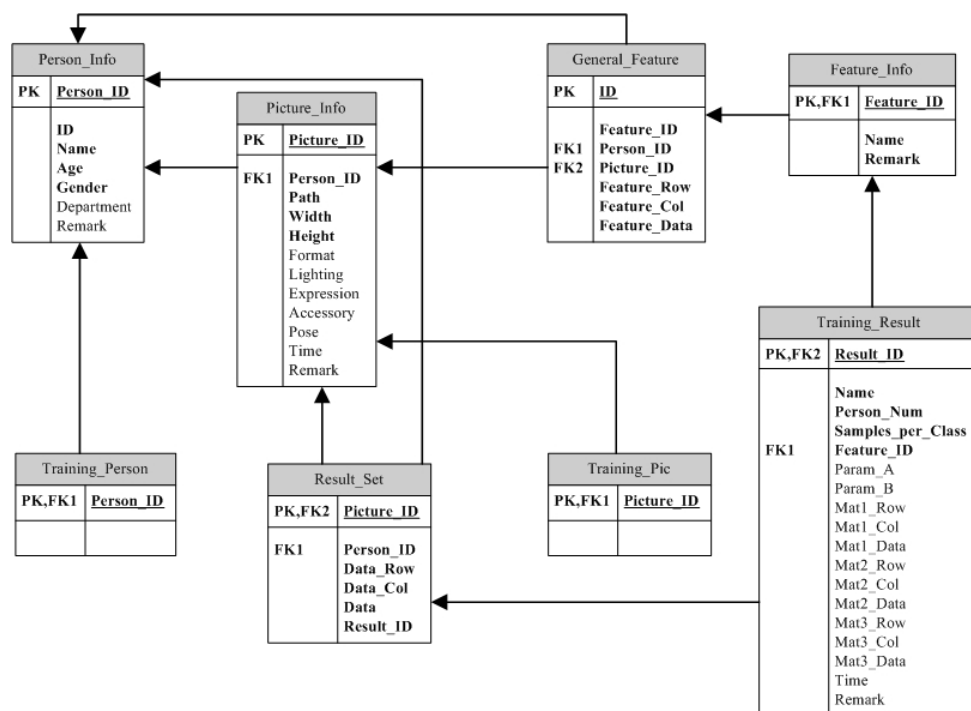


图 5.3 数据库结构

Fig 5.3 The structure of database

† *Person_Info* 表，记录人员信息。详参附录 A 表 B.1。

† *Picture_Info* 表，记录样本图片信息。详参附录 A 表 B.2。

† *Feature_Info* 表，记录体系框架中存在的特征描述的方法。详参附录 A 表 B.3。

† *General_Feature* 表，记录各样本图片在不同特征描述下的特征表示。详参附录 A 表 B.4。

† *Training_Result* 表，记录训练出来不同分类器信息。详参附录 A 表 B.5。

† *Result_Set* 表，记录使用子空间方法，经映射变换后得到人脸特征描述。详参附录 A 表 B.6。

† *Training_Person* 表和 *Training_Pic* 表，分类器的训练需要组织训练样本，这两张表为临时生成的表，用于记录用于训练分类器的人员与图片的信息。详参附录 A 表 B.7 和 B.8。

5.5.2 类实现

如本章 5.4 节中介绍, 开放式人脸识别体系框架选择面向对象的接口方法和 C++ 语言作为实现方法。类结构图见图 5.4。

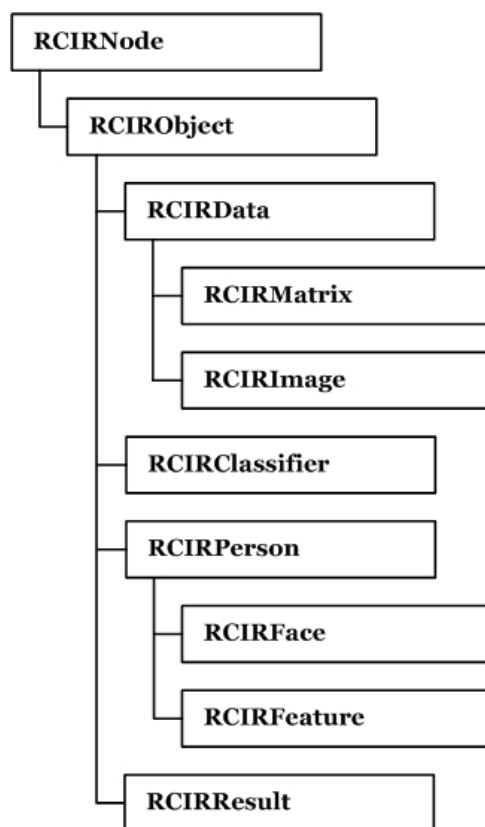


图 5.4 类结构图

Fig 5.4 Structure of classes

本节将概要的介绍各个类和接口, 详细内容参见程序文档与代码注释。

† *RCIRNode* 类为开放式人脸识别体系框架的基类, 该类为抽象类, 由它派生出的其他的子类;

† *RCIRObject* 类为 *RCIRNode* 的一级派生子类, 该类为抽象类, 为体系框架中实体类的基类;

† *RCIRData* 类为 *RCIRObject* 的子类, 实现体系框架的数据描述, 下面派生出 *RCIRMatrix* 类和 *RCIRImage* 类, 描述把数据描述具体化与体系框架中的核心数据类型矩阵和图像, 两个类中封装了 OpenCV 中的基本数据类型 *CvMat* 和 *IplImage*, 以简化实现。其接口包括了针对数据库的读写操作等;

† *RCIRClassifier* 类为 *RCIRObject* 的子类, 实现体系框架的中人脸分类器的描述, 并定义读写接口;

† *RCIRPerson* 类为 *RCIRObject* 的子类, 实现体系框架中核心“人”的描述及接口, 下面派生出 *RCIRFace* 类和 *RCIRFeature* 类, 为人脸描述和特征描述的抽象化描述, 体

系框架的开放性主要由这两个核心类实现；

† *RCIRResult* 类为 *RCIRObject* 的子类，为辅助类，用于描述识别结果，提供快速排序，查找等接口，辅助实现识别结果的分析与输出。

5.5.3 GUI 实现

良好的用户图形界面 (Graphic User Interface, GUI) 也是一个开放式系统框架的重要特征。本文中使用 MFC 作为开发工具，为各个模块提供了基于 Windows 平台的 GUI。若要开发 Linux 平台下的 GUI 可使用 KDE 或 GNOME，实现过程类似。下文将对体系框架中几个功能模块的 GUI 进行简要说明：

(1) 人脸样本采集模块

人脸样本采集模块的界面如图 5.5 所示。通过菜单项，可以设置采集摄像头的属性，采集视频流的格式，人脸采样的频率，采集人脸样本的大小，人脸采集的方式 (自动，手动)，人脸图像预处理方式等，这样就可以采集到满足各种需要的人脸样本图像。界面主要分为三个部分，上半部分为视频采集窗口和采集样本图片的大图预览，中间部分显示已经采集的样本图片，下半部分为人员信息框，通过此框可以选择库内已存在的人员或添加新的人员，以便使采集的人脸图像与人员身份对应起来。工作流程见附录图 A.3。

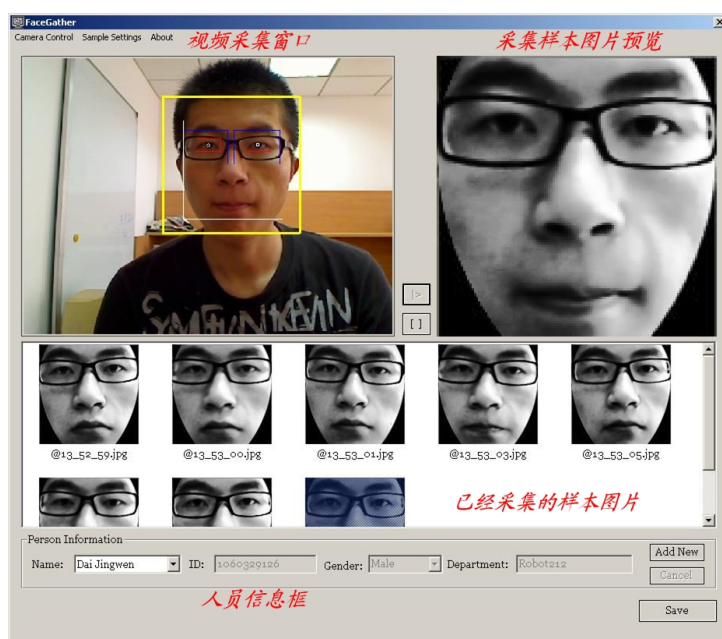


图 5.5 人脸样本采集界面

Fig 5.5 GUI of facial image sampling module

(2) 数据库管理模块

数据库管理模块的界面如图 5.6 所示。通过该界面用户可以完成对人员和样本图片的管理。界面主要分为两个部分：左侧为人员列表，人员信息框，以此用户可以管理人员信息 (浏览，添加，删除和修改)；右侧为样本图片大图，样本图片信息框，样本图片

预览缩略图,以使用户实现对样本图片的各种操作。

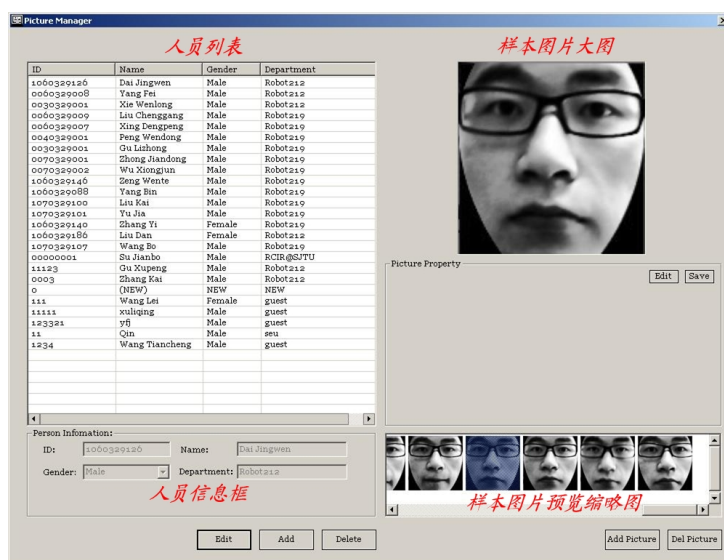


图 5.6 数据库管理界面

Fig 5.6 GUI of database management module

(3) 人脸分类器训练模块

人脸分类器训练模块的界面如图 5.7 所示。人脸分类器采用训练向导的模式,用户可在图 5.7(a) 所示不同的训练方式 (如开始新的分类器训练或更新已有分类器),启动不同的训练向导见图 5.7(b) 和 5.7(c),在向导界面下设置或修改分类器的参数并选择训练用人脸样本图片,以完成人脸分类器训练任务。工作流程见附录图 A.4。

(4) 静态图片人脸识别模块

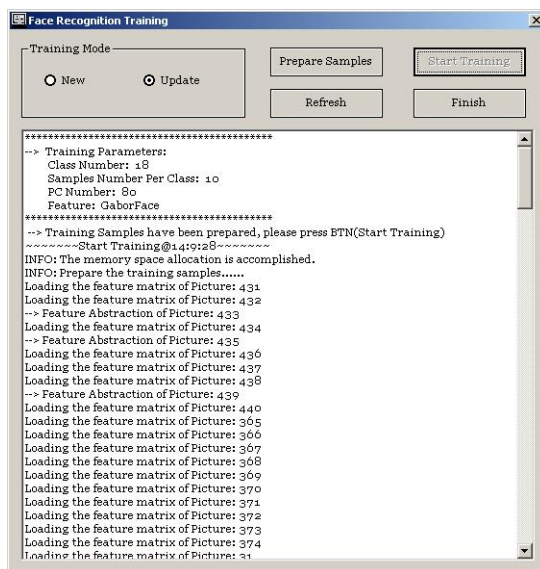
静态图片人脸识别模块的界面如图 5.8 所示。用户可选择静态样本图片完成识别,界面显示数据库内最相似的人的图像及身份,并且以列表的形式给出,库内其他样本与测试样本的相似度,以便帮助用户分析识别结果。工作流程见附录图 A.2。

(5) 动态实时人脸识别模块

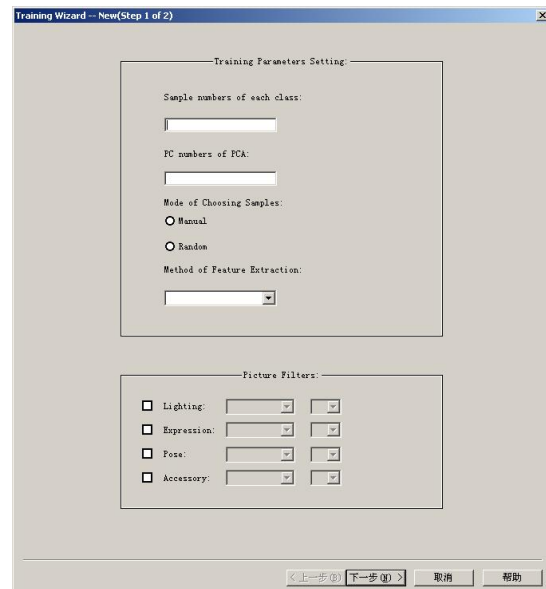
动态实时人脸识别模块的界面如图 5.9 所示。通过动态实时的采集视频流,并在视频流中的每帧图像中完成人脸检测,人脸特征点定位,人脸识别等过程,最终在用户界面上给出识别结果。工作流程见附录图 A.2。

5.6 本章小结

本章提出一种新的开放式人脸识别体系框架,根据开放式人脸识别体系框架的通用性、开放性和分布式的三大设计目标,具体描述了体系框架的设计思路和系统整体框架;并以面向对象技术、数据库技术和 OpenCV 为技术平台,具体实现开放式人脸识别体系框架。



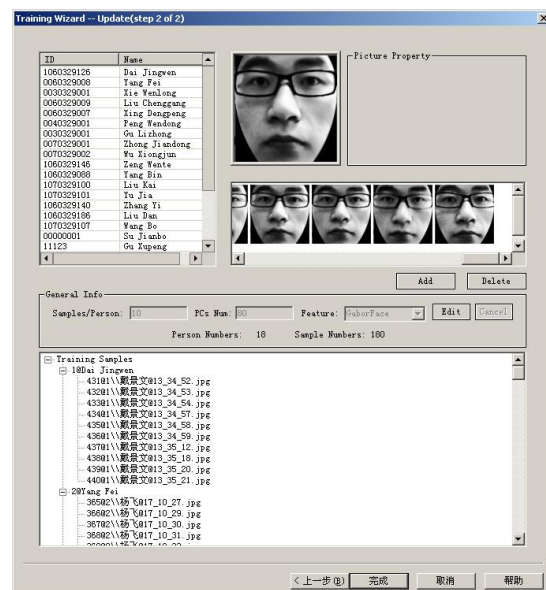
(a)



(b)



(c)



(d)

图 5.7 人脸分类器训练界面
Fig 5.7 GUI of facial classifier module

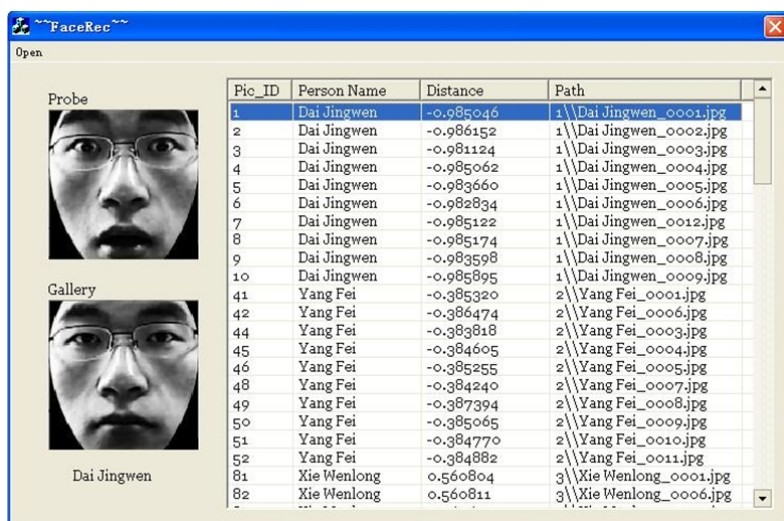


图 5.8 静态图片人脸识别界面

Fig 5.8 GUI of image face recognition module



图 5.9 动态实时人脸识别界面

Fig 5.9 GUI of real time face recognition module

第六章 基于人脸识别的考勤系统

上一章中，我们提出了开放式人脸识别体系框架，期望在此体系框架之下，不同的人脸识别算法可以完成不同的人脸识别任务。本章以办公室环境的日常考勤为应用目的，以开放式人脸识别体系为基础，采用 SFFS-Gabor 人脸识别方法实现基于人脸识别的考勤系统，以此为实施例，验证开放式人脸识别体系的开放性与通用性。

6.1 系统概述

6.1.1 编写目的

(1) 解决技术问题及软件优点

实现实时的人脸检测与识别，将人脸特征作为唯一 ID，实现稳定、高效的考勤，软件界面友好，使用简单方便。

(2) 应用领域行业

该系统不但可以用于各公司，工厂，学校的门禁与考勤；同时还可以用于各种人脸识别算法在实际应用环境下的长期测试与比较。

6.1.2 硬件及软件运行环境

基于人脸识别的考勤系统目前运行于 PC 机平台上，因此需要一定的硬件及软件运行环境。

(1) 硬件环境

系统运行所需的最低硬件配置和推荐硬件配置见表 6.1 和表 6.2。

表 6.1 最低硬件配置

Table 6.1 The minimum hardware requirements

CPU	Intel(R) Pentium(R) 4 1.8G (或AMD相应产品)
内存	128M DDR
剩余硬盘空间	200M
外设	鼠标、键盘等

表 6.2 推荐硬件配置

Table 6.2 The standard hardware requirements

CPU	Intel(R) Pentium(R) D 2.8G (或AMD相应产品)
内存	512M DDR2
剩余硬盘空间	500M
外设	鼠标、键盘等

从上述两张表格可以看出，目前主流的 PC 机的配置，完全可以胜任该人脸识别考勤系统的硬件要求，因此我们的系统具有普遍适用性。

(2) 软件环境

考勤系统的运行需要相关软件的支持，具体要求见表 6.3。

表 6.3 软件配置

Table 6.3 The standard software requirements

操作系统	Windows 2000, Windows XP, Windows Vista
支持软件	Microsoft Excel 2000 及以上版本 Microsoft Access 2000 及以上版本 相应网络摄像头或视频采集卡的驱动程序

6.1.3 系统功能与特点

该软件系统主要实现如下功能：

1. 用人脸作为唯一的 ID 实现门禁或考勤；
2. 方便快捷的采集新用户的人脸样本，并根据采集来的样本图片，通过分析学习，得到新添加用户的显著面部特征，作为识别依据；
3. 程序后台由 Microsoft Access 提供数据库支持，有效的实现数据管理；
4. 智能的出勤情况分析，按照多种方式输出 Excel 考勤报表。

该软件系统具有如下技术特点：

1. 快速准确的完成人脸检测和特征点的定位；
2. 快速的实现人脸面部显著特征的提取，通过分析比较，得出识别结果；
3. 采用多线程设计，各功能模块可同时工作；
4. 自动分析出勤情况，并自动生成 excel 的考勤报表。

经过长期测试，该软件系统可以实现如下技术指标：

1. 适用人群规模：30 – 1000 人；
2. 人脸识别的准确率：> 90%；
3. 识别每个用户的时间：< 1000ms。

6.2 系统实现

6.2.1 系统结构

考虑到 Microsoft SQL Server 数据库, 要求后台运行服务器程序, 因此不便于移植; 而一般考勤系统的数据量较小, Microsoft Access 数据完全可以胜任, 在软件发布时, 安装包内打包装入 Microsoft Access Engine 即可保证获得正常的数据库支持。

因为考勤系统的功能较为单一, 所以需要对开放式人脸识别体系框架进行裁减:

(1) 精简数据库

由于考虑了通用性的缘故, 开放式人脸识别体系的数据库设计中存在一定的冗余度, 在实用系统的实现中, 我们需要对数据库进行精简, 仅保留必要的表, 精简后的数据库结构见图 6.1 所示。

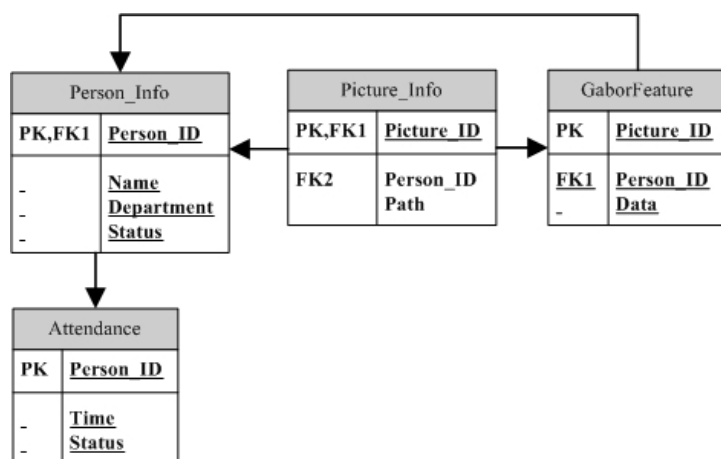


图 6.1 考勤系统数据库结构

Fig 6.1 Database structure of checking attendance system

由数据库结构库可见, 我们简化了 *Person_Info* 和 *Picture_Info* 表, 把 *General_Feature* 表具体化为 *GaborFeature* 表, 并增加了 *Attendance* 表以记录出勤情况。

(2) 精简功能模块

功能模块仅保留视频采集模块, 人脸样本采集模块, 动态实时人脸识别模块, 人脸分类器训练模块即可满足实用考勤系统的需要。

(3) 整合功能模块

为了方便用户使用, 我们把各个功能模块以线程的形式, 整合在同一个软件体系下。主要功能线程有: 视频采集线程, 人脸检测线程, 人脸识别线程和训练线程。系统启动时, 视频采集线程, 人脸检测线程, 人脸识别线程同时启动, 但后面两个线程处于休眠状态, 当视频采集线程采集到视频图像时, 激活人脸检测线程作人脸检测当检测到人脸时, 通过特征点定位, 完成人脸标准化之后, 激活人脸识别线程作特征提取与比对并输出识别结果, 然后人脸识别线程休眠; 当要新增加用户进入考勤系统时, 人脸检测线程, 人脸识别线程休眠, 启动训练线程, 训练完成之后, 终止训练线程。以多线程的方式协

同工作。系统即以上述多线程形式协同工作。

6.2.2 附加功能

实际考勤需要统计功能，因此我们的系统中增加了分析统计功能，并能够以 excel 表格的形式输出考勤表。

图 6.2 为考勤分析的界面，通过简单操作管理员可以了解到员工的出勤情况。

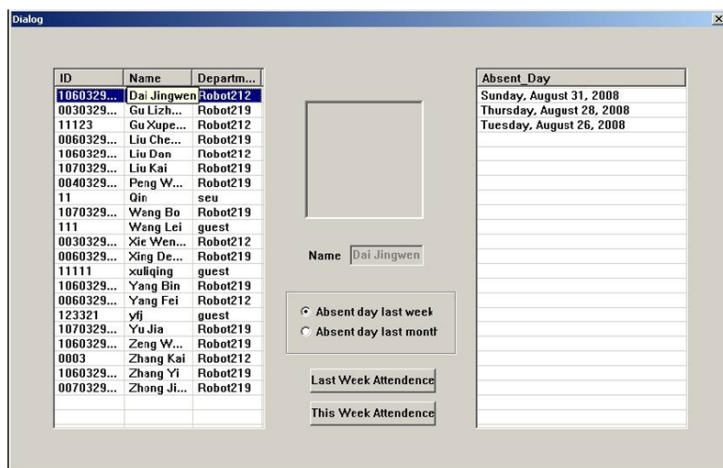


图 6.2 出勤分析界面

Fig 6.2 GUI of attendance analysis

图 6.3 为出勤情况表，Excel 表格中记录了各个用户每天到达及离开的时间。

	Monday Morning	Monday Afternoon	Monday Night	Tuesday Morning	Tuesday Afternoon	Tuesday Night	Wednesday Morning	Wednesday Afternoon	Wednesday Night	Thursday Morning	Thursday Afternoon	Thursday Night	Friday Morning	Friday Afternoon
1. Wang Lei														
2. xuliqing														
3. yfj														
4. 戴文														
5. 周立														
6. 周立														
7. 周立														
8. 周立														
9. 周立														
10. 周立														
11. 周立														
12. 周立														
13. 周立														
14. 周立														
15. 周立														
16. 周立														
17. 周立														
18. 周立														
19. 周立														
20. 周立														
21. 周立														
22. 周立														
23. 周立														
24. 周立														
25. 周立														
26. 周立														
27. 周立														
28. 周立														
29. 周立														
30. 周立														
31. 周立														
32. 周立														
33. 周立														
34. 周立														
35. 周立														
36. 周立														
37. 周立														
38. 周立														
39. 周立														
40. 周立														
41. 周立														
42. 周立														
43. 周立														

图 6.3 出勤表

Fig 6.3 Table of attendance in Excel

此外由于考勤系统运行的长期性,所以考勤系统是测试人脸识别算法的良好平台,为此我们增加了人脸识别准确率统计功能。图 6.4 为此项功能的界面,从可以看出,第四章提出的 SFFS-Gabor 的人脸识别方法,在运行 RCIR 实验室的考勤系统中达到 98.63% 的识别率。

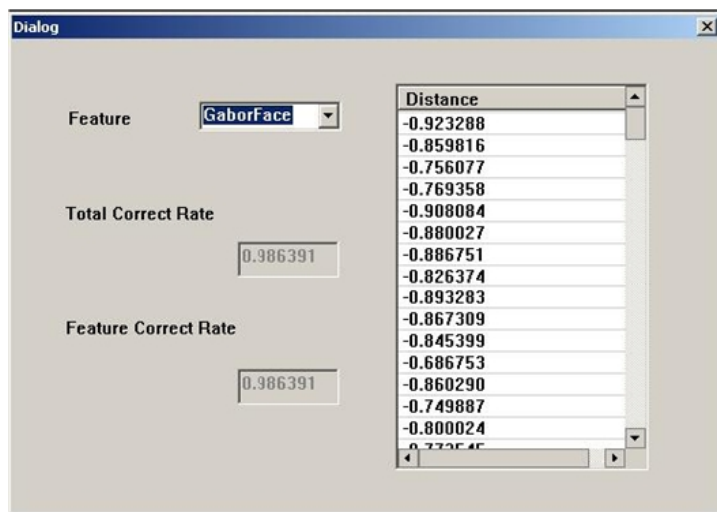


图 6.4 人脸识别准确率统计

Fig 6.4 Statistic of face recognition accuracy rate

6.2.3 系统界面

考勤系统界面见 6.5。右边小窗口为相似度的排序。



图 6.5 考勤系统界面

Fig 6.5 GUI of checking attendance system

6.3 本章小结

本章我们以办公室环境的日常考勤为应用目的，以开放式人脸识别体系框架为基础，采用 SFFS-Gabor 人脸识别方法实现基于人脸识别的考勤系统，以此为实施例，验证开放式人脸识别体系的开放性与通用性，同时也验证了 SFFS-Gabor 人脸识别方法的有效性。

第七章 总结与展望

7.1 总结

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, 本文以人脸识别的实用化, 实时化和系统化为目标, 对实用人脸识别系统中若干问题进行了初步探讨。

纵观全文的研究工作, 所取得的成果和结论主要体现在以下几个方面:

(1) 提出一种新的基于投影峰分析的眼睛快速定位方法

基于投影峰分析的眼睛快速定位方法, 充分利用了人脸五官特征的统计规律, 通过简单的运算实现眼睛中心坐标的准确快速的定位, 针对额发, 鬓角, 眉毛, 眼镜等干扰, 提出基于投影峰的定位方法, 有效地去除了上述干扰, 使得定位算法更具鲁棒性, 此外当人脸姿态发生俯仰($\pm 15^\circ$), 旋转($\pm 15^\circ$)等变化时, 同样可以实现精确定位。

(2) 改进基于 Gabor 特征的人脸识别方法

针对 Gabor 特征提取中, 二维卷积运算复杂度较高, 耗时较长的缺点, 提出了基于特征选择的 SFFS-Gabor 人脸识别方法。该方法通过 SFFS 方法通过训练集的学习, 经过两个步骤, 选择出最具分辨能力 Gabor 核的位置、尺度和方向, 以此描述人脸特征, 保证较高人脸识别率的同时, 降低了算法复杂度, 明显缩短了人脸识别的时间消耗, 使得基于 Gabor 特征描述的人脸识别方法, 更加适用于实用人脸识别系统。

(3) 改进基于 LBP 特征的人脸识别方法

针对一般 LBP 人脸描述维数较高, 不适合实时应用的缺点, 提出了压缩直方图的思想, 有效地降低了特征维数, 并保证了较高的人脸识别率。同时提出了一种新的基于全变分模型的 LBP 方法, 利用全变分模型对人脸边缘、轮廓和角点等显著特征的增强, 并结合了 LBP 算子对图像边缘、角点等局部变化特征敏感的特性, 提取对于外界因素变化特别是光照变化, 较为鲁棒的人脸特征用于人脸识别, 明显提高了光照变化情况下的人脸识别率。

(4) 构建开放式人脸识别体系框架

提出一种新的开放式人脸识别体系框架, 根据开放式人脸识别体系框架的通用性、开放性和分布式的三大设计目标, 具体描述了体系框架的设计思路和系统整体框架; 并以面向对象技术、数据库技术和 OpenCV 为技术平台, 具体实现开放式人脸识别体系框架。

(5) 实现基于人脸识别的考勤系统

以办公室环境的日常考勤为应用目的, 以开放式人脸识别体系框架为基础, 采用

SFFS-Gabor 人脸识别方法实现基于人脸识别的考勤系统，以此为实施例，验证开放式人脸识别体系的开放性与通用性，同时也验证了 SFFS-Gabor 人脸识别方法的有效性。

7.2 展望

虽然本文对实用人脸识别进行了初步的研究探索，对相关算法进行了改进，使其适于实用人脸识别系统的应用。但这些技术还不是很成熟，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。展望后续研究，我们认为还是有很多开放性的问题值得进一步开展：

(1) 鲁棒、准确的特征定位问题

人脸特征点的自动定位是人脸识别系统的必要且关键环节之一，在很大程度上影响着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，鲁棒、准确的特征定位问题值得进一步研究。

(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。提高学习算法的泛化能力还需要更加深入的研究。

(4) 光照变化问题

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。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

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) 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。

综上所述，人脸识别的实用化的研究还有诸多问题亟待解决，需要我们继续努力。

附录 A 流程图

本附录列出正文部分所述方法、步骤和过程的流程图。

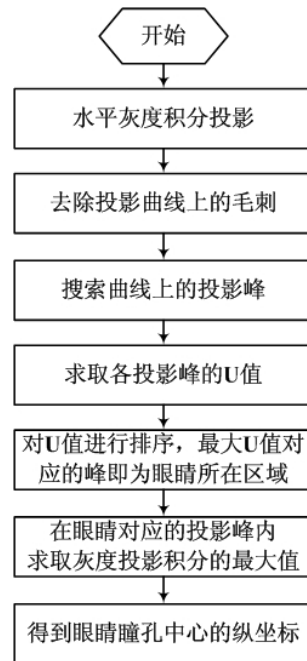


图 A.1 投影峰分析流程图

Fig A.1 Flow chart of projection peak analysis

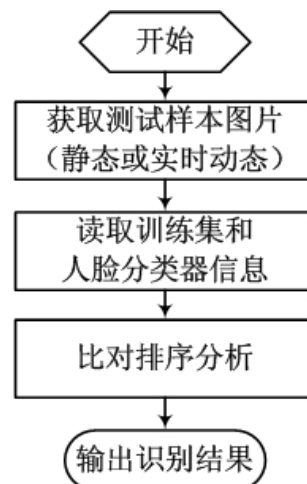


图 A.2 人脸识别过程流程图

Fig A.2 Flow chart of face recognition process

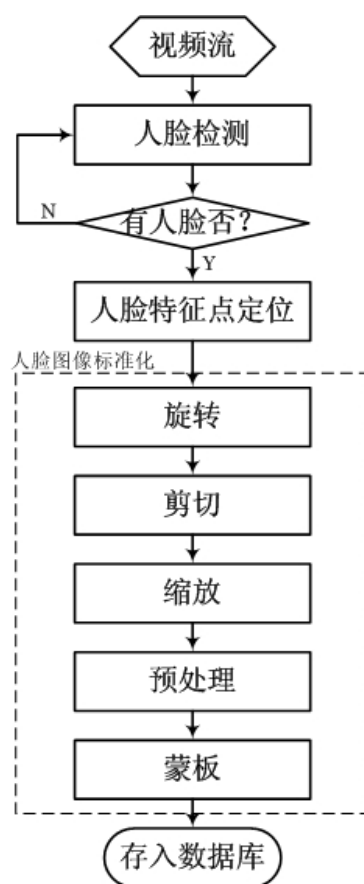


图 A.3 人脸采集过程流程图

Fig A.3 Flow chart of facial image sampling

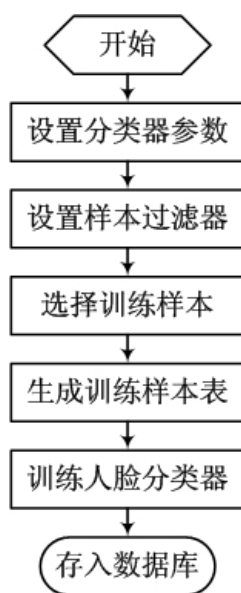


图 A.4 分类器训练流程图

Fig A.4 Flow chart of facial classifier training

附录 B 数据库表图

本附录列出了后台数据库中数据表的设计明细。

表 B.1 人员信息表

Table B.1 Table of person information

列名	数据类型	允许空	注释
<i>Person_ID</i>	bigint(8)		(PK)自动增长, 唯一标识
<i>ID</i>	varchar(50)		人员的ID, 如学号, 工号等
<i>Name</i>	varchar(50)		人员的姓名
<i>Age</i>	smallint(2)		人员的年龄
<i>Gender</i>	char(10)		人员的性别
<i>Department</i>	varchar(50)	●	人员的部门
<i>Remark</i>	text(16)	●	人员的相关说明

表 B.2 样本图像信息表

Table B.2 Table of picture information

列名	数据类型	允许空	注释
<i>Picture_ID</i>	bigint(8)		(PK)自动增长, 唯一标识
<i>Person_ID</i>	bigint(8)		(FK)与图片对应的人员ID
<i>Path</i>	text(16)		图片的保存路径
<i>Width</i>	smallint(2)		图片的宽度
<i>Height</i>	smallint(2)		图片的高度
<i>Lighting</i>	varchar(50)	●	人脸图像的光照情况
<i>Expression</i>	varchar(50)	●	人脸图像的表情
<i>Accessory</i>	varchar(50)	●	人脸图像的附件(如帽子, 眼睛等)
<i>Pose</i>	varchar(50)	●	人脸图像的姿态
<i>Time</i>	data(16)	●	样本采集的时间
<i>Remark</i>	text(16)	●	图片的相关说明

表 B.3 特征信息表

Table B.3 Table of feature information

列名	数据类型	允许空	注释
<i>Feature_ID</i>	bigint(8)		(PK,FK)自动增长, 唯一标识
<i>Name</i>	varchar(50)		特征描述的名称
<i>Remark</i>	text(16)	●	特征描述的相关说明

表 B.4 人脸特征描述表

Table B.4 Table of facial feature description

列名	数据类型	允许空	注释
<i>ID</i>	bigint(8)		(PK)自动增长, 唯一标识
<i>Feature_ID</i>	bigint(8)		特征描述的ID
<i>Person_ID</i>	bigint(8)		对应人员ID
<i>Picture_ID</i>	bigint(8)		对应样本图片ID
<i>Feature_Row</i>	bigint(8)		特征矩阵的行数
<i>Feature_Col</i>	bigint(8)		特征矩阵的列数
<i>Feature_Data</i>	image(16)		特征矩阵的数据

表 B.5 人脸分类器信息表

Table B.5 Table of facial classifier information

列名	数据类型	允许空	注释
<i>Result_ID</i>	bigint(8)		(PK,FK2)自动增长, 唯一标识
<i>Name</i>	varchar(50)		人脸分类器名称
<i>Person_Num</i>	bigint(8)		分类器对应的人员总数
<i>Samples_per_Class</i>	bigint(8)		每类的训练样本数
<i>Feature_ID</i>	bigint(8)		(FK1)特征描述的ID
<i>Param_A</i>	bigint(8)	●	分类器参数 A
<i>Param_B</i>	bigint(8)	●	分类器参数 B
<i>Mat1_Row</i>	bigint(8)	●	分类器矩阵 1 的行数
<i>Mat1_Col</i>	bigint(8)	●	分类器矩阵 1 的列数
<i>Mat1_Data</i>	bigint(8)	●	分类器矩阵 1 的数据
<i>Mat2_Row</i>	bigint(8)	●	分类器矩阵 2 的行数
<i>Mat2_Col</i>	bigint(8)	●	分类器矩阵 2 的列数
<i>Mat2_Data</i>	bigint(8)	●	分类器矩阵 2 的数据
<i>Mat3_Row</i>	bigint(8)	●	分类器矩阵 3 的行数
<i>Mat3_Col</i>	bigint(8)	●	分类器矩阵 3 的列数
<i>Mat3_Data</i>	bigint(8)	●	分类器矩阵 3 的数据
<i>Time</i>	data(16)	●	分类器训练的时间
<i>Remark</i>	text(16)	●	分类器的相关说明

表 B.6 人脸子空间特征表

Table B.6 Table of facial subspace feature

列名	数据类型	允许空	注释
<i>Picture_ID</i>	bigint(8)		(PK,FK2) 对应图片的 ID
<i>Person_ID</i>	bigint(8)		(FK1) 对应人员的 ID
<i>Data_Row</i>	bigint(8)		子空间下特征矩阵的行数
<i>Data_Col</i>	bigint(8)		子空间下特征矩阵的列数
<i>Data</i>	image(16)		子空间下特征矩阵的数据
<i>Result_ID</i>	bigint(8)		对应子空间的 ID

表 B.7 训练集合人员表

Table B.7 Table of person in training set

列名	数据类型	允许空	注释
<i>Person_ID</i>	bigint(8)		(PK,FK1) 对应人员 ID

表 B.8 训练集合样本图片表

Table B.8 Table of person in training set

列名	数据类型	允许空	注释
<i>Picture_ID</i>	bigint(8)		(PK,FK1) 对应图片 ID

参考文献

- [1] A. Pentland. Looking at people: sensing for ubiquitous and wearable computing. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(1):107 – 119, 2000.
- [2] H.Chan and W.W.Bledsoe. A man-machine facial recognition system: some preliminary results. Technical report, Panoramic Research Inc., 1965.
- [3] A.J.Goldstein, L.D.Harmon, A.B.Lesk. Identification of human faces. *Proceedings of the IEEE*, 59(5):748–760, 1971.
- [4] T. Kanade. *Picture Processing System by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto University, November 1973.
- [5] G.J. Kaufman, K.J. Breeding. The automatic recognition of human faces from profile silhouettes. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(2):113–121, 1976.
- [6] L. Harmon, W. Hunt. Automatic recognition of human face profile. *Computer Graphic and Image Process*, 6:135–156, 1977.
- [7] L.Harmon, M.K. Khan et al. Machine identification of human faces. *Pattern Recognition*, 13(2):97–110, 1981.
- [8] 山世光. 人脸识别中若干关键问题的研究. PhD thesis, 中国科学院计算技术研究所, 2004.
- [9] Thomas Vetter. Synthesis of novel views from a single face image. *International Journal of Computer Vision*, 28(2):103–116, 1998.
- [10] V. Blanz, S. Romdhani, T. Vetter. Face identification across different poses and illuminations with a 3d morphable model. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 202–207, 2002.
- [11] V. Blanz, T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 25(9):1063–1075, 2003.
- [12] A.S. Georgiades, P.N. Belhumeur, D.J. Kriegman. From few to many: Generative models for recognition under variable pose and illumination. In *Proceeding of*

- the 4th International Conference on Face and Gesture Recognition*, pages 277–284, Grenoble, France, March 2000.
- [13] R. Basri, D. Jacobs. Lambertian reflectance and linear subspaces. In *Proceeding of the the 8th IEEE International Conference on Computer Vision*, pages 383–390, Vancouver, British Columbia, Canada, July 2001.
- [14] R. Ramamoorthi. Analytic pca construction for theoretical analysis of lighting variability in images of a lambertian object. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24(10):1322–1333, 2002.
- [15] P.J. Phillips, H. Moon et al. The feret evaluation methodology for face-recognition algorithms. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.
- [16] P.J. Philips, P.J. Grother et al. Face recognition vendor test 2002: Evaluation report. Technical report, NISTIR 6965, National Institute of Standards and Technology, 2003.
- [17] G.J. Kaufman, K.J. Breeding. The automatic recognition of human faces from profile silhouettes. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(2):113–121, 1976.
- [18] J.C. Campos, A.D. Linney, J.P. Moss. The analysis of facial profiles using scale space techniques. *Pattern Recognition*, 26(6):819–824, 1993.
- [19] C. Wu, J. Huang. Human face profile recognition by computer. *Pattern Recognition*, 23:255–259, 1990.
- [20] R. Brunelli, T. Poggio. Face recognition through geometrical features. In *Proceedings of European Conference on Computer Vision*, pages 792–800, 1992.
- [21] R. Brunelli, T. Poggio. Face recognition: Features versus template. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052, 1993.
- [22] D. Beymer. Pose-invariant face recognition using real and virtual views. Technical report, AI lab, MIT, March 1996.
- [23] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

- [24] J.O. Toole, H. Abdi et al. A low-dimensional representation of faces in higher dimensions of the space. *Journal of Optical Society of America*, 10(3):405–411, 1993.
- [25] W. Zhao, R.Chellappa. A krishnaswamy discriminant analysis of principal components for face recognition. Technical report, University of Maryland, 1998.
- [26] Zi-Quan Hong. Algebraic feature extraction of image for recognition. *Pattern Recognition*, 24(3):211–219, 1991.
- [27] Y.Q. Cheng, K. Liu, J.Y. Yang. A novel feature-extraction method for image recognition based on similar discriminate function(sdf). *Pattern Recognition*, 26(1):115–125, 1993.
- [28] Y.Q. Cheng, K. Liu, J.Y. Yang. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. *Pattern Recognition*, 26(6):115–125, 1993.
- [29] D.Valentin, H.Abdi et al. Connectionist models of face processing: A survey. *Pattern Recognition*, 27(7):1209–1230, 1994.
- [30] R.Brunelli and T. Poggio. Hyperbf networks for gender classification. In *Proceedings of DARPA Image Understanding Workshop*, pages 311–314, 1992.
- [31] D. Beymer, T. Poggio. Face recognition from one example view. Technical report, AI lab, MIT, Spet. 1995.
- [32] L. Wiskott et al. Face recognition by elastic bunch graph matching. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.
- [33] P.J. Phillips, Y. Vardi. Data-driven methods in face recognition. In *Proceeding of the International Workshop on Automatic Face- and Gesture-recognition*, pages 65–69, Zurich, 1995.
- [34] Ian Craw et al. Automatic face recognition: Combining configuration and texture. In *Proceeding of the International Workshop on Automatic Face- and Gesture-recognition*, pages 53–58, Zurich, 1995.
- [35] G. Gordo. Face recognition based on depth and surface curvature. In *Proceeding of Geometric Methods in Computer Vision*, volume 1570, pages 234–237, 1991.
- [36] J.C. Campos, A.D.Linney, J.P. Moss. The analysis of facial profiles using scale space techniques. *Pattern Recognition*, 26(6):819–824, 1993.

- [37] J. Ben-Are, D. Nandy. A volumetric/iconic frequency domain representation for objects with application for pose invariant face recognition. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(5):449–457, 1998.
- [38] A. Samal, P.A. Iyengar. Automatic recognition and analysis of human faces and facial expressions: a survey. *Pattern Recognition*, 25(1):65–77, 1992.
- [39] D. Valentin et al. Connectionist models of face processing: A survey. *Pattern Recognition*, 27(9):1209–1230, 1994.
- [40] R. Chellappa, C.L. Wilson, S. Sirohey. Human and machine recognition of faces: a survey. In *Proceedings of the IEEE*, volume 83, pages 705–741, 1995.
- [41] M.A. Grudin. On internal representations in face recognition systems. *Pattern Recognition*, 33(7):1161–1177, 2000.
- [42] 张长水 李衍达周杰, 卢春雨. 人脸自动识别方法综述. 电子学报, (4), 2000.
- [43] 苏光大张翠平. 人脸识别技术综述. 中国图形图像学报, 5, 2000.
- [44] W. Zhao, R. Chellappa et al. Face recognition: A literature survey. *ACM Computing Surveys*, pages 399–458, 2003.
- [45] B. Moghaddam, A. Pentland. Probabilistic visual learning for object representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(7):696–710, 1997.
- [46] Y.Tian, T.Tan, Y.Wang. Do singular values contains adequate information for face recognition? *Pattern Recognition*, 36(3):649–655, 2003.
- [47] C. Liu and H. Wechsler. A shape and texture based enhanced fisher classifier for face recognition. *IEEE Transaction on Image Processing*, 10(4):598–608, 2001.
- [48] P.Penev, J.Atick. Local feature analysis: A general statistical theory for object representation. *Network: Computation in Neural Systems*, 7:477–500, 1996.
- [49] J. Lai, P. Yuen, G. Feng. Face recognition using holistic fourier invariant features. *Pattern Recognition*, 34(1):95–109, 2001.
- [50] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transaction on Image Processing*, 11(4):467–476, 2002.

- [51] C. Liu and H. Wechsler. Independent component analysis of gabor features for face recognition. *IEEE Transaction on Neural Networks*, 14(4):919 – 928, 2003.
- [52] A.S. Georgiades, P.N. Belhumeur, D.J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.
- [53] H. Wang, S. Li, Y. Wang. Generalized quotient image. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2004.
- [54] Y. Yang, J.O. Pedersen. A comparative study on feature selection in text categorization. In *Proc. ICML*, volume 12, page 412 – 420, 1997.
- [55] I. Guyon, A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(3):1157 – 1182, 2003.
- [56] B. Krishnapuram, A.J. Hartemink, L. Carin, M. Figueiredo. A bayesian approach to joint feature selection and classifier design. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26(9):1105 – 1111, 2004.
- [57] D. Deodhare, D. Erdogmus et al. Feature selection in mlps and svms based on maximum output information. *IEEE Transaction on Neural Networks*, 15(4):937 – 948, 2004.
- [58] H. Peng, F. Long, C. Ding. Feature selection based on mutual information: criteria of maxdependency, max-relevance, and min-redundancy. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 27(8):1226 – 1238, 2005.
- [59] P. Somol, P. Pudil, J. Kittler. Fast branch and bound algorithms for optimal feature selection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26(7):901 – 912, 2004.
- [60] F. Fleuret. Fast binary feature selection with conditinal mutual information. *Journal of Machine Learning Research*, 5(11):1531 – 1555, 2004.
- [61] I. Guyon, J. Weston, S. Barnhill, V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46:389 – 422, 2002.
- [62] K.Z. Mao. Feature subset selection for support vector machines through discriminative function pruning analysis. *IEEE Transaction on Systems, Man, and Cybernetics*, 34, 2004.

- [63] T. Evgeniou, M. Pontil, C. Papageorgiou, T. Poggio. Image representations and feature selection for multimedia database search. *IEEE Transaction on Knowledge and Data Engineering*, 15(4):911–920, 2003.
- [64] B. Heisele, T. Serre, S. Prentice, T. Poggio. Hierarchical classification and feature reduction for fast face detection with support vector machines. *Pattern Recognition*, 36:2007 – 2017, 2003.
- [65] R. Brunelli, D. Falavigna. Person identification using multiple cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(10):955–966, 1995.
- [66] A.R. Mirhosseini, H. Yan. Human face image recognition: An evidence aggregation approach. *Computer Vision and Image Understanding*, 71(2):213–230, 1998.
- [67] G. Gordon. Face recognition from frontal and profile views. In *Proceeding of the International Workshop on Automatic Face- and Gesture-recognition*, pages 47–52, Zurich, 1995.
- [68] E. Hjelm, K. L. Boon. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, 2001.
- [69] M. H. Yang, D. Kriegman, N. Ahuja. Detecting faces in images: A survey. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24(1):34–58, 2002.
- [70] M. H. Yang, N. Ahuja. *Face Detection and Gesture Recognition for Human Computer Interaction*. Kluwer Academic Publishers Group, 2001.
- [71] R. Meir, G. Raetsch. *Advanced Lectures on Machine Learning*, chapter An Introduction to Boosting and Leveraging, pages 118–183. Berlin: Springer, 2003.
- [72] Y. Freund, R. E. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771–780, 1999.
- [73] P. Viola, M. Jones. Robust real time object detection. In *Proceedings of 8th IEEE International Conference on Computer Vision*, pages 747–747, Vancouver, British Columbia, July 2001.
- [74] S. Z. Li, Z. Q. Zhang. Floatboost learning and statistical face detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 28(9):1112–1123, 2004.
- [75] R. Xiao, M. J. Li, H. J. Zhang. Robust multipose face detection in images. *IEEE Transaction on Circuits and Systems for Video Technology*, 14(1):31–41, 2004.

- [76] 林维训, 潘纲, 吴朝晖等. 脸部特征定位方法. *中国图像图形学报*, 8(8):849–859, 2003.
- [77] S. G. Shan, Y. Z. Chang, W. Gao. Curse of mis-alignment in face recognition: Problem and a novel mis-alignment learning solution. In *Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 314–320, Korea, May 2004.
- [78] Z. H. Zhou, X. Geng. Projection functions for eye detection. *Pattern Recognition*, 37(5):1049–1056, 2004.
- [79] G. H. Li, X. P. Cai, X. S. Li. An efficient face normalization algorithm based on eyes detection. In *Proceedings of the 2006 International Conference on Intelligent Robots and Systems*, pages 3843–3848, Beijing China, October 2006.
- [80] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 511–518, Hawaii USA, December 2001.
- [81] Rafael C. Gonzalez, Richard E. Wood 著, 阮秋琦, 阮宇智等译. 数字图像处理 (第二版). 电子工业出版社, 2007.
- [82] O. Jesorsky, K. Kirchberg, R. Frischholz. *Audio- and Video-Based Biometric Person Authentication*, chapter Robust Face Detection Using the Hausdorff Distance, pages 90–95. Berlin: Springer, 2001.
- [83] P. J. Phillips, H. Wechsler, J. Huang. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.
- [84] The BioID face database. <http://www.bioid.com>.
- [85] M. J. Lyons, S. Akamatsu, M. Kamachi. Coding facial expressions with gabor wavelets. In *Proceeding of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pages 200–205, Nara Japan, April 1998.
- [86] 崔锦泰[美]著, 程正兴译. 小波分析导论. 西安交通大学出版社, 1991.
- [87] R. Mehrotra. Gabor filter-based edge detection. *Pattern Recognition*, 25(12):1479 – 1494, 1992.
- [88] V. Kruger, A. Happe, G. Sommer. A fine real-time tracking using gabor wavelet networks. *ICCV' 99 Workshop Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems(RATFG-RTS' 99)*, 1999.

- [89] V. Kruger. Wavelet subspace method for real-time face tracking. *Proc. Pattern Recognition, 23rd DAGM Symposium*, 2001.
- [90] 董学智. *Gabor小波特征跟踪方法研究*. PhD thesis, 中科院长春光学精密机械与物理研究所, 2003.
- [91] S. Pittner and S. V. Kamarthi. Feature extraction from wavelet coefficients for pattern recognition tasks. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 21(1):83 – 88, 1999.
- [92] Z.Y. Zhang and Michael Lyons. Comparison between geometry-based and gabor-waveletsbased facial expression recognition using multi-layer perception. In *The third IEEE international conference on automatic face and gesture recognition*, page 454 – 459, 1998.
- [93] L. Wiskott, J. Fellous, N. Krüger, C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(7):775 – 779, 2003.
- [94] Lay David. *Linear Algebra and It's Applications*. Addison-Wesley, New York, 2000.
- [95] M. Turk and A. Pentland. Face recognition using eigenfaces. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, 1991.
- [96] A. Jain, D. Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(2):153–158, 1997.
- [97] A. Jain, R.P.W. Duin, J. Mao. Statistical pattern recognition: a review. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(1):4 – 37, 2000.
- [98] P. Pudil, J. Novovicova, J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.
- [99] T. Ahonen, A. Hadid, M. Pietikainen. Face recognition with local binary patterns. In *European Conference on Computer Vision*, pages 496–481, 2004.
- [100] Y. Rodriguez, S. Marcel. Face authentication using adapted local binary pattern histograms. In *European Conference on Computer Vision*, 2006.

- [101] T. Ojala, M. Pietikainen, T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.
- [102] B. Zhang, S. Shan, X. Chen, W. Gao. Histogram of gabor phase patterns(hgpp): A novel object representation approach for face recognition. *IEEE Transactions on Image Processing*, 16(1):57–68, 2007.
- [103] X. Tan, B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, 2007.
- [104] Bjarne Stroustrup. *The C++ Programming Language: Special Edition*. Addison-Wesley, 2000.
- [105] Bruce Eckel. *Thinking in C++, Volume One: Introduction to Standard C++*. Prentice Hall Inc., 2000.
- [106] Open Computer Vision Library. <http://sourceforge.net/projects/opencvlibrary>.

致 谢

本文是在我的导师苏剑波教授的悉心指导下完成的，从论文的选题，方案的确定，实际的设计直到论文的撰写和修改，无不凝聚着导师心血和无私的劳动，在此，谨向导师表示衷心的感谢！

苏老师治学严谨，一丝不苟，视角宽远，待人宽和，不仅是我学业上的老师，同时也是我生活中的朋友。”良师益友”，苏老师当之无愧。

苏老师老师对事业执着追求的精神、渊博的学识、敏捷的思维、谦逊的人格，无时无刻不深深地感染着我，鼓舞着我，成为我终身学习的榜样。再次向苏老师表示深深的敬意和感激！

智能机器人研究中心是一个温馨和睦的大家庭，与实验室每一位兄弟姐妹朝夕相处的日子令人难忘。感谢人脸识别组的杨飞师兄，在我刚刚接触科研的懵懂时期，杨师兄给予了我很多帮助和指导；感谢我的搭档刘丹同学，在研究工作中对我的帮助和鼓励，我们相互扶持合作愉快；感谢师弟张凯和我一起完成系统搭建。感谢张怡，曾温特，杨斌，谢谢你们分享我的快乐与烦恼，共同度过紧张充实而愉快的两年多研究生生活，祝你们一切顺利。感谢一起工作学习过的博士师兄，周玮，高振东，谢文龙，顾立忠，彭文东，邢登鹏，刘成刚，你们敏捷的思维，深厚的知识积累让我叹服，每当我遇到问题和困难时，你们总会热情地帮助我，在和你们的轻松愉快的交流讨论中，许多学术问题迎刃而解。同时还要感谢其他一起学习生活的实验室成员，钟建冬，王永雄，陈健，刘飞，龙敏，赵丽敏，余佳，刘凯，王博，顾徐鹏，和你们一起开心快乐的日子值得留念。

感谢自动化系 B0603293 班的全体同学，有缘和你们一起共度三年时光，是我的荣幸。尤其感谢我的室友武江涛同学，在我迷茫失落的时候，总有你的关心与帮助，深深的祝福送给你。

仅以此文献给我最亲爱的父亲和母亲，以答谢他们对我的养育之恩。他们无私而不求回报的爱永远激励着我勇于面对任何困难和挑战。

最后，向所有关心、爱护和帮助过我的人致以最衷心的感谢！

攻读学位期间发表的论文与成果

发表的论文:

1. J. Dai, D. Liu and J. Su, "Projection Peak Analysis for Rapid Eye Localization," Accepted by the International Conference on Computer Vision Theory and Applications, Feb 5-8, Lisboa, Portugal, 2009. **(VISAPP09)**
2. 戴景文, "基于全变分模型的 LBP 人脸识别方法", 微型电脑应用, 已录用
3. 戴景文, 刘丹, 苏剑波, "基于投影峰的眼睛快速定位方法", 模式识别与人工智能, 已投

申请的发明专利:

苏剑波, 戴景文, "开放式人脸识别体系框架", 待公开

申请的软件著作权:

苏剑波, 戴景文, 张凯, "SJTU RCIR 人脸识别考勤系统"