

分类号_____密级_____

UDC^{注1}_____

学 位 论 文

基于扩展 LBP 特征的人脸识别系统研究

(题名和副题名)

蒋思洋

(作者姓名)

指导教师姓名_____解 梅 教 授

电子科技大学 成 都

(职务、职称、学位、单位名称及地址)

申请专业学位级别_____硕士_____专业名称_____信号与信息处理

论文提交日期_____2012-5-21_____论文答辩日期_____2012-5-25_____

学位授予单位和日期_____电子科技大学

答辩委员会主席_____

评阅人_____

2012 年 月 日

注 1：注明《国际十进分类法 UDC》的类号。

独 创 性 声 明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

签名：_____ 日期： 年 月 日

论 文 使 用 授 权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将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、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

签名：_____ 导师签名：_____

日期： 年 月 日

摘要

人脸作为一种具有独特特征的生物特征，具有直接性、唯一性、方便采集等特点，但是又由于人脸的可塑多变和在采集过程中受诸多外部因素的影响，使得人脸的机器识别这一课题非常具有挑战性。现有算法的显著不足在于对环境的变化鲁棒性较差，尤其是与人脸识别的能力相比而言，自动人脸识别对环境和姿态的变化适应性能低，现有的方法大多数情况下是要求待检测人脸在规定环境下采集，这显然不满足实用性的要求。本文对经典的人脸识别算法进行了研究和改进，在预处理，特征采集，特征约简几个方面做了相应的研究，并且初步实现了一个基于视频的人脸识别系统原型。本论文所做的主要工作如下：

1. 研究了 LBP 算子的几种扩展形式，包括统一模式，旋转不变的统一模式和改进为三值形式的 LTP 模式。通过研究，可以看到 LBP 算子所具有旋转不变和灰度鲁棒这两个特性对于人脸识别来说很重要，同时 LBP 也具有描述图像局部特征和全局特征的能力，这对于人脸识别来说是也非常有利的；本文还将 LBP 的三值扩展形式 LTP 模式运用到人脸识别中来，其相较于传统 LBP 模式对于局部的光照变化和噪声有更高的鲁棒性。

2. 深入研究了 LBP 原始模式的降维算法。因为 LBP 统一模式本身所存在的一些缺陷，所以我们用降维生成的 LBP 子模式来代替 LBP 统一模式；本文比较了局部保持投影（LPP）、主成份分析（PCA）和线性判别分析（LDA）这三种降维算法在人脸领域的应用，最终选取的是 LPP 模式在构建 LBP 子模式。

3. 研究了人脸图像的预处理方法，特别是人脸校正。本论文方法是：先通过平均精确合成滤波器（ASEF）取得人眼的位置然后再通过仿射变化对人脸进行旋转，最后得到近似标准的正面人脸。ASEF 定为人眼的精度较 Haar-like+Adaboost 方法高很多，并且其对于戴了眼镜或者有遮挡的人脸的鲁棒性也更强，最重要的是其速度非常快，几乎在检测到人脸的同时就能定位人眼。

4. 基于 OpenCV 开源库构造了一个基于视频的人脸识别系统原型。并在上面对于算法进行实际的测试。

关键词：人脸识别，局部二值模型，局部保持投影，平均精确合成滤波器

ABSTRACT

Face recognition has been a hot research problem of pattern recognition and machine vision. The face is a kind of biological characteristics, directly, uniqueness, convenient acquisition, but because of the fluidity of the face and vulnerable to external factors, makes the subject of face recognition is very challenging. Existing algorithms is significantly poor robustness to environmental changes, especially compared with the face recognition capabilities, automatic face recognition performance on the environment and attitude change adaptation is low, the majority of the existing methods case is to be detected face collection in the specified environment, which obviously does not meet the practical requirements. This paper studies and improve the classic face recognition algorithm, including image preprocessing, characteristic collection, feature reduces, etc, and the initial realization of a video-based face recognition system prototype. The main work is as follows:

1. Research LBP operator of several extended form, including the unified model, the rotation invariant unified model and improve the three-valued form of LTP model. Through research, we can see the LBP operator has two features: rotation invariant and grayscale robust. It is very important to face recognition, LBP also has the ability to describe the local and global features of the image, This is also very beneficial to face recognition; the LBP value of this article will be an extended form of LTP mode is applied to face recognition, compared to the traditional LBP model for the local illumination variations and noise, a higher level of robustness.

2. In-depth study of the dimensionality reduction algorithm of the original model of LBP. Some of the defects that exist because of the the LBP unified model itself, so we use the LBP sub-model generated by the dimensionality reduction instead of the LBP unified model; This paper compares three dimensionality reduction algorithm (locality preserving projection (LPP), principal component analysis (PCA) and linear discriminant analysis (LDA)) in the field of face, and ultimately selected the LPP mode in building the LBP sub-mode.

3. Research face image preprocessing methods, especially the face

ABSTRACT

correction. This paper show a method: first through the average of synthetic exact filter (ASEF) to obtain the location of the human eye then change the face rotation by affine transformation, and finally to the frontal face of the approximate standard. ASEF given the precision of the human eye is much higher than the Haar-like+Adaboost method, and it's more robustness for wearing glasses or the block face, the most important is its very fast, almost to the detection face at the same time be able to locate the human eye.

4. Open source library based on OpenCV Constructed a prototype of a video-based face recognition system. In the above algorithm, the actual test.

Keywords: Face recognition, Local Binary Pattern, Locality Preserving Projection, Average of Synthetic Exact Filters

目 录

目 录	IV
第一章 绪论	1
1.1 课题研究的背景和意义	1
1.2 现阶段的人脸识别技术	2
1.3 本文研究的主要内容及全文结构安排	3
1.3.1 研究的主要方向和内容	3
1.3.2 本文的结构安排	4
第二章 基本 LBP 算子的原理及应用	6
2.1 原始 LBP 特征概述	6
2.2 LBP 算子的扩展	8
2.2.1 LBP 统一模式	8
2.2.2 旋转不变的 LBP 算子	9
2.3 LTP 特征概述	11
2.4 基于相似度测量的距离转换	12
2.5 本章小结	14
第三章 降维生成 LBP 子模式	16
3.1 LBP 统一模式所存在的问题	16
3.2 维度约减方法概述	17
3.2.1 主成份分析算法介绍	18
3.2.2 线性判别分析算法介绍	19
3.2.3 局部保持投影算法概述	20
3.2.4 LPP 的统计特性	21
3.3 LPP, PCA, LDA 的理论分析	22

3.3.1 局部保持投影 (LPP) 与主成份分析 (PCA) 的关系	22
3.3.2 局部保持投影 (LPP) 和线性判别分析 (LDA) 的关系	23
3.4 基于 LPP 降维算法构建 LBP 子模式	26
3.5 LBP 统一模式与子模式的比较	26
3.6 实验结果与分析	27
3.7 本章小结	28
第四章 图像预处理及人眼检测	29
4.1 抑制图像噪声	29
4.2 人脸图像直方图均衡	30
4.3 人脸图像校正	31
4.3.1 人眼定位简介	31
4.3.2 眼睛的定位	32
4.3.3 ASEF 相关滤波器	33
4.3.4 ASEF 滤波器定义	34
4.4 基于 ASEF 的人脸校正	36
4.4.1 基于 ASEF 的人眼检测	36
4.4.2 人脸校正	37
4.4.3 ASEF 的特点	38
4.5 本章小结	40
第五章 基于级联分类器的训练算法	41
5.1 分类器结构	41
5.1.1 Boosting 原理	41
5.1.2 分类器级联	42
5.2 弱分类器设计	43
5.2.1 矩形特征选取	43
5.2.2 LBP 子模式特征提取	44
5.3 Adaboost 构建强分类器	45

5.4 本章小结	47
第六章 基于 LBP 和 Adaboost 的人脸系统	48
6.1 算法简介	49
6.2 实验结果与分析	53
6.2.1 ORL 人脸数据库算法结果	54
6.2.2 FERET 数据库算法结果	54
6.2.3 CAS-PEAL 数据库算法结果	55
6.3 基于 OpenCV 的系统原型实现	56
6.4 本章小结	57
第七章 总结与展望	58
7.1 全文总结	58
7.2 后续工作展望	59
致 谢	60
参考文献	61
攻读硕士期间取得的研究成果	66

第一章 绪论

1.1 课题研究的背景和意义

对他人身份和辨识是我们日常生活中很重要的一个问题。传统的认证手段，如钥匙和证件等都需要随身携带，一旦丢失就会非常不便，同时这些都可以伪造。而密码也有其固有的缺陷。用密码进行身份验证最大的问题就是密码的长度和安全性之间的平衡，密码越长越复杂，也就越安全，但也越不便于记忆。生物特征识别技术作为 21 世纪对人类社会带来革命性影响的十大技术之一，是目前最简便、有效的身份识别技术，它对人的本身进行认定，不需要身外的其他标识物。目前主要的生物识别技术包括指纹、虹膜等生物特征的认证，但是这些技术都或多或少地需要用户有主动的验证意愿。基于人脸固有特征（包括几何和纹理特征）的机器人脸识别技术可以不接触人脸就能进行识别，因而具有对待识别用户干扰低的特点，被认为是最有发展前景的自动识别技术^[1]。

生物特征（biometrics）包括生理特征和行为特征，生理特征主要有指纹、掌纹、面容、静脉纹理、虹膜、视网膜、DNA 等，这些特征基本是属于先天拥有的。行为特征主要有签名、声音、步态等，这些特征是人类后天养成的。在实际应用中，作为表示身份的生物特征必须具备身份标识的特点，生物特征识别系统要求所用生物特征具有广泛性、唯一性、稳定性、可采集性外，还存在以下附加要求：1、性能：识别的速度、鲁棒性，系统运行效率；2、可接受性：是否易于接受，不侵犯个人隐私；3、安全性：是否能够保证该应用场合的安全性要求；4、成本：硬件投入、系统架设及数据录入成本是否在可接受范围。

使用得最早、最广泛的生物特征识别技术是基于指纹的。每个人的指纹都是唯一的，且不会随着时间的变化而改变，并且指纹的生物学特征易于采集，与一般的图像相比，指纹图像有非常独特的特征，而且其特征的形式比较简单^[2]。

虹膜是人眼内部的一个器官，处于角膜的保护之下，不易受到外界的伤害，从外部看去是黑色瞳孔和白色巩膜之间的有色圆环状部分。虹膜识别被认为是最安全、最精确的识别方法，它利用摄像机捕获用户眼睛的图像，从中解析并分离出虹膜图像，并对图像中所包含的特征信息进行提取编码，用于匹配。迄今为止，虹膜识别系统的识别率仍然位居生物特征识别系统首位。但虹膜因位于人眼内部，

采集过程中需要用户配合，而且难免会受到眼睑及睫毛的遮挡，不易准确捕获虹膜图像。同时虹膜图像的采集需要专门的设备，构造复杂成本昂贵，一定程度上限制了虹膜识别的应用范围。

人脸识别是近年来非常活跃的一个研究领域，发展十分迅速，用人脸进行身份识别符合人类的直觉，也是人类本身所非常习惯的形式。人脸识别包括静态图像识别和视频图像识别，前者通常用于照片处理，后者进行实时跟踪检测。

检测和识别是人脸识别体统的两个主要组成部分，其中只有检测到了人脸才能进一步地去识别人脸。在实际应用中，人脸采集时光照、角度、距离、遮挡物等的环境复杂性，以及表情、姿态的变化，使得人脸识别系统目前还只能在小样本数据库中运用，并且其识别率还远低于指纹及虹膜识别。与此同时，肤色、胡须、化妆等因素，加上不可避免的双胞胎识别问题，使得人脸识别技术变得更富有挑战性。

表 1-1 常用生物特征识别技术的比较^[2,3]

类型	是否接触	采集成本	采集速度	实现难易	复制难易	可靠性	应用性
人脸	非接触	高	最快	最难	不易	高	最广泛
指纹	接触	低	快	较难	可能	高	广泛
虹膜	非接触	高	较快	难	难	最高	较广泛
视网膜	非接触	最高	较慢	难	难	最高	不广泛
手型	接触	低	快	较难	可能	高	不广泛
签名	接触	较低	慢	易	可能	低	广泛
语音	非接触	低	快	易	可能	低	较广泛

1.2 现阶段的人脸识别技术

对于人脸识别技术的研究很早之前就已经出现了，近年来许多创新的技术和方法被应用在了人脸识别领域，尤其是上世纪九十年代以来，随着对于生物识别的需求越来越多，要求越来越高，人脸识别技术得到广泛的关注和研究探讨^[1]。如今，人脸识别的研究已经成为各科研机构和技术公司的主要项目之一。

人脸识别系统一般主要包括人脸检测、人脸跟踪、人脸特征提取、人脸识别这几个方面。人脸识别系统的整个框架如图 1-3 所示。

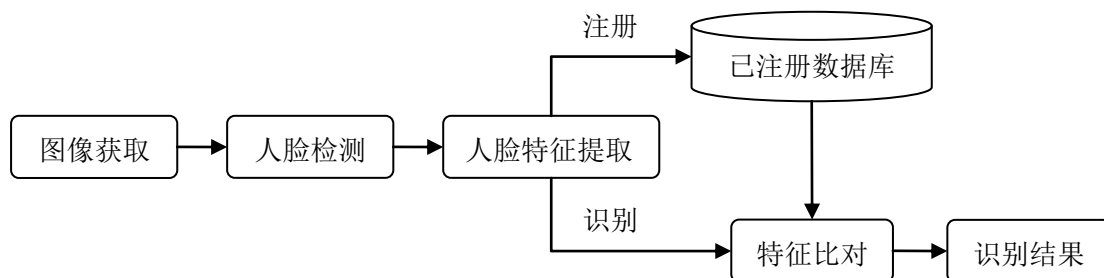


图 1-1 人脸识别系统大致流程

早期的人脸识别算法主要是对于人脸的外部轮廓特征或者通过训练样本形成的统一的人脸模版来进行匹配。现在主流的人脸识别研究方向分为基于局部特征的方法、基于主元分析的特征脸（Eigenface）、用人工神经网络构建分类器的方法和基于隐马尔科夫链（HMM）的方法。

能不能提取到人脸特征，提取的特征是否符合要求，这些都是人脸特征提取所必须要面对的问题，而特征提取又是识别的非常关键的一步，所以怎样提取人脸特征值得我们去好好研究。根据特征提取的方法、特征形式的不同，可以运用以下几种特征来进行人脸识别：局部特征，全局特征，混合特征（即混合全局和局部特征）。

总得说来，现阶段存在有很多人识别的方法，但每一种方法都不能很好的对每一种情况进行适应，所以就要各种方法进行融合和改进。

1.3 本文研究的主要内容及全文结构安排

1.3.1 研究的主要方向和内容

人脸系统各部分（预处理、检测、定位、识别）的研究虽然已经进行了很长的时间，但是由于现阶段的人脸识别系统对于外部光照条件、人脸表情、头部姿势都有比较严苛的要求，以及文本本身所具有的复杂性，所以还很难找到一个很完备的方法和很有效的单一特征来鉴别人脸。再则，随着人脸数据库的扩大，人脸数据的储存也是很大的一个难题，因此，怎样提取结构比较简答简单、能有效描述人脸的特征也成为了目前人脸识别邻域的研究重点。

Ojala 于 1996 年首次提出了局部二值模式（Local Binary Pattern, LBP），其算子能够准确且快速的描述图像纹理，其扩展模式对于图像旋转和光照变化适应性都很强^[22]。Ahonen 等首先将 LBP 算子运用到人脸识别技术当中来。本文重点

研究方向是将 LBP 及其扩展特征运用到现有人脸检测和人脸识别算法中，主要目标是找到一种切实有效的扩展 LBP 特征，其对于光照的变化、人脸表情、姿态的变化有很高的鲁棒性，且能够满足视频流中人脸识别的实时性。具体工作如下：

1. 研究了一种新的 LBP 子模式算法。对 LBP 原始模式的降维算法进行了研究，包括传统的 PCA^[25]、LDA^[26]和基于流形理论的 LPP (Locality Preserving Projections, 局部保持投影)^[27]，分析了 LBP 统一模式的缺点，并结合 LPP 算法进行降维，开拓了一种新的思路。

2. 提出将局部三值模式 (LTP, Local Ternary Pattern)^[28]运用到人脸识别中。LBP 特征对于全局的光照变化或者噪声是具有一定的鲁棒性的，但是对于人脸局部的光照变化和噪声，其性能就十分低下，LTP 特征引入了阈值的概念，其对于局部区域的噪声更不敏感。

3. 研究了多分辨率图像的多尺度 LBP 算法。文中将图像金字塔的概念融合到 LBP 特征提取中，用图像分辨率的缩放来代替 LBP 算子本身的缩放。其与以前的方法相比能够提取到更加有效且丰富的人脸特征，同时计算量也较小。

4. 对人脸识别前的图像预处理进行了研究，应用平均精确合成滤波器 (ASEF) 找到人眼，并通过人眼的坐标来对人脸进行了校正。

5. 用多尺度 LBP 子模式特征作为基础构建出了人脸识别系统。该算法在各个主流数据库中都能得到很好的测试成绩，说明其对人脸位置的偏移和环境光照的变化都具有较高的鲁棒性。

1.3.2 本文的结构安排

本文共七章，第一章为全篇的绪论，第二到第六章详细的介绍了本文所做的研究工作，第七章是对本为所做工作的总结与未来需要做的工作的展望。

第一章概述了人脸识别技术的历史背景，现阶段发展形势以及本文所做工作的一些简介。

第二章详细介绍了 LBP 算子以及其发展过程，且将其在模式识别和图像处理领域中的应用做了相关概述。针对于 LBP 模式的一些固有缺陷，提出了 LTP 模式来取得较好的光照和噪声鲁棒性。

第三章研究了 LBP 子模式特征。首先简要介绍了 LBP 统一模式，及其所存在的一些缺陷；然后简介了几种主流的特征提取算子；最后结合 LBP 模式和 LPP 算子得出 LBP 子模式，其大大减少了原始模式的数量，且有效模式在新的投影空

间中还更加的密集。

第四章研究了图像预处理技术。主要应用 ASEF 对人脸进行校正。

第五章介绍了经典的基于级联分类器的人脸图像训练算法。

第六章对于基于本文算法的人脸识别系统进行了系统介绍，并与主流算法在数据库中进行了测试和分析。

第七章总结了现阶段已完成的研究工作，并对未来的工作做了一些规划。

第二章 基本 LBP 算子的原理及应用

局部二值模式^[22](Local Binary Pattern, LBP)是一种纹理描述算子,其模型简单,但是对于纹理的描述十分有效,其可以提取出灰度图像中局部邻域的纹理信息,按照规则生成的灰度直方图特征也被证明是一种非常有用的高阶统计信息,其能够适应不同的旋转与光照。LBP 最开始是用于纹理分析方面,现在在图像检索,纹理分割、分类,人脸识别等领域也都得到了很好的运用。

2.1 原始 LBP 特征概述

局部二值模式是运用于灰度图像中的一种纹理描述算子,为了更好地度量图像的局部对比度,Ojala 等人提出了最初的局部二值模式 (LBP),最原始的 LBP 算子被定义为一个 3X3 的窗口,以窗口的最中心像素的灰度值作为门限值,将其与相邻的 8 个像素进行比较,比中心像素大的像素被标记为 1,反之则记为 0,这样就可以按照一定的规则计算出该窗口的 LBP 值^[22]。如图 2.1 所示:

example	thresholded	weights																											
<table border="1"> <tr><td>6</td><td>5</td><td>2</td></tr> <tr><td>7</td><td>6</td><td>1</td></tr> <tr><td>9</td><td>8</td><td>7</td></tr> </table>	6	5	2	7	6	1	9	8	7	<table border="1"> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td></td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	1	0	0	1		0	1	1	1	<table border="1"> <tr><td>1</td><td>2</td><td>4</td></tr> <tr><td>128</td><td></td><td>8</td></tr> <tr><td>64</td><td>32</td><td>16</td></tr> </table>	1	2	4	128		8	64	32	16
6	5	2																											
7	6	1																											
9	8	7																											
1	0	0																											
1		0																											
1	1	1																											
1	2	4																											
128		8																											
64	32	16																											
Pattern = 11110001 LBP = 128 + 64 + 32 + 16 + 0 + 0 + 0 + 1 = 241																													

图 2.1 原始的 LBP 算子

原始 LBP 算子被固定在一个 3×3 区间范围内的,其覆盖面积明显较小,这显然不能满足于不同尺度纹理的要求。为了适应和提取不同尺度的纹理特征,并且达到旋转不变和灰度不变的目标,Ojala 等改进了 LBP 算子,将最初的 3X3 邻域扩展到了任意邻域,抛弃了原本的正方形邻域,用圆形的邻域来替代,改进后的 LBP 模式在半径 R 的圆形邻域里允许有任意个取值点^[24]。

为了方便,我们假设局部区域的中心像素点为 g_c ,对应的圆形邻域窗口内的 N 个像素分别为 g_0, \dots, g_{N-1} ,我们用局部区域各个像素的联合密度分布函数来定义这个区域内的图像纹理,其定义如下:

$$T = t(g_c, g_0, \dots, g_{N-1}) \quad (2-1)$$

假设中心像素 g_c 与差值 $g_n - g_c (n = 0, 1, \dots, N-1)$ 无关, 那么上式就可以改写为:

$$T \approx t(g_c)(g_0 - g_c, \dots, g_{N-1} - g_c) \quad (2-2)$$

算子 $t(g_c)$ 描述的是整幅图像的明亮分布情况, 而与图像的局部纹理没有关系, 因此可以将其忽略, 从而得出式 (2-3):

$$T \approx (g_0 - g_c, \dots, g_{N-1} - g_c) \quad (2-3)$$

将周围像素与中心像素做比较, 如果大于中心像素赋值为 1, 反之则赋予 0, 如公式 (2-4) 所示:

$$T \approx t(s(g_0 - g_c), \dots, s(g_{N-1} - g_c)) \quad (2-4)$$

其中 $s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$ 。通过上式可以得到一个 N 位的二进制数, 再通过加权求和

就可以得到一个唯一的值, 这个值就是 LBP 的模式。这个值描述了以 (x_c, y_c) 为中心的一个局部区域的纹理特征, 可以用公式 (2-5) 来表示:

$$LBP(x_c, y_c) = \sum_{n=0}^{N-1} s(g_n - g_c) 2^n \quad (2-5)$$

根据上述 LBP 算子所展示出来的特性我们可以看出, LBP 算子对于单调的灰度变化 (人脸中就是人脸所处环境的光照变化) 具有鲁棒性, 只要各像素不单独发生改变, 那么得到的 LBP 值就不会改变。假设在半径为 R 的圆形邻域内有 N 个像素点 $g_n (n = 0, \dots, N-1)$, 其 LBP 算子用 LBP_N^R 表示, 常见的有:

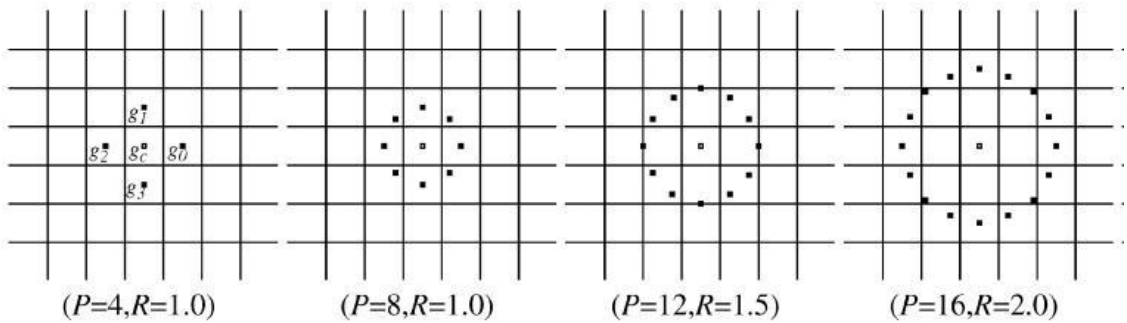


图 2-2 几种常用 LBP

2.2 LBP 算子的扩展

2.2.1 LBP 统一模式

从 LBP 算子的定义可以发现，一个 LBP 算子的二进制模式数量与其采样点数成正比，比如，对于 LBP_N^R 就有 2^N 种二进制模式。我们实际需要的算子是简单、快速、空间占用小，但是原始 LBP 模式产生的此多的二进制模式对于纹理特征的处理是非常不利的。我们在进行人脸识别时，常常采用的是 LBP 算子的统计直方图来表达图像的全局信息，如果 LBP 算子数量过多，不仅将使得数据量庞大，而且也会使直方图过于稀疏，不利于特征比对。因此，要得到既精又好二进制模式，就需要对原始的 LBP 模式进行维度的约简。

Ojala 等提出了一种叫做“统一模式”（Uniform Pattern）的方法来对原始的 LBP 算子的模式种类进行降维，其认为在实际的图像中，大部分的 LBP 算子最多只包含两次从 0 到 1 或者 1 到 0 的跳变，因此，Ojala 将“统一模式”定义为：当某个局部二进制模式所对应的循环二进制数从 0 到 1 或者 1 到 0 最多有两次跳变，那么该局部二进制数就被称做是一个统一模式类^[24]。比如 11111111，00000000，10001111 都是统一模式类。是否是统一模式类可以由式（2-6）进行检验：

$$U(G_n) = |s(g_{N-1} - g_c) - s(g_0 - g_c)| + \sum_{n=1}^{N-1} |s(g_n - g_c) - s(g_{n-1} - g_c)| \quad (2-6)$$

如果某种模式计算得到的 $U(G_n)$ 小于或者等于 2，则其就是统一模式。还有一类模式称为混合模式，其是除开统一模式类之外的所有模式。通过生成统一模式，LBP 算子的二进制模式种类大大减少，且几乎保留了全部的信息，模式数量由原来的 2^N 减少为 $N(N-1)+2$ ，N 代表邻域内的采样点数。Ojala 认为，统一模式也能很好的表达了包括图像边缘信息在内的大多数关键模式，如图 2-3 所示，白色表示 1，黑色表示 0。从图中我们可以看出，统一模式类在所有的模式中占了大多数，所以使用混合直方图统计（包括统一模式和混合模式）能够更加有效的提取出人脸的纹理特征。

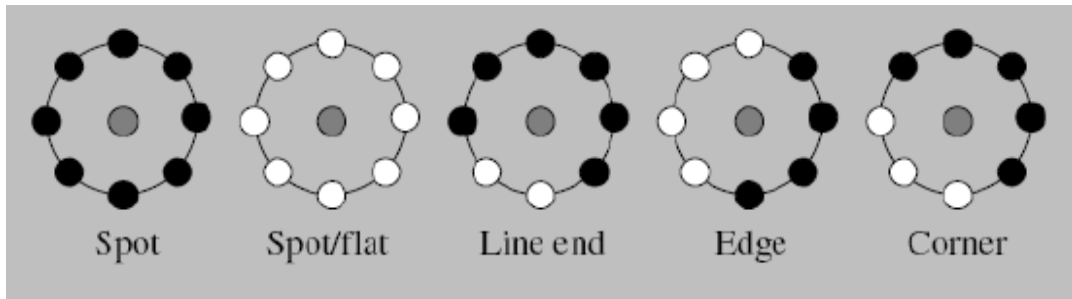


图 2-3 LBP 模式所检测的基本纹理单元

以 LBP_8^1 为例，其所有的二进制模式中统一模式占 23%。但是实验表明，对于一般图像而言，统一模式在所有模式中占主导作用。如图 2-4，在图 (a)，(b)，(c) 中，统一模式所占的比重分别是 86%、92%、78%。Ojala 等^[24] 也指出可以再图像上叠加一层低通滤波器来增加统一模式所占的比重。

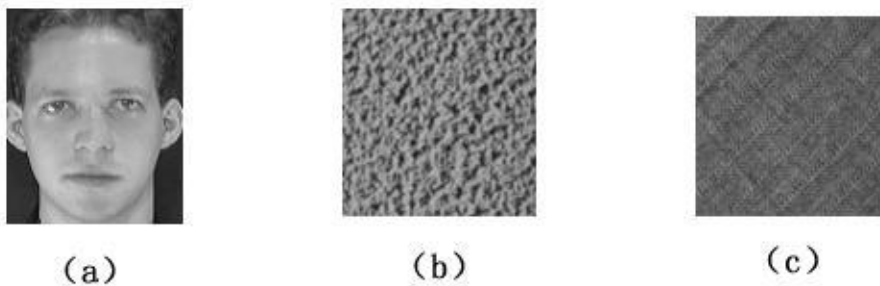


图 2-4 不同类型图像

2.2.2 旋转不变的 LBP 算子

从 LBP 算子的定义我们可以看出，图像整体的灰度改变其 LBP 算子是不变的，但是当图像旋转时其 LBP 值也会随之改变。只要 $s(g_i - g_c)$ 不全为 0 或者 1，那么旋转后的图像（旋转度数不是 360 的整数倍）就肯定会得到不同的 LBP 值，由图 2-5 我们也可以清楚的看到这一点。因此，Maenpaa 等人将 LBP 特征进行了扩展，给 LBP 算子带来了旋转不变的特性，不断旋转的圆形邻域可以得到的很多不同的 LBP 值，而取其中的最小值做为代表该邻域的 LBP 值^[29]，可用如下公式表达：

$$LBP_{N,R}^{ri} = \min(ROR(LBP_{N,R}^{ri}, i) | i = 0, 1, \dots, N-1) \quad (2-7)$$

旋转不变的 LBP 算子用 LBP^{ri} 表示， $ROR(x, y)$ 代表旋转函数，表示将 x 循环右移 i 位 ($i < N$)。图 2.5 给出了求取旋转不变的 LBP 的过程示意图，图中所示的

LBP 算子中的最小值是 15，按照旋转不变 LBP 算子的定义，此局部纹理得到的具有旋转不变性的 LBP 值为 15，即 00001111。那么，就用 00001111 来代表以下 8 种 LBP 模式的二进制编码^[29]。

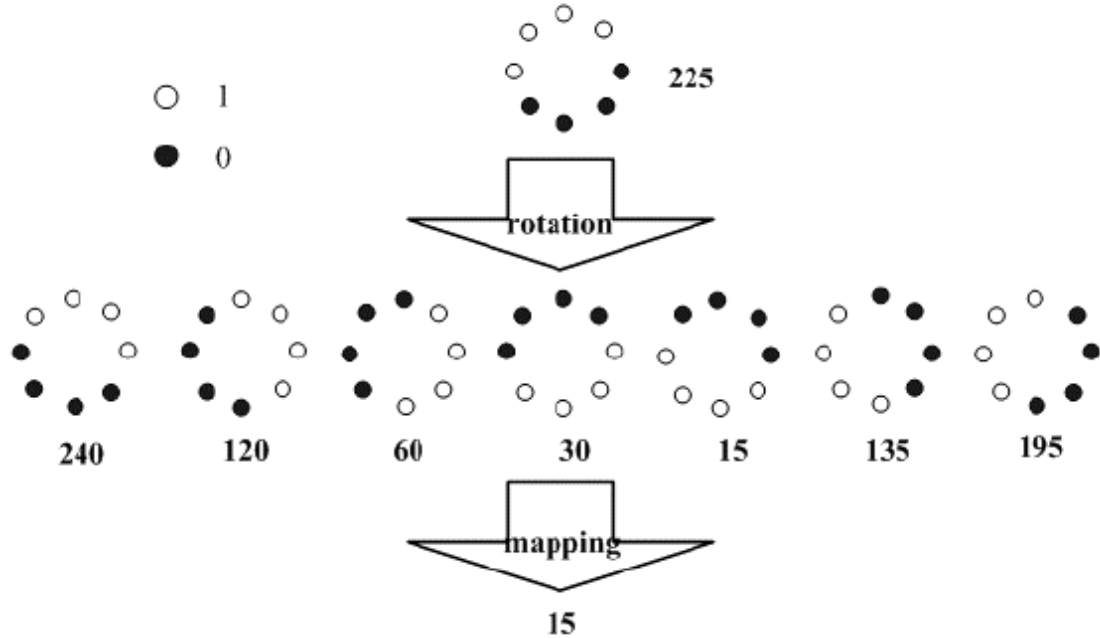


图 2-5 旋转不变的 LBP 算子

旋转不变的引入，使得 LBP 算子不仅对于图像的偏转有了更强的鲁棒性，同时也使 LBP 算子的种类进一步的减少。当然，丢失了方向信息使得旋转不变的 LBP 算子不再适用于所有场合。不过，对于具有相似纹理图像的分析， LBP^r 算子是很有效的。

旋转不变统一模式将可能的模式数量从 2^N 减少到 $N+1$ 。如公式 (2-8) 所示：

$$LBP_{N,R}^{riu2} = \begin{cases} \sum_{n=0}^{N-1} s(g_n - g_c) & U(G_n) \leq 2 \\ N+1 & U(G_n) > 2 \end{cases} \quad (2-8)$$

其中， $U(G_n)$ 表示 1 到 0 或者是 0 到 1 的跳变次数，具体定义在公式 (2-6)， LBP^{riu2} 被称做旋转不变的模式。通过统一模式和旋转不变的扩展过后，LBP 算子不仅具有了旋转不变和灰度鲁棒的特点之外，原始 LBP 算子的模式种类也大大的减少。表 2-1 展示了三种 LBP 模式在经过扩展前后的模式数量比较：

表 2-1 三种 LBP 算子扩展前后维度比较

	原始模式数	统一模式	旋转不变统一模式
LBP_8^1	256	58+1	9
LBP_{16}^2	65536	242+1	17
LBP_{24}^3	16777216	554	25

2.3 LTP 特征概述

根据 LBP 编码公式中的 $g_n - g_c$ ，我们可以看出 LBP 算子在灰度范围内是单调不变的，这使得它能适应不同的光照变化，再则，LBP 算子能够很有效的描述图像的纹理信息，是因为其并不是从一个单一的像素点来描述纹理，而是通过一个局部区域的模式来表示纹理，每一个像素点都是通过与其匹配的局部邻域来描述的^[22]。这些优点使得 LBP 特征能够很好的区分不同图像的纹理特征。但是提取中心像素 g_c 时，在统一模式图像区域，中心像素周围的噪声会有较大的影响，既然图像的绝大多数区域是统一模式区域，那么要提高 LBP 算子在局部的鲁棒性，以及提高特征空间的分类能力就必须的抑制噪声对于纹理算子的影响。

Tan 等人^[28]将 LBP 模式在一次进行扩展，将 LBP 模式扩展为三值的编码，提出了 LTP (Local Ternary Pattern 局部三值模式) 模式。LTP 模式不仅延续了 LBP 模式的主要优点而且对于局部区域的噪声更加的鲁棒。

LTP 的计算方法是将与中心像素 g_c 的像素差值在 $\pm t$ 之内的邻域像素量化为 0，除此之外大于 g_c 的量化为 1，小于 g_c 的量化为 -1，此时 $s(x)$ 被扩展成了三值形式，如公式 2-9 所示：

$$s'(g_n, g_c, t) = \begin{cases} 1, & g_n \geq g_c + t \\ 0, & |g_n - g_c| < t \\ -1, & g_n \leq g_c - t \end{cases} \quad (2-9)$$

由此，二值的 LBP 算子由三值的 LTP 算子所代替，式中 t 是自定义的一个阈值。正是因为阈值 t 的存在，使得 LTP 算子对于局部噪声更加的鲁棒，但是也使得 LTP 算子不再像 LBP 算子一样具有严格的灰度不变性。LTP 算子的编码过程如图 2-6 所示，将阈值 t 设置为 5：

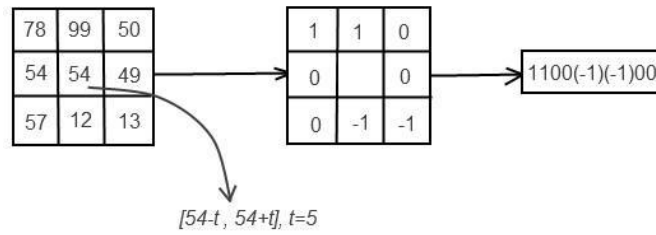


图 2-6 LTP 算子编码过程示意图

虽然在使用 LTP 模式进行人脸识别时我们使用的是三值编码，但是统一模式在此同样适用。为了计算简单并且运用现有 LBP 算子的一些研究成果，我们将 LTP 算子的分为正负两个部分，（如图 2-7 所示）将其看作两个独立通道的 LBP 算子对待，并且分别计算其分割后的灰度直方图和相似度矩阵，最后才将这两个部分结合在一起作为一个特征来对待。

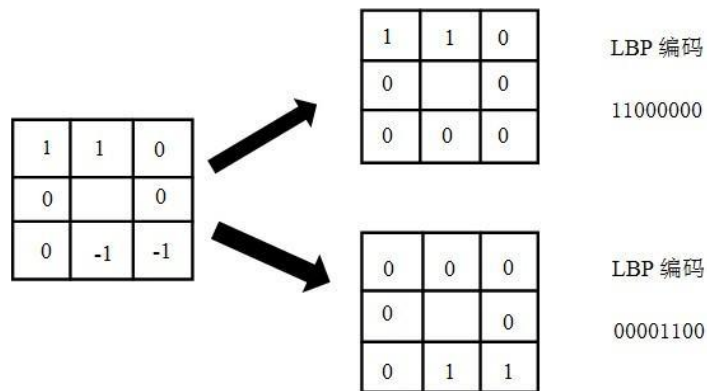


图 2-7 LTP 算子分解为正负两个 LBP 算子过程

LTP 算子有些类似于 1990 年代的纹理图谱技术^[30]，只是纹理图谱不包括图像预处理，阈值化，局部灰度直方图和基于降维的统一模式，而且其也不适用于人脸图像。

2.4 基于相似度测量的距离转换

T. Ahonen^[31]等提出了基于 LBP 的人脸识别算法，其将人脸图像等分为单独的几个部分，分别计算每个部分的 LBP 统一模式的直方图，然后再形成整幅图像的统计直方图，最后在基于 χ^2 概率统计算子的统计直方图距离中使用最近邻分类来进行人脸识别，公式如下：

$$\chi^2(p, q) = \sum_i \frac{(p_i - q_i)^2}{p_i + q_i} \quad (2-10)$$

在这里， p, q 是两幅图像的统计直方图向量。此方法在 FERET 数据库中取得了很好的训练和测试结果^[31]。

可能有人会认为将图像细分为一个个子区域有些随意（认为这些子区域不能很好地代表整个人脸的特征），同时认为这些子区域很可能导致量化噪声（因为随意的空间量化）和空间分辨率的丢失（子区域中某些位置的像素没有被编码）。假如采用 LBP 或者 LTP 特征的目标是为了提高光照和噪声的鲁棒性，那么因为其算子的非线性将导致一些小的空间偏差。比较图像 X 和图像 Y 在相近位置是否有相似的 LBP 或者 LTP 编码，使用 Hausdorff 距离等相似度度量算子是一个很好的方法，此方法能够更好的控制基于表观的图像匹配的空间松弛度。

假设有一个二维图像 X ，我们计算出其 LBP 或者 LTP 算子，然后将其结果转换成一系列的稀疏的二进制图像 b^k ， k 表示每一种可能的编码（比如，对于 LBP_8^1 的统一模式 k 有 59 种取值）。计算每一个 b^k 的距离变换图像 d^k ， d^k 的每一个像素表示了图像 X 的最相邻的像素和编码值 k 之间的距离（通常使用二维欧几里德距离）。那么图像 X 和图像 Y 的距离或者说是相似度由式 2-11 给出：

$$D(X, Y) = \sum_{pixels(i, j) of Y} \omega(d_X^{k_Y(i, j)}(i, j)) \quad (2-11)$$

这里， $k_Y(i, j)$ 是图像 Y 中像素 (i, j) 的编码值， $\omega()$ 是用户自定义的一个概率统计函数，较常用的是高斯相似度测量 $\omega(d) = \exp\{-(d/\sigma)^2/2\}$ 。

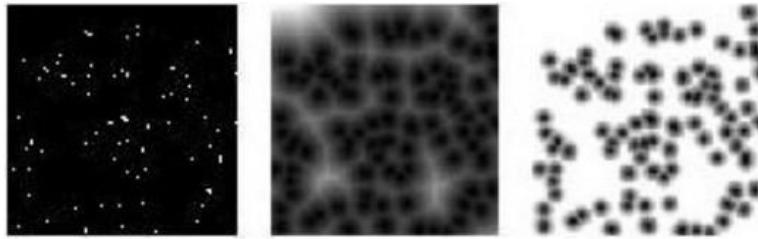


图 2-8 距离转换图示

图 2-8 从左到右表示一个二进制编码层，其距离变换图像，其被约减后的线性图像。

2.5 本章小结

本章介绍了 LBP 的发展过程，从诞生到应用到人脸识别领域；同时也介绍了 LBP 算子的几种扩展形式，包括统一模式，旋转不变的统一模式和改进为三值形式的 LTP 模式。通过这章对于 LBP 算子的详尽描述，我们可以看到其具有旋转不变和灰度鲁棒这两个特性，同时 LBP 也具有描述图像局部特征和全局特征的能力，这对于人脸识别来说是非常有利的。在 Yale 数据库中我们使用本章所介绍的 LBP/LTP 特征提取算法来进行特征的提取，然后进行人脸识别率的对比。我们采用的 Yale 数据库共分为五部分，每一部分的光照环境都不太一样，图 2-9 是同一人在五种光照环境下的人脸图像。



图 2-9 Yale 数据库中五组不同光照人脸图像（每一组 3 张）

我们将基于相似度测量的距离转换（Distance Transform based Similarity Metric）简称为 DT，预处理（Preprocessing）简称为 PP。分别采用 Haar-like, LBP/ χ^2 , LTP/ χ^2 , PP+ LTP/ χ^2 , PP+ LTP/DT 五种方法在 Yale 数据库中测试人脸识别率，结果如表 2-3 所示；在光照条件很好的第一组中进行测试时，跑完算法所用时间如表 2-2 所示：

表 2-2 各算法所用时间（S）

	Haar-like	LBP/ χ^2	LTP/ χ^2
测试组 1(363 幅)	11.5	5.8	8.2

表 2-3 在 Yale 数据库中各方法的识别率 (%)

特征提取/判 别方法	各分组以及其人脸图片数量				
	1	2	3	4	5
	363 幅	456 幅	455 幅	526 幅	714 幅
Haar-like	100	100	86.2	56.6	22.6
LBP/ χ^2	100	100	97.3	65.4	44.4
LTP/ χ^2	100	100	99.2	76.8	54.6
PP+ LTP/ χ^2	100	100	100	98.1	96.3
PP+ LTP/DT	100	100	100	99.2	97.2

从结果我们可以看出，基于原始 LBP 和 LTP 算法的人脸检测系统，其在检测速度上领先于 Viola 提出的基于 Haar-like 特征的检测算法，且在对于人脸的识别率以及在复杂光照环境下的鲁棒性都较 Viola 的方法好。

第三章 降维生成 LBP 子模式

最初的 LBP 模式所生成的二进制模式数量非常的庞大，为了解决这个问题，Ojala 提出了 LBP 统一模式，其与原始 LBP 模式相比，模式的种类大大的减少。由于在图像中提取 LBP 统一模式很简单，并且对于大多数的人脸图像都适用，所以其受到了很多的关注。但是，由于图像的纹理信息是多种多样的，适用于大部分图像的统一模式并不适用于所有的人脸图像，其也不是在所有图像中都占大多数。对于哪些不适用于统一模式的人脸图像，我们应该在减少其模式种类的时候尽可能的保留其主要的用于区别不同图像的特征。

3.1 LBP 统一模式所存在的问题

至今对于 LBP 降维的方法都缺乏深入研究，大多都是使用统一模式来对算子进行维度约减。然而，通过实际的研究我们发现对于一部分图像统一模式并不是最合适的，甚至对于有些图像统一模式是完全不适用的。“统一模式”本质上应该看作是一种特征筛选，它从原始的 2^N 种编码模式中挑选了其中 $N(N-1)+2$ 种模式来描述人脸，再将其他的 $2^N - N(N-1)+2$ 种归为一种混合模式中^[24]。由 LBP 算子的特性我们知道，Ojala 等提出的“统一模式”及其特征子集只与训练时采用的数据库的统计特性相关，这就有可能出现一下两个问题：1、纹理数据库中的纹理不能很好地表示所有的纹理图像，在其数据库中得到的最优特征子集，不一定适用于其他的纹理图像；2、Ojala 等对于特征的选取也没有给出成熟的理论依据，都是凭研究人员自己去尝试。在我们的研究中，发现采用统一模式的 LBP 算子时，混合模式的数量常常会比较大，特别是让我们将实验条件变得比较“恶劣”的情况下（光照情况不好，人脸姿态发生变化，人脸带有表情），这对于我们准确和稳定地提取人脸特征都是很不利的。

以 LBP_8^1 算子为例，其总共有 256 种模式，图 3-1 是一副人脸图像的 LBP 算子统计直方图，其中白色的条表示统一模式，黑色的表示混合模式。通过统计我们可以发现所有模式中统一模式的占比达到了 68%，这表示图像的大部分有效信息都可以由统一模式来表示，但是我们发现混合模式也占到了 32%，其数量也是非常的大，这就说明这幅图像不能只用统一模式来表示。

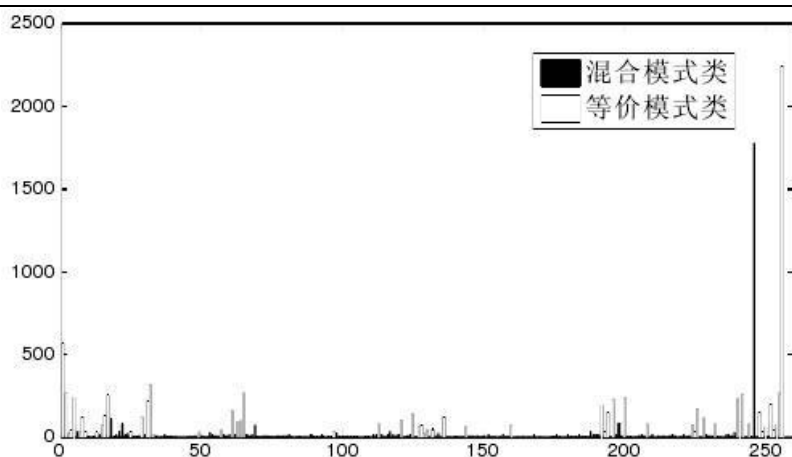


图 3-1 一副图像中 LBP 统一模式和混合模式

3.2 维度约减方法概述

在过去的几十年中，诞生了许多人脸识别算法，但是当中最成功的还是基于表观（appearance-based）的方法^[31]。当使用基于表观图像的方法时，我们通常将一个大小为 $n \times m$ 像素的图像用一个 $n \times m$ 维的向量来代替。但是这 $n \times m$ 维空间对于鲁棒和快速的人脸识别来说，其维数通常都过高，并且人脸图像在高维空间中是稀疏分布的，一个普遍的方法是采用降维技术来解决这些问题^{[32]~[36]}。目前最流行的两种技术分别是：主成份分析（PCA）^[4]和线性判别分析（LDA）^[26]。

与 PCA 的编码需在正交线性空间中进行不同，LDA 的编码也可在非正交的线性空间中进行。正是因为此，普遍认为 LDA 算法是优于 PCA 的，但是在训练集较小时，PCA 是优于 LDA 的，并且 PCA 对于不同的训练集都有很好的适应性。

通过近年来的一些研究表明，人脸图像因为表情、姿态和光照的影响其所表现出的是一种非线性流形结构^{[37]~[40]}。然而，PCA 和 LDA 只能适用于欧几里德结构（即线性子空间），它们无法发现图像空间中的非线性流形结构。近年来，一些非线性子空间的方法（如 Isomap^[41]，LLE^{[42][43]}，Laplacian Eigenmap^[44]）在特征匹配的基准的测试项目上取得了非常不错的成绩。但是因为流形学习方法没有明确的投影矩阵，所以很难直接提取图像的特征，怎样将非线性流形学习算法运用到人脸识别当中来还是一个很大的难题。

本文，我们提出了一种新的方法来进行人脸分析（包括表现和识别人脸），其明确的考虑了人脸的流形结构。我们将其命名为局部保持映射（Locality Preserving Projections, LPP），其用最近邻图来近似流形结构、建立映射，且能够很好的保

留图像的局部结构。具体的讲，就是用一个保留了图像空间局部结构的最近邻图来代表图像的流形结构，由此，通过 LPP 算法就获得了一个人脸图像的子空间图像。在图像空间中的每一个人脸图像都被映射到了低维的人脸空间，其是由一系列叫做 Laplacianfaces 来构建的，这些人脸子空间保留了人脸的局部特征，并且比 PCA 分类能力更强。

3.2.1 主成份分析算法介绍

其主要的算法思想是：将原始的 N 维数据投影到由其协方差矩阵的主要特征向量构成的 K ($\ll N$) 维正交子空间中，在这个过程中数据集的特征维数大大的减少，其有两大优点：1、将样本的数据量大大地进行了压缩；2、因为其采用了正交的特征向量，所以消除了模式样本之间的相关性^[25]。

假设 x 是输入的样本图像，其本质上就是一个二维矩阵，PCA 的目的就是要将二维矩阵向量 x 变化为互相正交的向量 y ，那么 x 与 y 之间的函数关系就可以由式 3-1 来表示：

$$y = A^T x \quad (3-1)$$

则 x 可以转换为：

$$x = Ay = \sum_{i=1}^D y_i a_i \quad (3-2)$$

其中 $A = (a_1, a_2, \dots, a_d)$ 是 R^D 空间的一个基底，如果只用前面的 d 项对 x 进行近似，那么就可以改写成 $\hat{x} = \sum_{i=1}^d y_i a_i$ ，定义的均方误差为：

$$\varepsilon = E[(x - \hat{x})^T (x - \hat{x})] = E[\sum_{i=d+1}^D y_i^2] = \sum_i^D a_i^T E(x\hat{x}) a_i = \sum_{i=d+1}^D a_i^T S a_i \quad (3-3)$$

式中 $y_i = a_i^T x, i=1, \dots, D$ ，随机变量 x 的协方差矩阵用 S 表示。在使得上式正交归一化的前提下，可以求得取均方误差 ε 极值的条件。

$$g = \sum_{i=d+1}^D a_i^T S a_i - \sum_{i=d+1}^D \lambda_i (a_i^T a_i - 1) \quad (3-4)$$

式 3-4 取得极值应该满足条件：

$$(S - \lambda_i I) a_i = 0 \quad i = d+1, d+2, \dots, D \quad (3-5)$$

由上式可得出， λ_i 为协方差矩阵 S 的特征值，特征值对应的特征向量是 a_i 。

当将图像 x 投影到以 S 的特征向量构建的坐标系时，能取得其截断均方误差的极值，当取前面 n 个特征向量来表示 x 时，其均方误差为：

$$\varepsilon = \sum_{i=n+1}^N \lambda_i \quad (3-6)$$

上式表明，当 $\lambda_{n+1}, \lambda_{n+2}, \dots, \lambda_N$ 的算术总和越小，以 a_1, a_2, \dots, a_d 为正交基展开的图像 x 损失的信息量越少。因此在 PCA 算法中一般都只保留主要的特征值对应的特征向量，以此来构建新的投影空间。

3.2.2 线性判别分析算法介绍

LDA (Linear Discriminant Analysis) 也可称做 Fisher 线性判别 (Fisher Linear Discriminant, FLD)，因其最常使用的是 Fisher 线性判别函数，LDA 是模式识别中很经典的一种算法，其是由 Belhumeur 最先引入到模式识别和人脸识别领域的，基本思想是：寻找在降维后的新的特征子空间中，原始特征有最大的类间距离和最小的类内距离，因此，LDA 是一种非常有效的特征抽取方法^[34]。

假设对于一个 R^n 空间有 l 个样本分别是 x_1, x_2, \dots, x_l ，每一个 x 是一个 d 维的样本， n_i 表示属于 i 类样本的个数， S_b 表示类间离散度矩阵， S_w 表示类内离散度矩阵， m 表示所有样本的均值， m_i 表示类 i 的样本均值。

所有样本的均值如下式所示：

$$m = \frac{1}{n} \sum_{i=1}^n x_i \quad (3-7)$$

类内离散度矩阵和类间离散度矩阵如下所示：

$$S_w = \sum_{i=1}^l \sum_{j=1}^{n_i} (x_j^{(i)} - m^{(i)})(x_j^{(i)} - m^{(i)})^T \quad (3-8)$$

$$S_B = \sum_{i=1}^l n_i (m^{(i)} - m)(m^{(i)} - m)^T \quad (3-9)$$

式中， $x_j^{(i)}$ 表示第 i 类的第 j 个样本。那么我们所需要的投影矩阵 W 应该满足式 3-10 的要求：

$$\max_w \frac{W^T S_B W}{W^T S_W W} \quad (3-10)$$

3.2.3 局部保持投影算法概述

在实际研究中，局部纹理有时比全局的特征更加重要，而 PCA 和 LDA 的目标是保留图像的全局信息，所以本节我们描述了一种叫做 LPP (Locality Preserving Projection) 的算法，它是局部子空间保持学习算法的一种扩展算法，LPP 算法主要的目标是保留图像数据和局部纹理的原始集合特征，主要保留的是数据中近邻点之间的几何信息，本质上可以将 LPP 算法看作是对 LE (Laplacian Eigenmap) 的线性逼近^[45]。LPP 的目标函数如下所示：

$$\min \sum_{ij} (y_i - y_j)^2 S_{ij} \quad (3-11)$$

y_i 是经过 $y_i = w^T x_i$ 变换后得来的低维向量， w 是一个投影矩阵，其使得 y_i 在低维空间能够很好的描述 x_i ，并且在原图中相邻的点经过了变化后仍然应该相邻。 S_{ij} 是一个相似度矩阵，其定义由下式 3-12 给出：

$$S_{ij} = \begin{cases} \exp(-\|x_i - x_j\|^2 / t), & \|x_i - x_j\|^2 < \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (3-12)$$

或者

$$S_{ij} = \begin{cases} \exp(-\|x_i - x_j\|^2 / t), & \text{如果 } x_i \text{ 在 } x_j \text{ 的 } k \text{ 邻域或者相反} \\ 0 & \text{otherwise} \end{cases} \quad (3-13)$$

ε 足够的小的常量，并且大于 0， ε 定义了局部邻域的半径。

因此从目标函数 3-11 可以看出，降维后的特征子空间需要能够保持原始数据的主要结构信息，式 3-11 进行简单的代数变换：

$$\begin{aligned} & \frac{1}{2} \sum_{ij} (y_i - y_j)^2 S_{ij} \\ &= \frac{1}{2} \sum_{ij} (w^T x_i - w^T x_j)^2 S_{ij} \\ &= \sum_{ij} w^T x_i S_{ij} x_i^T w - \sum_{ij} w^T x_i S_{ij} x_j^T w \\ &= \sum_i w^T x_i D_{ii} x_i^T w - w^T X S X^T w \\ &= w^T X D X^T w - w^T X S X^T w \\ &= w^T X (D - S) X^T w \\ &= w^T X L X^T w \end{aligned} \quad (3-14)$$

其中 $X = \{x_1, x_2, \dots, x_n\}$ 是训练矩阵, D 是一个对角矩阵, $D_{ij} = \begin{cases} \sum_{k=1}^n S_{ik} & i = j \\ 0 & i \neq j \end{cases}$,

$L = D - S$ 是一个拉普拉斯矩阵, 如果把原始数据中的像素点作为一个定点, 将其近邻图中近邻点相连, 则构成了一个用 S 构成的带权图 G , 图 G 的各个顶点的值是由对角矩阵 D 提供, D_{ii} 对应着第 i 幅训练图像, D_{ii} 越大反映出第 i 幅图像越重要, 因此决定投影过后的 y_i 就越重要, 故为了去除上述最小化问题中的图像缩放, 并取消空间平移的任意性^[46], 我们做出如下限制:

$$\sum_i D_{ii} y_i^2 = 1 \Rightarrow w^T XDX^T w = 1 \quad (3-15)$$

则最小化的问题转换为:

$$\arg \min_{w^T XDX^T w = 1} w^T XLX^T w \quad (3-16)$$

现在我们需要求得一变换矩阵, 其能最小化目标函数, 可通过求解式 3-17 的广义特征值来得到。需要注意的是矩阵 XLX^T 和 XDX^T 都是半正定并且对称的。

$$XLX^T w = \lambda XDX^T w \quad (3-17)$$

3.2.4 LPP 的统计特性

LPP 特征也具有很多统计方面的特性, 假设 x 和 y 是两个随机的变量, 我们定义一个线性映射 $x \rightarrow w^T x$, 其很好的保留了局部纹理结构, 并且最小化了 $w^T x$ 和 $w^T y$ 之间的距离 L^2 , 也就是说:

$$\min_w E(|w^T x - w^T y|^2 | \|x - y\| < \varepsilon) \quad (3-18)$$

这里 ε 是一个极小值, 并且为正。令 $z = x - y$, 那么上式可以改写为:

$$\min_w E(|w^T z|^2 | \|z\| < \varepsilon) \quad (3-19)$$

展开可得:

$$\begin{aligned} & E(|w^T z|^2 | \|z\| < \varepsilon) \\ &= E(w^T z z^T w | \|z\| < \varepsilon) \\ &= w^T E(z z^T | \|z\| < \varepsilon) w \end{aligned} \quad (3-20)$$

已知一个样本点集 x_1, x_2, \dots, x_n ，我们定义一个指示函数 S_{ij} ，如下式：

$$S_{ij} = \begin{cases} 1 & \|x_i - x_j\|^2 < \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (3-21)$$

假设 d 是 S_{ij} 中的非零值， D 是一个对角矩阵，其每一项是 S 的每一列之和，

即 $D_{ii} = \sum_j S_{ji}$ ，根据大数定律， $E(zz^T | \|z\| < \varepsilon)$ 能被以下样本点估计：

$$\begin{aligned} & E(zz^T | \|z\| < \varepsilon) \\ & \approx \frac{1}{d} \sum_{\|z\| < \varepsilon} zz^T \\ & = \frac{1}{d} \sum_{\|x_i - x_j\| < \varepsilon} (x_i - x_j)(x_i - x_j)^T \\ & = \frac{1}{d} \sum_{i,j} (x_i - x_j)(x_i - x_j)^T S_{ij} \\ & = \frac{1}{d} \left(\sum_{i,j} x_i x_i^T S_{ij} + \sum_{i,j} x_j x_j^T S_{ij} - \sum_{i,j} x_i x_j^T S_{ij} - \sum_{i,j} x_j x_i^T S_{ij} \right) \\ & = \frac{2}{d} \left(\sum_i x_i x_i^T D_{ii} - \sum_{i,j} x_i x_j^T S_{ij} \right) \\ & = \frac{2}{d} (XDX^T - XSX^T) \\ & = \frac{2}{d} XLX^T \end{aligned} \quad (3-22)$$

这里 $L = D - S$ 是拉普拉斯矩阵，矩阵 X 的第 i 列是 x_i ，由上式我们可以得到同式 3-14 一样最小化描述。

3.3 LPP, PCA, LDA 的理论分析

3.3.1 局部保持投影（LPP）与主成份分析（PCA）的关系

由上一节我们知道 XLX^T 是一个协方差矩阵，如果拉普拉斯矩阵 $L = \frac{1}{n}I - \frac{1}{n^2}ee^T$ ， n 代表样本点个数， I 是单位矩阵， e 是每一项的列向量，事实上这里的拉普拉斯矩阵去掉了样本向量的样本均值。我们将每一项的权值矩阵 S

取值 $1/n^2$, $S_{ij} = 1/n^2, \forall i, j. D_{ii} = \sum_j S_{ji} = 1/n$, 因此, 拉普拉斯矩阵 $L = D - S = \frac{1}{n}I - \frac{1}{n^2}ee^T$ 。 m 代表所有样本的均值, $m = 1/n \sum_i x_i$ 。那么我们就得到下式:

$$\begin{aligned}
 XLX^T &= \frac{1}{n}X(I - \frac{1}{n}ee^T)X^T \\
 &= \frac{1}{n}XX^T - \frac{1}{n^2}(Xe)(Xe)^T \\
 &= \frac{1}{n}\sum_i x_i x_i^T - \frac{1}{n^2}(nm)(nm)^T \\
 &= \frac{1}{n}\sum_i (x_i - m)(x_i - m)^T \\
 &\quad + \frac{1}{n}\sum_i x_i m^T + \frac{1}{n}\sum_i m x_i^T - \frac{1}{n}\sum_i m m^T - m m^T \\
 &= E[(x - m)(x - m)^T] + 2m m^T - 2m m^T \\
 &= E[(x - m)(x - m)^T]
 \end{aligned} \tag{3-23}$$

这里的 $E[(x - m)(x - m)^T]$ 就是样本集的协方差矩阵。

由以上的分析我们可以看出在 LPP 算法中最核心的就是权值矩阵 S , 当我们需要保留全局的结构的时候, 我们将 ε 设置为无穷大, 然后选择具有最大特征值的特征向量来构建矩阵 XLX^T ; 当我们需要保留图像的局部结构的时候, 我们将 ε 设置为无穷小, 然后选择具有最小特征值的特征向量来构建矩阵 XLX^T 。很重要的一点是, 当 ε 无穷小的时候, 这拉普拉斯矩阵不再是样本的协方差矩阵, 因此, 保留局部特征的方向向量不再是指向最小协方差的方向。

3.3.2 局部保持投影 (LPP) 和线性判别分析 (LDA) 的关系

LDA 寻找的是通过有效的判别式来提取特征, 其表示的是样本类内和类间的关系, 下式是解决广义特征值问题的投影算子:

$$S_B w = \lambda S_w w \tag{3-24}$$

这里 S_B 和 S_w 分别表示的是类间和类内离散度。假如样本有 l 类, 第 i 类包含有 n_i 个样本点, 让 $m^{(i)}$ 代表第 i 类的平均向量, $x^{(i)}$ 代表与第 i 类相关的随机向量, $x_j^{(i)}$ 代表第 i 类的第 j 个样本点, 那么我们重写类内离散度矩阵:

$$\begin{aligned}
 S_w &= \sum_{i=1}^l \left(\sum_{j=1}^{n_i} (x_j^{(i)} - m^{(i)})(x_j^{(i)} - m^{(i)})^T \right) \\
 &= \sum_{i=1}^l \left(\sum_{j=1}^{n_i} (x_j^{(i)} (x_j^{(i)})^T - m^{(i)} (x_j^{(i)})^T \right. \\
 &\quad \left. - x_j^{(i)} (m^{(i)})^T + m^{(i)} (m^{(i)})^T \right) \\
 &= \sum_{i=1}^l \left(\sum_{j=1}^{n_i} (x_j^{(i)} (x_j^{(i)})^T - n_i m^{(i)} (m^{(i)})^T \right) \\
 &= \sum_{i=1}^l \left(X_i X_i^T - \frac{1}{n_i} (x_1^{(i)} + \dots + x_{n_i}^{(i)})(x_1^{(i)} + \dots + x_{n_i}^{(i)})^T \right) \quad (3-25) \\
 &= \sum_{i=1}^l \left(X_i X_i^T - \frac{1}{n_i} X_i (e_i e_i^T) X_i^T \right) \\
 &= \sum_{i=1}^l X_i L_i X_i^T
 \end{aligned}$$

式中 $X_i L_i X_i^T$ 是第 i 类样本的协方差矩阵, $X_i = [x_1^{(i)}, x_2^{(i)}, \dots, x_{n_i}^{(i)}]$ 是一个 $d \times n_i$ 维矩阵, $L_i = I - 1/n_i e_i e_i^T$ 是一个 $n_i \times n_i$ 维矩阵, I 是一个单位矩阵, $e_i = (1, 1, \dots, 1)^T$ 是一个 n_i 维向量。为了进一步简化式 3-25, 我们定义 $X = (x_1, x_2, \dots, x_n)$, 那么:

$$W_{ij} = \begin{cases} 1/n_k & \text{如果 } x_i \text{ 和 } x_j \text{ 都是属于第 } k \text{ 类} \\ 0 & \text{其他} \end{cases} \quad (3-26)$$

我们又知道 $L = I - W$, 那么我们就得到:

$$S_W = X L X^T \quad (3-27)$$

其中 W 是权值矩阵, 比如 W_{ij} 是边缘 (x_i, x_j) 的权重, W 反映的是数据集的类内离散度, 矩阵 L 也被叫做拉普拉斯矩阵, 其在 LPP 中的作用很重要^[45]。

同样的, 我们也可以计算类间离散度矩阵 S_B , 算式如下:

$$S_B = \sum_{i=1}^l n_i (m^{(i)} - m)(m^{(i)} - m)^T$$

$$\begin{aligned}
 &= \left(\sum_{i=1}^l n_i m^{(i)} (m^{(i)})^T \right) - m \left(\sum_{i=1}^l n_i (m^{(i)})^T \right) \\
 &\quad - \left(\sum_{i=1}^l n_i m^{(i)} \right) m^T + \left(\sum_{i=1}^l n_i \right) m m^T \\
 &= \left(\sum_{i=1}^l \frac{1}{n_i} (x_1^{(i)} + \dots + x_{n_i}^{(i)}) (x_1^{(i)} + \dots + x_{n_i}^{(i)})^T \right) \\
 &\quad - 2n m m^T + n m m^T \\
 &= \left(\sum_{i=1}^l \sum_{j,k=1}^{n_i} \frac{1}{n_i} x_j^{(i)} (x_k^{(i)})^T \right) - 2n m m^T + n m m^T \\
 &= X W X^T - 2n m m^T + n m m^T \\
 &= X W X^T - n m m^T \\
 &= X W X^T - X \left(\frac{1}{n} e e^T \right) X^T \\
 &= X \left(W - \frac{1}{n} e e^T \right) X^T \\
 &= X \left(W - I + I - \frac{1}{n} e e^T \right) X^T \\
 &= -X L X^T + X \left(I - \frac{1}{n} e e^T \right) X^T \\
 &= -X L X^T + C
 \end{aligned} \tag{3-28}$$

这里 $e_i = (1, 1, \dots, 1)^T$ 也同样是一个 n 维的向量， $C = X(I - \frac{1}{n} e e^T) X^T$ 是数据集的协方差矩阵。因此 LDA 的广义特征向量问题可以改写为式 3-29:

$$\begin{aligned}
 S_B w &= \lambda S_W w \\
 \Rightarrow (C - X L X^T) w &= \lambda X L X^T w \\
 \Rightarrow C w &= (1 + \lambda) X L X^T w \\
 \Rightarrow X L X^T w &= \frac{1}{1 + \lambda} C w
 \end{aligned} \tag{3-29}$$

因此 LDA 的映射可以通过求解下列方程得到:

$$X L X^T w = \lambda C w \tag{3-30}$$

如果我们所采用的测试训练样本的均值为 0，那么协方差矩阵就可以改写为 XX^T ，这就是 LPP 算法中的 XX^T 的矩阵，那么 LPP 和 LDA 就存在相类似的算式，所以我们可以采用 LPP 算子来表示类间关系。

3.4 基于 LPP 降维算法构建 LBP 子模式

LPP 有很好的数据降维能力和基于流形的保持人脸局部纹理能量的能力，这就使得 LPP 的特征约减能力很强。对于待分类的人脸，设模式的类别有 l 个，每类的训练样本图像有 n_i 个，总样本数为 N ，假设第 i 类第 j 幅图像为 $f_{ij}(x, y)$ ，计算出其 LBP_p^R 算子后，统计出其整个直方图，整个直方图信息可以有下式给出：

$$H_{ij}(m) = \sum_{x,y} I\{f_{ij}(x, y) = m\}, m = 0, \dots, B-1 \quad (3-31)$$

式中，我们用 B 来代表所拥有的 LBP 模式的个数， m 表示一个 LBP 模式，它的取值取决于我们所采用的 LBP 算子的模式，比如采用 LBP_8^1 算子，那么就一共有 256 种模式可能， m 的取值范围就是 0~255。 I 是如式 3-32 所示的函数：

$$I(A) = \begin{cases} 1 & A = true \\ 0 & A = false \end{cases} \quad (3-32)$$

将图像的 LBP 模式直方图 H_{ij} 看为一个一维向量，用矩阵 H 来描述所有图像的统计直方图，那么 H 就是一个 $B \times N$ 维的矩阵。LPP 的主要思想就是通过一定的性能要求来求得线性变化 W ，以此来实现对原始数据 H （维数较高）进行降维得到 \bar{H} ： $\bar{H} = WH$ ， \bar{H}_{ij} 的总 bin 数肯定是要远远小于 H_{ij} 的。

根据上面几节我们可知，变化矩阵 W 可通过求解 $XLX^T w = \lambda XDX^T w$ 的广义特征值和特征向量来得到，我们将由 LPP 降维过后剩下的近邻流形空间称做 LBP 模式的子空间，那么 \bar{H} 就是 H 在这个子空间中的投影，我们就称 \bar{H} 为 LBP 子模式。

3.5 LBP 统一模式与子模式的比较

LBP 子模式对于不同的图像都能够占到很高的比重，而且图像的种类繁多，不同类型的图像其关键信息又千差万别，所以仅仅依靠固定的统一模式是很难适用于所有图像的，但是 LBP 子模式不同，其所采用的降维算法能够用很少的特征维度来表示原始图像，而且根据不同图像的纹理特点选择不同的降维算法就更能找到更合适的 LBP 子模式，如果我们采用的是 PCA 方法对算子进行降维的话，还消除掉了 LBP 子模式各模式之间的相关性，而且 LBP 子模式都舍弃掉了一些特征值较小的特征向量（这表示其在图像中所占比重较小并且特征辨别性较差），这也从侧面抑制了噪声对于人脸图像的影响。

我们在 ORL 数据库中分别测试了 LBP 子模式和统一模式在模式数量和特征提取方面的性能。我们选取一个人的 10 幅人脸图像，训练样本选取前面的 5 幅图像，其余的作为测试图像。实验结果如表 3-1 所示：

表 3-1 LBP 子模式和统一模式

	图像分块 2X2		图像分块 3X3		图像分块 4X4		图像分块 5X5	
	识别率	特征维数	识别率	特征维数	识别率	特征维数	识别率	特征维数
LBP 统一模式	0.87	236	0.908	531	0.934	944	0.933	1475
LBP 子模式	0.93	128	0.968	288	0.955	512	0.946	800

从上表我们可以看出，在 ORL 人脸数据库中，如果使用相同的图像分块，LBP 子模式在特征维数的数量和识别率上都要优于统一模式。

3.6 实验结果与分析

我们还是在 Yale 数据库中进行测试，选取 15 个人一共 165 张灰度图像，其中随机选择每个人的 6 幅图像进行训练（一共就 90 幅训练图像），剩余的作为测试图像。对于 PCA 和 LPP 来说，其识别率是跟其降维过后的特征维数有关的，维数越多性能越好（但不能过多，会造成训练结果的过适应，到时识别率下降），但是其所耗时间会成倍增加。下表 3-2 就展示了三种方法的在数据库中的错误率以及得到最低错误率时的算子维度（采用最近邻分类来进行识别的判定）。

从下面结果我们可以看出，基于 LPP 降维的方法比 PCA 和 FLDA 能够得到更低的错误率，虽然相较于 FLDA 所需的特征维度较高，但是换来的是更低的错误率。从图 3-2 我们可以看到，PCA，FLDA 和 LPP 在特征维度为 33，14，28 时就能得到最优子空间，更多的特征维度并不能带来更低的错误率，对于 LPP 来说，随着维度的增加还有可能提高错误率。

表 3-2 在 Yale 数据库中的性能比对

降维算法	维度	错误率 (%)
PCA	33	25.3
FLDA	14	20
LPP	28	11.3

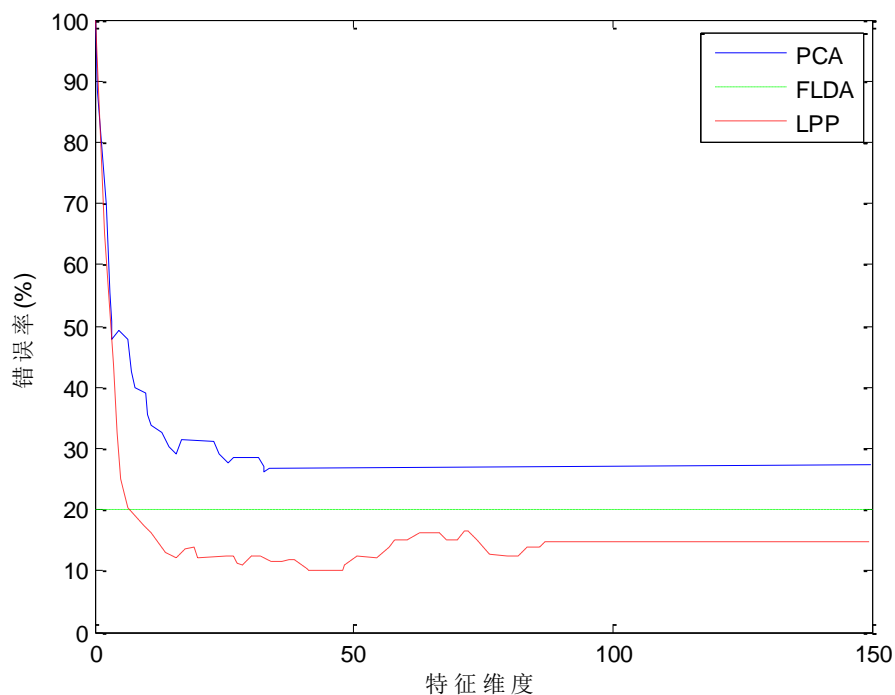


图 3-2 识别率与特征维度的关系

3.7 本章小结

本章我们主要研究了几种思路不同的降维算法。首先总结了当前 LBP 统一模式算法结构上所存在的缺陷，然后讨论了 PCA，LDA 和 LPP 三种降维算法，并且选中了基于流形理论的 LPP 算法作为降维算子，以此得到 LBP 子模式。通过最后的实验我们可以看出，LBP 子模式相较于统一模式具有更低的特征维数，并且维数约简采用的方法更加灵活、效果更好。

第四章 图像预处理及人眼检测

由于背景、姿态、光照等环境因素的影响，人脸在不同的视觉条件下，表现出不同的形态，这给人脸的检测和识别带来很大的困难。为了控制这种图像的可变性，一个有效的并且最实用的方法是让人脸识别系统在一种受控的状态下进行，但是这就失去了使用人脸来进行身份识别的意义。因此，在目前主流的人脸识别算法中，提取人脸特征之前都要将人脸图像标准化，也就是需要对人脸图像进行预处理。

4.1 抑制图像噪声

图像噪声对于人脸系统的每一个环节都有非常大的影响。所以，我们在进行人脸识别之前需要尽可能的消除掉噪声对于图像的影响，图像处理中通常的去噪算法对于人脸识别都使用，主要有一下几种：

1、空间平滑滤波

平滑滤波主要采用的是邻域平均法，大部分是利用卷积运算的图像平滑方法。主要思想是通过一像素点及其邻域内像素点求平均来去除突变的像素点，从而滤掉某些跳变较大的噪声，其优点是算法简单、计算速度快，但是会造成图像（特别是边缘图像）一定程度的模糊^[47]。

2、中值滤波

中值滤波是一种非线性滤波器，在一定程度上可以克服线性滤波器所带来的图像模糊问题，在滤除噪声的同时，较好地保留了图像的边缘信息^[47]。

3、二维离散卷积滤波

众所周知，人脸图像实际上是一个二维矩阵，所以我们主要采用二维离散卷积算子来对人脸图像进行卷积滤波^[47]，其卷积定义如下：

$$(x, y) \otimes F(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G(m, n) F(x-m, y-n) \quad (4-1)$$

在公式 4-1 中， $M \times N$ 表示图像区域大小， $G(x, y)$ 表示人脸图像， $F(x, y)$ 表示所用的卷积算子。

4、频域滤波

频率域滤波是利用傅立叶算子将图像从空间或时间域转换到频率域进行图像增强的方法^[47]。一般来说噪声在频率域其所在的频带较大，而人脸图像所处的频率区域较低。

频域常用的几种低通滤波器为理想低通滤波器(Ideal circular low-pass filter)、巴特沃思(Butterworth)低通滤波器、指数低通滤波器及高斯低通滤波器。这些低通滤波器，都能在图像内有噪声干扰成分时起到改善的作用。

4.2 人脸图像直方图均衡

主要是利用直方图均衡调整人脸的亮度分布，其加大了人脸图像整个区域灰度的对比度，使得灰度的分布更加的均匀，在一定程度上消除了变化的光照环境对于人脸图像的影响。

计算步骤如下：

- 1、 设原始图像灰度级为 f_i ， $i = 0, 1, \dots, L-1$ ，这代表着图像中共具有 L 个灰度等级；
- 2、 统计各个灰度级的像素个数，记为 n_i ， $i = 0, 1, \dots, L-1$ ；
- 3、 原始图像直方图每一个灰度级的频数用 $P_f(f_i) = \frac{n_i}{n}$ ， $i = 0, 1, \dots, L-1$ 来表示；
- 4、 计算累积分布函数 $C(f) = \sum_{i=0}^k P_f(f_i)$ ， $i = 0, 1, \dots, k, \dots, L-1$ ；
- 5、 应用以下公式计算输出图像的灰度级 g_i ， $i = 0, 1, \dots, P-1$ ，输出图像灰度级级数我们用 P 来表示： $g_i = INT[(g_{\max} - g_{\min})C(f) + g_{\min} + 0.5]$ ，其中 INT 为取整的符号；
- 6、 映射后各灰度级的数量 n_j ， $j = 0, 1, \dots, P-1$ ；
- 7、 输出图像直方图 $P_g(g_j) = \frac{n_j}{n}$ ， $j = 0, 1, \dots, P-1$ ；
- 8、 用 f_i 和 g_j 之间的映射关系来获得直方图分布近似均匀的输出图像，这就使得整幅图像的灰度值更加的平滑，光照的影响更加的小。

图 4-1 是灰度均匀后的效果，我们可以看到，将图像经过直方图均衡化处理过后，图像的明暗度明显相较于之前的图像有了很大的提升，整幅图像明暗均匀，没有区域明显过亮。



(a) 直方图均衡化前

(b) 直方图均衡化后

图 4-1 人脸图像直方图均衡化事例

4.3 人脸图像校正

经过人脸检测并成功分割出人脸区域后，为了进一步提高人脸识别的质量，则需要对人脸区域进行校正。因为人脸图片不可避免的存在倾斜、偏转，如果能有效的克服倾斜和偏转的干扰，那么人脸识别的准确度则会取得较大提升。经过长期的研究实验，针对人眼进行人脸校正具有速度快、精确度高的效果，因为人眼特征比较明显，而且本身具备对称性，校正过程会比较简单快速。所以人脸校正的问题就变成了如何进行人眼定位的问题。人眼定位及人眼特征提取在模式识别领域有着非常重要的地位，目前使用比较多且比较成熟的人眼定位技术主要有三种：Adaboost+Haar 小波算法，AAM 算法以及本文使用的 ASEF 滤波器。

4.3.1 人眼定位简介

从图像中检测某个特定特征时，我们通常使用相关实例模版进行匹配检测的方法，这种方法的优点在于：1.较为简单；2.具有较高的效率。但是这种方法也有非常明显的弱点，其中影响最大的弱点便是：如果检测的目标与模版几乎完美匹配，那么得到的响应就会非常高。但是，当目标和模版匹配不匹配时，那么得到的结果响应是不可测的。使用实例模版进行人眼检测的方法有很多，Adaboost 算法以及 AAM 算法都属于这类。

AdaBoost 算法是 Viola 和 Jones 提出的主要应用于人脸检测方法，人脸检测

技术才真正得以实用，AdaBoost+Haar 小波算法的主要目的就是提取 Haar-like 特征，这种特征可以很好的反映人脸的灰度分布^[9]。那么使用 Adaboost 的方法来找人眼的过程就被转化为找到描述人眼图像分布的 Haar-like 特征，然后从大量的 Haar-like 特征中挑选出最优的描述人眼图像的灰度分布特征。而使用 Adaboost 训练分类器，核心思想便是：将众多不同的分类器（弱分类器）在同一个训练样本集中进行训练，选取出分类效果好的分类器，再采用级联或者并联的方式将这些分类器集合起来，最终构成一个新的分类器具备更好的分类效果（强分类器）。基于 AdaBoost 算法的人眼检测方法主要包括三部分内容：Haar-like 特征提取、AdaBoost 学习算法和运用级联分类器来将弱分类器生成强分类器，核心内容为 AdaBoost 算法。

AAM 模型算法，即主动表观模型算法，在 1998 年 Cootes 等在 ASM 的基础上提出了 AAM^[11-16]。AAM 近几年在人脸识别、表情识别等领域有了广泛的应用。同理，AAM 也适用于人眼检测。

AAM 的核心思想是：模型的建立和匹配搜索。模型的建立包括了形状模型的建立、纹理模型的建立，最后合成表观模型。匹配算法使用梯度下降法。如此针对建立好的模型，即使模型有形态上的变动，比如人的表情变化，在匹配时也会根据模型的形态进行调整，是模型在变化后进行匹配。所以，AAM 在人眼检测中，可以很好的匹配人眼的种种变化，在检测精度上有很大提升。

为了解决实例模版所存在的一些问题，一系列相关滤波器被开发出来，在保留匹配目标得到高响应结果的同时，尽量抑制类似或者诱导目标的结果响应。而这些滤波器的不同，在于他们是如何通过训练样本进行构建的。

虽然这些方法大大增长相关滤波器的性能,但仍有进一步改善的余地。因此，我们提出了 Average of Synthetic Exact Filters (ASEF) 。首先，滤波器的相关响应面是通过所有训练样本构建的。第二，结果过滤器对每个训练图像得到简单的平均值，因此结果滤波器器不易过学习能更大包容训练集。

4.3.2 眼睛的定位

一般来说,眼睛定位算法会返回正面像（略微有些偏转的侧面像）左右眼的中心坐标。作为一个成功的算法，必须要返回眼睛真正的位置。特别是在需要计算距离的时候，比如说两眼之间的距离，精确的得到眼睛位置就显得尤为重要了。通常有一种人眼定位的方法，是使用经验数据，利用事先已经得到的精确位置。

这本质上就是推断出眼睛在脸上的坐标位置。在人脸检测非常准确的时候，这种方法是准确并且容易的。然后最难的地方在于如何精确的定位眼睛位置而不依靠之前的约束，ASEF 滤波器就是为了解决这类问题而出现的。

在 Viola & Jones 的 AdaBoost 算法中，实际上是使用了近似区域约束这一前置条件。如果没有这一前置约束条件，那么级联分类器就会生成许多错误检测因此导致结果变差。此外，Gabor 小波也被应用在人眼定位当中，当我们把 ASEF 滤波器和 Gabor 算法放在一起比较的时候，发现小波算法的计算复杂度至少是 ASEF 的 20 倍。

4.3.3 ASEF 相关滤波器

近年来，为了应对不同输入信号专家们已经提出了一系列相关滤波器。首先是 SDF 滤波器，它对正样本有很好的响应并且会抑制负样本。在 SDF 滤波器之后，MVSDF 滤波器、MACE 滤波器和 OTF 滤波器都被提出来。针对滤波器这些各种方面的改进，目的都是为了提高滤波器应对噪声时的性能，并且提高滤波器的空间解析度。这四种滤波器的训练方式类似，样本在训练前都需要就行标注（是否有目标），但标注对相关滤波器来说，其实是不需要的，甚至会影响滤波器的鲁棒性。非相关约束滤波器比如 MACH 滤波器和 UMACE 滤波器，则不需要约束条件，它们使用平均图像代替相关性响应。

ASEF 滤波器与以前的相关滤波器不同，最主要的原因是在于卷积定理的应用，这大大简化了输入训练图像与输出相关面之间的映射。在频率域对相关性的操作变得非常简单，因此也可以独立的对每个傅立叶集的系数进行计算，再将结果转换回空间域。作为一个整体的相关输出可以代表每个单个图像。

ASEF 滤波器与之前所说的几个滤波器最大的不同在于：ASEF 滤波器是过约束的，而 SDF 仅仅对于每个训练图片输出一个合成相关值。ASEF 滤波器对图片的每个部分进行训练来得到期望的响应。这种响应会在感兴趣的目标中心形成一个波峰。

这种相关性输出所带来的一个结果就是：在对每张训练图片的约束和自由度之间取得了完美的平衡，因此对每张训练图片完成了一个“精确滤波器”。过拟合问题也会因为采用了平均滤波器而得到避免。虽然 UMACE 滤波器也是通过平均避免过拟合，但 UMACE 只是对所有图片的平均而 ASEF 是对平均图片集做精确滤波。

总的来说，ASEF 滤波器在选择训练图片集和指定综合输出时提供了更多的自由度。好处之一是不再再以目标为图片的中心。对于每张训练图像，我们指定一个需要的输出并且标出目标出现的波峰。因为相关峰会步调一致的随着目标移动，所以所有精确滤波器都会对着相关过程进行反向注册。这增强了训练的灵活性，允许我们对每张图片去自定义需要的输出。例如，训练图片含有多个目标时，那么综合输出就会含有多个相匹配的峰值。

4.3.4 ASEF 滤波器定义

表 4-1 描述了 ASEF 的工作原理和状态。 f_1 是训练集中的单幅图片， g_1 是期望获得的相关滤波输出。 h_1 则是从 f_1 到 g_1 的傅氏变换中所需要的滤波器。最终的滤波器，是通过多个滤波器求平均而得到的。需要记住的是 f_i, g_i 是一对训练对，包含了训练图像和需要的相关输出。 g_i 是在相关图像里用一个很亮的光点表示的，位置在目标的中心，比如在这里就表示人的左右眼。特别的是，我们将 g_i 定义为一个二维高斯分布，其中心就是目标，半径为 σ ：

$$g_i(x, y) = e^{-\frac{(x-x_i)^2 + (y-y_i)^2}{\sigma^2}} \quad (4-2)$$


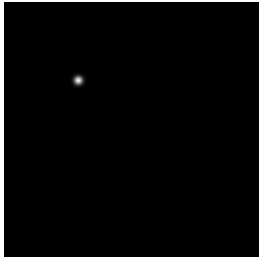
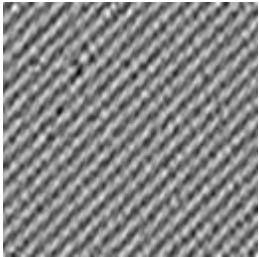

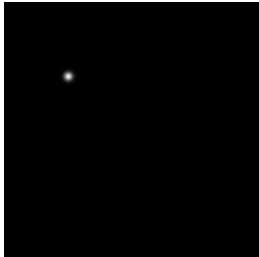
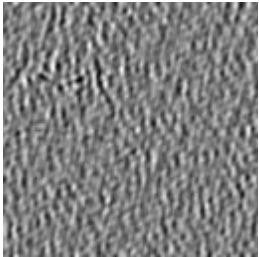

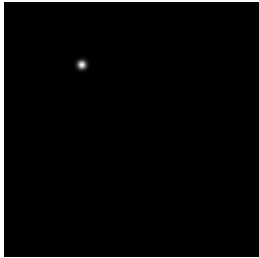
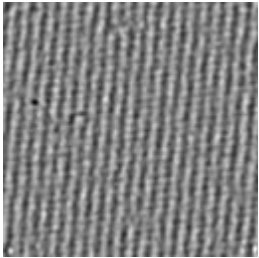
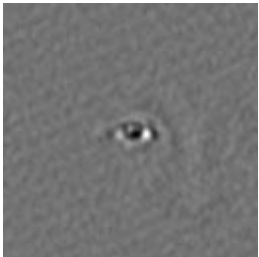
其中 σ 的定义与是用来平衡噪声容忍度和信号的冒刺程度。

根据卷积定理中，我们知道：

$$g(x, y) = (f \otimes h)(x, y) = \Gamma^{-1}(F(w, v)H(w, v)) \quad (4-3)$$

精确滤波器便基于以上关系式而成立： f 是输入图像， h 是滤波器， g 就是相关输出在空间域的表示。而等式中的大写 F 、 H 和 G 则是相关参数在频域空间中的表示。

表 4-1 ASEF 滤波器原理

		
f_1	g_1	h_1
		
f_2	g_2	h_2
		
f_3	g_3	h_3
		
		$1/N \sum_{i=0}^N h_i$

我们发现，当训练的图片越多时，ASEF 滤波器的效果越好。一般来说，图

像变换包括旋转、缩放和转换，这些都增强了 ASEF 滤波器的鲁棒性，因为它们给 ASEF 滤波器提供了更多变化的图片。这一系列的变换同样增强了滤波器在目标区域的响应使得更容易构建效果更好的滤波器。

4.4 基于 ASEF 的人脸校正

使用 ASEF 进行人脸的校正，是指对原始图片使用人脸检测算法后，对固定尺寸的人脸图像使用 ASEF 算法，获得人脸的眼睛坐标。通过提前预设的人脸模版，利用眼睛坐标对图片进行旋转，从而得到标准的正面人脸像。

4.4.1 基于 ASEF 的人眼检测

得到人脸的图片后，我们就可以通过人脸图片确定眼睛的位置。在滤波器训练好的情况下，使用滤波器方法的速度和精度都是高于模版匹配方法的。通过 ASEF 检测，可以精准的定位到人眼中心坐标，而且速度小于 8ms，基本达到实时检测的效果。由于得到的是人眼的中心坐标，因此可以用 ASEF 来直接定位瞳孔，针对人眼坐标进行各种图形变换也更加精确。如图 4-2 所示：



图 4-2 通过 ASEF 得到的人眼坐标

对于其他情况，比如：1) 人脸姿势偏转 2) 眼珠位置不在中间等情况，ASEF 算法也要较好的适应性，具体见图 4-3。由于 ASEF 人眼检测很大程度上取决于人脸检测效果的好坏，而人脸检测现在应用最广泛的技术，就是 Viola 和 Jones

提出的 AdaBoost 算法, 这个算法对人脸的偏转适应性不好, 因此如果无法检测到人脸, 那么在原始图片上应用 ASEF 算法, 就会造成很大的运算负担。但是在 AdaBoost 能检测到的人脸图片中, ASEF 算法都能很好的检测到眼睛的中心坐标。

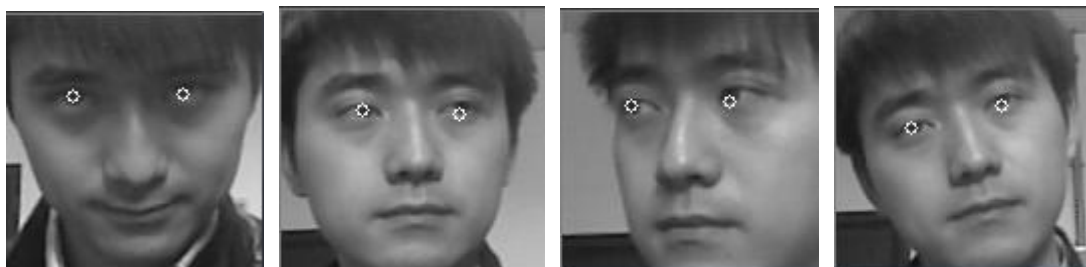


图 4-3 不同姿势 ASEF 检测效果图

4.4.2 人脸校正

正面人脸图像在会存在一定角度的倾斜, 虽然倾斜的角度不大 (太大后人脸检测算法会失效), 但是这会给人脸的识别带来很大的麻烦, 因为人脸识别的主要过程就是将训练样本和待测样本进行比较, 但如果人脸姿势存在巨大差异, 就势必会对匹配造成巨大影响, 所以我们就要在识别之前将人脸进行校正, 尽量保证进行识别的人脸是端正的。

在上一节中, 我们通过 ASEF 算子得到人眼中心的坐标, 我们可以利用人眼中心的坐标, 对整个图片进行旋转, 从而得到正面的人脸像。图 4 说明了校正过程, 假设图 4-4(a)是校正前的图像, 这里我们规定图像的尺寸是 120×120 , 假设中间两个黑点是检测到的人眼中心坐标。图 4-4(b)是我们要求的人脸输出图像, 尺寸是 128×160 , 眼睛坐标限定在 $(26, 40)$ 和 $(102, 40)$ 两点, 通过确定眼睛坐标, 可以保证人脸会保持正向, 因此, 需要的只是通过两点而将图像进行旋转的过程。

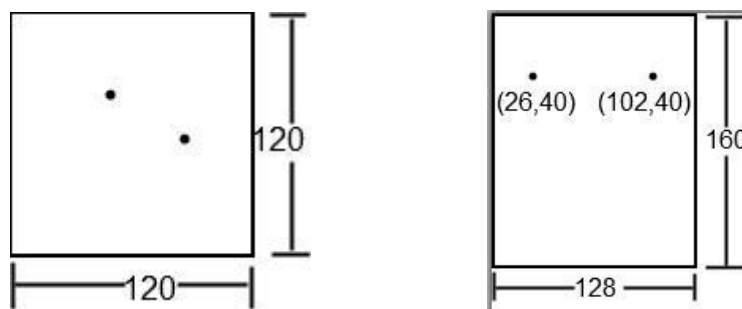


图 4-4 人脸的校正过程

对于图像的几何操作，这里的操作包括各种方式的拉伸，包括一致性缩放和非一致性缩放。这些拉伸、扭曲、旋转图像的操作被叫做图像的几何转换。对于平面区域，有两种方式的几何转换：一种是基于 2×3 的矩阵进行的变换，也叫仿射变换；另一种是基于 3×3 矩阵的变换，又称透视变换或者单映射变换。

一个任意的仿射变换可以表达为乘以一个矩阵在加上一个向量的形式，这里使用矩阵的定义如下：

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}, B = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \Rightarrow T = [A \quad B], X = \begin{bmatrix} x \\ y \end{bmatrix}, X' = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (4-4)$$

仿射变换 $A \cdot X + B$ 等效于将向量 X 扩展成 X' ，并且只是将 X' 左乘一个 T ，我们需要做的就是求出仿射矩阵 T ，再利用这个矩阵将整张人脸图片进行旋转，从而得到校正后的人脸。



图 4-5 人脸校正过程

从左至右：原始图像，通过仿射变换后的人脸，将变换后的人脸转为灰度图像

4.4.3 ASEF 的特点

通过使用 ASEF 人脸校正，我们可以让待检测样本更加标准，提高识别率。与其他人脸校正方法相比，ASEF 人脸校正有以下特点：

1、人脸位置的更精准定位

传统的 Adaboost 算法，在检测人脸时，检测到的人脸还包含有背景信息，而且会包括头发等人脸特征。但研究表明，在人脸识别中，最重要的人脸信息集中在人脸的中心部位，尤其是眼睛、鼻子两侧等部位，而人脸的四周部位成分变化

太大的话，反而会影响人脸识别的效果，造成不小的干扰。因此在人脸识别过程中我们应该尽量去除这些部分造成的干扰。由于 ASEF 基于人眼进行校正，在保证眼睛定位准确的前提下，就可以根据眼睛的位置确定人脸的中心部位，通过“由内而外”的方式来定位人脸，从而提高人脸图片的质量。如图 4-5 所示，原始图片里面包含了很多除人脸外的背景信息，这些信息就在识别上就是噪声。而且除了人脸的中心部位以外，还包括了人的头发、衣领等信息，这些都会提高后面特征提取过程中运算复杂度，增加提取的特征数量增加以及存储空间的消耗。第二三张图是 ASEF 人脸校正后图片。我们可以看出这些图片基本上都是人脸的中心细节，以鼻子为中心，包括眼睛眉毛脸颊和嘴巴，这些信息都是人脸识别中最突出的人脸特征，所以保留这些特征剔除其他特征对识别效果有巨大提升。

2、人脸校正的精度

从图 4-3 可以看出，ASEF 定位人眼的效果非常出色。针对不同的人脸姿态和眼球位置，都可以很好的定位人眼位置。所以，基于眼球位置进行人脸的定位，准确度是可以保证的。

3、人眼定位的宽容度

从图 4-6 我们可以看到，对于人脸有部分遮挡物（这里是眼镜）ASEF 算法也能很准确的定为到人眼。



图 4-6 旋转人脸的人眼定位

4、人脸校正的速度

由于人脸校正的关键步骤——人眼定位部分使用的是 ASEF 滤波器进行人眼定位，而人脸校正因为使用的仿射变换，是一个运算复杂度很低的线性运算，所以整个人脸校正的速度非常之快。在实验室的人脸运算平台上，我们测出的数据显示，人脸定位所需要的时间远远小于人脸检测的时间，CPU 占用率也很低。

表 4-2 人脸检测和 ASEF 人眼定位时间比较

	时间	CPU 占用
人脸检测	32ms	23%
人眼定位	8ms	1%
人脸检测总花费时间		

4.5 本章小结

本章的内容是围绕人脸预处理展开的，去噪和直方图均衡是进行人脸检测前很关键的步骤，而人脸校正是衔接人脸检测和人脸识别的重要步骤，通过人脸校正，我们可以进一步扩大图片中人脸成分的比重，减少运算数据量以及对人脸姿态进行纠正，对人脸识别过程有着重大意义。

对人脸图像的精确注册是人脸识别的第一个重要步骤,而建立人脸注册的一般方法就是找到眼睛。经过长期的研究实验，针对人眼进行人脸校正具有速度快、精确度高的效果，因为人眼特征比较明显，而且本身具备对称性，校正过程会比较简单快速。所以人脸校正的问题就变成了如何进行人眼定位的问题。我们使用 ASEF 滤波器进行人眼的定位。ASEF 人眼定位实际上是通过 ASEF 对样本图片进行处理，ASEF 滤波器是在频域空间对图片进行滤波，现代数字信号处理中，由于有专门的快速傅立叶变换算法，所以 ASEF 滤波器的处理速度非常之快，可以满足实时的需求。

第五章 基于级联分类器的训练算法

5.1 分类器结构

在人脸检测中，我们将构造一个分类器来区分人脸和非人脸。我们将一个子窗口在包含有人脸的图像中进行移动，其将遍历整个图像，然后把此子窗口所遍历得到的图像信息送入到分类器中去，分类器将按照设计的算法去判定每个子窗口中是否包含得有人脸。在众多的被检测子窗口中，绝大多数的窗口都不包含人脸，最后剩下的包含人脸的窗口数量极少，因此我们也可以将人脸检测看作是一个小概率事件。我们在进行分类器设计的时候需要注意以下几个方面问题：

- 1、我们将采用的分类器结构是什么；
- 2、针对被检测的图像我们应该选用怎样的分类器；
- 3、在训练集中怎样有效地训练分类以及在识别中怎样选用分类器；
- 4、分类器的分类正确率能否满足我们的实际要求。

5.1.1 Boosting 原理

Schapire^[51]最先提出了 Boosting 算法^[50]，Boosting 的整个学习算法的框架是基于 PAC（Probably Approximately Correct，可能近似正确）模型的^[52]。

X 表示训练样本的集合， C 表示所有目标的学习结果集合，目标函数被定义在了其上。人脸检测是一个分类结果只可能有两种的分类问题，那么 C 的取值范围为 $\{-1, 1\}$ 。 C 中每个目标 c 对应于 X 的一个等效布尔函数 $c: X \rightarrow \{-1, 1\}$ ，若 x 是 c 的一个正例，即其代表人脸，则 $c(x)=1$ ，若 x 是 c 的一个负例，即非脸，则 $c(x)=-1$ ^[60]。

在 Boosting 中，不要求学习器 L 具有多高的正确分类能力，只要求其能进行近似正确的假设就可以了，也称其为弱 PAC。Boosting 的中心思想就是：将一系列比随机假设性能稍好的弱分类器在样本数据库（需要样本数据足够充足）中进行学习提升，当其满足我们设定的性能指标时，就将其合并为一个性能优秀的强分类器^[60]。

设 h_1, h_2, \dots, h_n 是一系列的假设，其按照各自权值结合起来的总体假设是：

$$f(x) = \sum_{i=1}^n \alpha_i h_i(x) \quad (5-1)$$

其中 α_i 表示假设 h_i 的系数， α_i 和 h_i 的数字同时在每一次的 Boosting 过程中进行更新。其中， α_i 和 h_i 的初始值的选择和随后的更新方法将可以区分不同的 Boosting 算法。

Boosting 基本思路可以描述为以下步骤：

- 1、将训练样本导入到训练集中，此时数据是无规律地分布的；
- 2、根据各训练样本在训练集中的重要程度来初始化个样本的权重；
- 3、将带有初始权重的训练样本输入弱分类器，弱分类器根据其自身的定义来对训练样本进行划分；
- 4、根据弱分类器的分类能力更新其在整个分类数中的权重；
- 5、看是否达到我们设置的性能指标，如果达到就继续下一步，如果没有达到就转到步骤 2，继续更新弱分类器权重；
- 6、将各个弱分类器的权重加权组合起来，构成有强分类能力的强分类器。

5.1.2 分类器级联

要让单一的分类器达到很好的分类能力是非常困难的，就算能够设计出这样的分类器，其实现的难度以及所付出的代价也是很大的。级联的想法是首先构造多个只具有简单分类能力的弱分类器，然后将其结合起来构成具有高分类能力的分类器，其结构简单（都是由简单分类器结合得到），计算复杂度小（只需要计算简单分类器）；但是其性能却相对较差。

本节我们设计的系统是一个简单分类器的集合 $\{h_i\}_{i=1}^n$ ，我们将采用串行的方式将每一个弱分类器（weaker classifier） h_i 级联在一起，如图 5-1 所示。不能预期构造的最佳分类函数是否能很好地训练数据是弱分类器最大的软肋^[60]，所以我们采用级联系统将多个弱分类器以一定权重结合起来提升为一个强分类器。

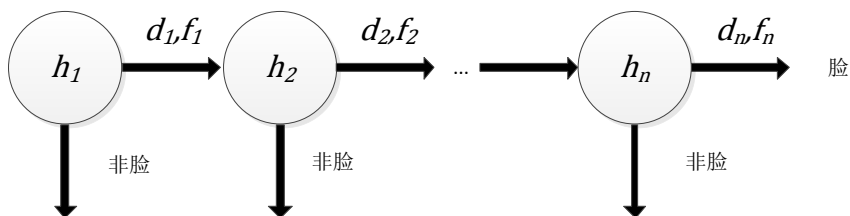


图 5-1 n 节点的级联结构

一个分类系统，评估其分类性能的主要指标是：

正确检测率（correct detection rate, CDR），虚假正确率（false positive rate, FPR）和虚假错误率（false negative rate, FNR）。

我们知道在整幅图像中人脸窗口所占的比重较小，而绝大多数都是“非脸”的，为了减低整个系统的负担，级联结构应该在最开始的几级就将非人脸的子窗口从系统中排除，而用剩下的分类器去对“人脸”子窗口进行进一步的筛选。

所以在设计弱分类器 h_i 时要求几乎能够全部检测到人脸，同时也允许将非人脸错误地判定为人脸。如果图像子窗口在级联系统第 i 级被分类为“脸”，则将图像子窗口从 h_i 传递到下一个节点 h_{i+1} ，否则就判断为“非脸”，设 h_i 的检测率为 d_i ，假阳率为 f_i ，那么整个系统的检测率为 D ，假阳率为 F ，则：

$$\begin{cases} D = \prod_i^n d_i \\ F = \prod_i^n f_i \end{cases} \quad (5-2)$$

在一个拥有 10 级结构的级联系统中，假设 $d_i=0.999$ ， $f_i=0.5$ ，那么 $D \approx 0.98$ ， $F = 9.76 \times 10^{-4}$ ，按照上面的检测指标来说，这个分类器拥有很高的分类性能。寻找有效级联的方法和较好的弱分类器是设计级联系统的最重要的工作。

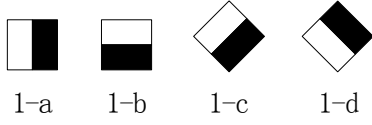
5.2 弱分类器设计

分类器自身的结构决定了其进行分类所需花的时间，而在人脸系统中，实时性又是我们所必须要关注的问题，所以我们就要使得分类器的结构不能太复杂，同时其分类精度还不能降低。

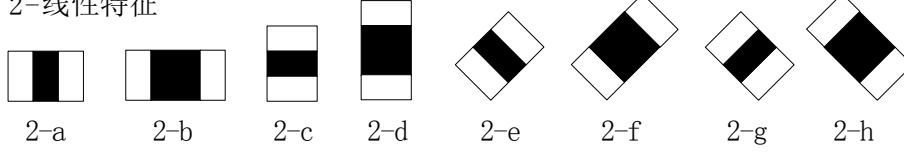
5.2.1 矩形特征选取

矩形特征对于人脸图像的边缘、线段等基本的图形结构效果很好，也就是说可以用矩形特征来匹配人脸图像的眼镜，鼻子，嘴巴等几何结构，最经典的几何特征是由 Viola 提出的一种 Haar-like 矩形特征，其原始模型最初只有五种^[9]；Lienhart 根据需要增加了旋转的 Haar-like 特征^[53]，图 5-2 是目前通用的 Haar-like 特征模型：

1-边缘特征



2-线性特征



3-中心环绕特征

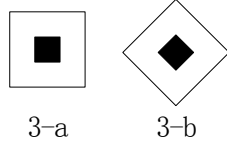


图 5-2 Haar-like 扩展特征

5.2.2 LBP 子模式特征提取

首先提取出原始图像的 LBP 特征,原始的人脸 LBP 特征是在高维空间 R^D 中,我们用 $X = \{x_1, x_2, \dots, x_N\}$ 来表示,我们需要将这些样本投影到一个相对低维的特征空间 R^d ($d \ll D$), 假设投影矩阵为 A , 样本在低维空间中被表示为 $Y = \{y_1, y_2, \dots, y_N\}$, 其中 $y_i = A^T x_i$ 。提取流程如下:

step1:

构建原始数据的近邻图。根据给定的数据集 X , 构建一个带有 N 个顶点的带权图 G , 对于是近邻的两个样本 x_i 和 x_j , 我们在其间建立一条边, 对于其近邻关系的度量我们采用阈值 ε , 如果 $\|x_i - x_j\|^2 < \varepsilon$, 则在 x_i 和 x_j 之间建立一个连接; 还有一种是采用像素 K 近邻准则, 若两个定点是 K 近邻的话就建立连接。

step2:

初始化权重。构造一个 $N \times N$ 的稀疏对称矩阵 W , 如果 x_i 和 x_j 存在连接, 则计算其连接的权值 w_{ij} , 否则就将 w_{ij} 设为 0, 有如下权值计算表示方式:

$$w_{ij} = \begin{cases} 0 & \text{otherwise} \\ \exp(-\|x_i - x_j\|^2 / 2\sigma) & x_i \text{ 和 } x_j \text{ 近邻} \end{cases} \quad (5-3)$$

或则,

$$W_{ij} = \begin{cases} 0 & otherwise \\ 1 & x_i \text{ 和 } x_j \text{ 近邻} \end{cases} \quad (5-4)$$

其中， σ 为用户根据训练情况自定义。

step3:

特征约减。首先根据式 5-5 计算新特征空间的特征值和特征向量：

$$XLX^T a = \lambda XD X^T a \quad (5-5)$$

其中 D 是 N 维对角矩阵， $D_{ii} = \sum_j W_{ji}$ ， $L = D - W$ 为拉普拉斯矩阵，假设 a_1, a_2, \dots, a_l 是公式 5-5 的特征向量，其对应的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_l$ ，得出的特征投影矩阵 $A = [a_1, a_2, \dots, a_l]$ 。从而降维得到的低维特征 $y_i = A^T x_i$ ， y_i 是 l 维向量。

弱分类器由通过 LPP 降维后的 LBP 特征向量构成，因此我们将弱分类器构造为：

$$h_j(x) = \begin{cases} 1 & f_j g(x) = f_j h_j \\ 0 & otherwise \end{cases} \quad (5-6)$$

其中 $g(x) = a_i^T x$ ， a_i 是特征向量， x 是训练样本集， h_j 是第 j 个弱分类器的阈值， f_j 是方向向量。函数 $g(x)$ 将所有样本都投影到新的特征空间中，可得到 N 个特征值，选取其中最合适的作为弱分类器的阈值。

5.3 Adaboost 构建强分类器

Adaboost 是 Adaptive Boosting 的简称，其是一种自适应的提升算法，对于弱学习算法返回的弱假设误差，其能够进行自适应的调节^[60]。Freund^[54]等在 1995 年提出了最初的 Adaboost 算法，Adaboost 的训练算法，采用循环的方式，直到最后输出的分类器达到了所要求性能指标才停止循环。

Adaboost 用一个特定的学习算法来循环更新整个分类器。整个提升算法的详细流程如下：

- 1) 对于样本集中 n 个样本 $\{x_1, y_1\}, \dots, \{x_n, y_n\}$ ，其中 $x_i \in X, y_i \in Y = \{-1, 1\}$ ，假设 n 个样本中有 p 个正样本和 q 个负样本。
- 2) 初始化，对于 $y_i = 1, -1$ ，权重分别赋予 $D_{1,i} = 1/2p$ 和 $1/2q$

3) 对于 $t=1, \dots, T$

a) 权值归一化: $D_{t,i} \leftarrow \frac{D_{t,i}}{\sum_{j=1}^n D_{t,i}}$, $D_{t,i}$ 可以看作一个概率分布算子

b) 对于每一个特征 j , 训练其对应的分类器 h_j , 设与 $D_{t,i}$ 相关的误差为 ε_t ,

$$\text{计算方式为 } \varepsilon_t = P_{i \sim D_t}[h_j(x_i) \neq y_i] = \sum_{i: h_j(x_i) \neq y_i} w_j(i) = \sum_i D_i |h_j(x_i) - y_i|$$

c) 选择具有最小的 ε_t 所对应的 h_t

d) 更新权重: 令 $\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$, 则: $D_{t+1,i}(i) = D_{t,i} \beta_t^{1-e_i} = \begin{cases} D_{t,i} \beta_t & h_t(x_i) = y_i \\ D_{t,i} & h_t(x_i) \neq y_i \end{cases}$, 当

分类正确时 $h_t(x_i) = y_i, e_i = 0$; 当错误分类时 $h_t(x_i) \neq y_i, e_i = 1$

e) 输出最后的假设

$$H(x) = \text{sgn}\left(\sum_{t=1}^T a_t h_t(x)\right) = \begin{cases} 1 & \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{n=1}^T a_t \\ -1 & \text{otherwise} \end{cases}$$

其中 $a_t = \ln\left(\frac{1}{\beta_t}\right) = \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)$ 。

上述流程中 h_t 是弱假设; H 是最终输出的假设, H 是 T 个权重为 a_t 的弱假设 h_t 加权求和所得出的强假设; a_t 是 h_t 的权值; 训练误差用 ε_t 来表示; h_t 的概率分布函数是 D_t 。

我们用弱分类器的误差来度量其是否能很好的找到弱假设 h_t 来表述概率分布 D_t :

$$\varepsilon_t = P_{i \sim D_t}[h_t(x_i) \neq y_i] \quad (5-7)$$

一旦弱学习假设 h_t 成立, Adaboost 就选择一个参数 β_t , β_t 直接和 a_t 有关, 而

a_t 是 h_t 的权重。 $\varepsilon_t \leq 1/2$, 则 $\frac{\varepsilon_t}{1 - \varepsilon_t} < 1, \beta_t < 1$, 即减小了正确分类样本在整个训练

过程的权重, 而且随着 ε_t 越小, β_t 就越小; 同时, $a_t \geq 0$, 并且 ε_t 越小, a_t 越大。

在整个系统初始化时, 将所有的样本的权值都归一。在每次循环过后, 更新所有样本的权重(对于能够被系统正确分类的样本减小权重, 反之则增加其权重)。

这样是为了使得不能被正确分类的样本能进行“再学习”，以使得分类器最终能正确分类。各个弱分类器按照其在系统中的权重进行“投票”，然后按照一定的规则线性地合成最终的强分类器。

$$H(x) = \begin{cases} 1 & \text{sgn}(\sum_{t=1}^T a_t h_t(x)) > 0 \\ -1 & \text{otherwise} \end{cases} \quad (5-8)$$

5.4 本章小结

本章首先讨论了 Adaboost 算法中分类器的结构，然后对于基于矩形特征的 Haar-like 特征和基于纹理的 LBP 特征进行了介绍。Haar-like 特性因为其特征简单，运算速度较快（因为“积分图”的存在）得到了广泛的应用，但是其对于光照和姿势变化的鲁棒性较差，所以我们提出了基于 LBP+LPP 特征的弱分类器构造方法，其对于光照变化和人脸旋转都具有一定的鲁棒性。最后介绍了 Adaboost 算法将弱分类器提升为强分类器的流程。

第六章 基于 LBP 和 Adaboost 的人脸系统

人脸技术广泛应用于门禁、视频监控、图像检索等领域，主成份分析（PCA，Principle Component Analysis）、Adaboost、支撑矢量机（SVM，Support Vector Machine）是基于人脸统计模型的经典方法。PCA 方法是由 Turk 和 Pentland 于 1991 年引用到人脸检测和识别领域的，该方法是通过 K-L（正交）变换将人脸图像从高维空间变化到由一组正交向量构成的低维特征空间中，该方法识别速度快，实现简单，但是就是检测、识别率不高^[25]。Viola 等于 2002 年提出了一种基于 Adaboost 的人脸检测算法，其创新的提出使用“积分图像”的方法来快速的计算图像的多尺度 Haar 特征，将最优的弱分类器按照一定权重结合在一起“提升”为强分类器，该算法的特点是简单，实时性好，但是该算法存在训练样本需求较大，训练时间长等问题，而且对于小样本的分类能力十分低下^[24]。支撑向量机（SVM，Support Vector Machines）是 Vapnik 等提出的基于结构风险最小化原理（SRM，Structural Risk Minimization Principle）的统计学习理论^[48]，用于分类与回归问题，SRM 使得 VC 维数的上限最小化，因为 SVM 有很好的分类性能，所以 Osuna 等将 SVM 方法用于人脸检测^[49]。

上述算法可以在环境条件很好的情况下得到较好的人脸检测结果，但是对于复杂光照或者人脸旋转等情况，其性能就急剧下降，近年来，一些基于非线性子空间方法（LLE，Isomap，Laplacian Eigenmaps）也被用于人脸检测，虽然其能更好的对于人脸图像进行非线性建模，但是这些方法都存在计算量较大，不能满足实时性的缺点^[41-44]。

本文提出的基于 LBP+LPP+Adaboost 方法进行人脸检测/识别，其继承了 LBP 特征旋转不变和灰度鲁棒的特点，并且结合 LPP 算法降低算子的维数，也在一定程度上保留了人脸图像的非线性流形特征，但是因为 LPP 算法是在线性空间中的，所以较上述非线性算法在计算复杂度和计算速度上面都要快。整个算法我们采用的是 Adaboost（Adaptive Boosting），将经过 LPP 降维后的 LBP 算子作为弱分类器，然后将多个弱分类器提升（boosting）合并为一个强分类器，最后采用级联的方法将强分类器合为最终的检测分类器。大致算法流程如下：

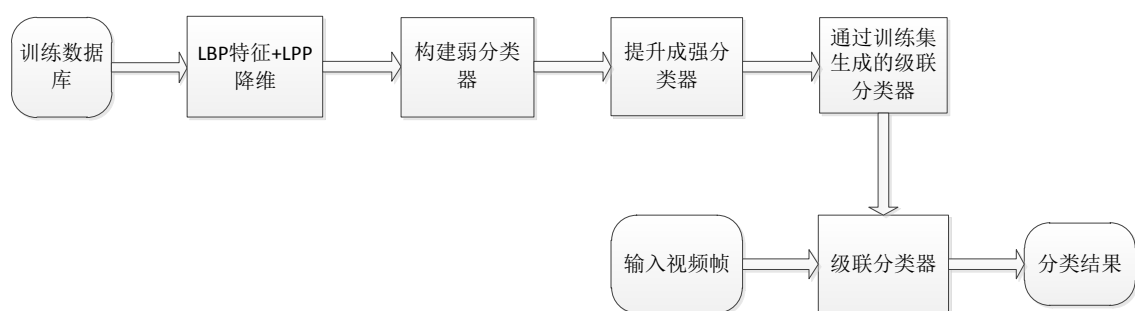


图 6-1 人脸训练和检测算法流程图

6.1 算法简介

在评估了识别率和实时性两方面的需求后，我们选取 LBP+Adaboost 的方法来生成强分类器，并将其用于人脸检测。

图像预处理我们首先采用 Gamma 校正来减弱光照的影响，然后对图像进行直方图唯一化，最后使用 ASEF 算法对人脸进行校正，这一步为下面的人脸特征提取提供了尽可能标准的环境。

人脸特征提取主要由以下步骤构成：

步骤 1、选取合适的滤波器（如高斯滤波器，拉普拉斯滤波器）对待识别图像进行滤波分解，形成不同大小、不同分辨率的各层图像，然后按照下层图像比上层图像尺寸较大、分辨率较高的原则构建图像金字塔。具体包括以下步骤：

步骤 1-1、将原始图像表示为 $I(x, y)$ ， $x \leq M$ ， $y \leq N$ ，其中 M ， N 分别表示图像的行数和列数； G_i 表示图像金字塔的第 i 级，则将 G_0 表示为：

$$G_0(x, y) = I(2^{-\text{expand}}x, 2^{-\text{expand}}y) \quad (6-1)$$

上述公式中 expand 表示采样扩展因子，这里 expand 取值 -1，表示对图像进行下采样。

步骤 1-2、为了构造金字塔的第 i 层 G_i ，先将 G_{i-1} 与一个具有低通特性的滤波器 $W(m, n)$ 进行卷积操作（这里的滤波器可以取以下几种：（1）高斯滤波器；（2）DOG 滤波器；（3）小波多尺度滤波器（其包含有所有小波分析中的尺度滤波器）），再把卷积结果作隔行隔列的下采样，如下面公式所示：

$$G_i(x, y) = \sum_m \sum_n W(m, n) G_{i-1}(2x+m, 2y+n) \quad (6-2)$$

式中， $0 < i < N$ ， N 为金字塔的总层数， $0 < x < C_i$ ， $0 < y < R_i$ ，将金字塔的第 i 层

图像的列数设为 C_i ，行数为 R_i ，窗口函数用 $W(m,n)$ 表示。

步骤 1-3、将图像金字塔的各级图像设为 G_1, \dots, G_N ，因为我们选用的是分辨率为 128×128 的人脸图像，所以通过下采样过后所得到的图像层次较多，后续的计算量较大，所以在满足准确性的我们选取其中的 G_k, G_{k+1}, \dots, G_N 来构成所需要的图像金字塔（这里 N 的实际取值为 8，所以 K 我们取 3）。整个过程可以用简化为下式：

$$G_k = C_{(\downarrow 2)} G_{k+1} \quad (6-3)$$

$C_{(\downarrow 2)}$ 表示平滑算子，其中 “ \downarrow ” 表示下采样，“2” 表示采样率。

步骤 2、采用 LTP 算子计算各尺度图像的 LTP 特征谱，并且构建具有多尺度特性的 LBP 金字塔，具体步骤如下：

步骤 2-1、用下述 LTP 算子去遍历各尺度的图像，得到各个尺度图像的 LTP 特征谱

$$s(u, i_c, t) = \begin{cases} 1 & g_p \geq g_c + t \\ 0 & |g_p - g_c| < t \\ -1 & g_p \leq g_c - t \end{cases} \quad (6-4)$$

$$LTP_p^R = \sum_{p=0}^{P-1} s(g_p - g_c, t) \quad (6-5)$$

其中， R 表示 LTP 特征的半径， P 表示 LTP 特征的采样点数， g_c 表示 LTP 特征中心像素值， g_p 表示 LTP 特征中心像素四周的像素值。 t 是自定义的一个阈值，使得 LTP 编码对于噪声更有抑制效果。

步骤 2-2、因为 LTP 算子的编码特性使其对于灰度值的转换不再严格，所以将 LTP 编码分成正值和负值两个部分，正负值各是一个完整的 LBP 编码。因为正负值所分别产生的 LBP 编码特性是一致的，所以下面步骤对于 LBP 特征的处理是通用的。

步骤 2-3、将各个尺度图像的 LBP 特征图按照原始图像在图像金字塔中的位置排列，所得到的图像简称 LBP 金字塔。

步骤 3、将各尺度图像的 LBP 特征图划分为单独的多个区域，各自统计各区域的 LBP 直方图，然后再分别合成每一级图像的 LBP 直方图。这里将图像进行

分区块的处理是为了充分利用 PC 平台的并行处理能力, 提高特征的提取速度(当然也可以将整幅图像的 LBP 特征进行直方图统计, 这就视需求而定)。具体步骤如下:

步骤 3-1、对于待分类的图像, 将其等分为互不重叠的多个区域。

步骤 3-2、设第 i 幅图像第 j 分块为 $f_{ij}(x, y)$, 采用 LBP_p^R 算子计算后做直方图统计, 其 LBP 模式直方图可以表示为公式:

$$H_{ij}(m) = \sum_{x,y} I\{f_{ij}(x, y) = m\}, m = 0, \dots, B-1 \quad (6-6)$$

LBP 模式总的数量用 B 来表示, B 同时也是统计直方图横轴的长度。 m 表示一个 LBP 模式, 他的取值完全决定于选用的 LBP 算子, 如采用 LBP_8^1 算子, 有 8 个采样点, 它的取值为 1~256。 I 是这样一个函数:

$$I(A) = \begin{cases} 1 & A = \text{ture} \\ 0 & A = \text{false} \end{cases} \quad (6-7)$$

步骤 3-3、将各个分块图像的 LBP 直方图结合成为整层图像的 LBP 特征谱。

步骤 4、将各尺度的 LBP 模式直方图分别进行投影变换得到各尺度的 LBP 特征的约简模式, 简称为 LBP 子模式特征, 具体步骤如下:

步骤 4-1、将每一级图像的 LBP 模式直方图 H_i 看成一个向量, 并把所有 H_i 合并为一矩阵 H , 则 H 为一 $B \times N$ 的矩阵。 B 为直方图的 bin 数, N 为图像金字塔的层数。

步骤 4-2、求第 i 级图像 H_i 的协方差矩阵 \sum_{H_i} 的本征值和本征向量, 选取 m 个最大的本征值对应的本征向量构成变换矩阵 A_i 。多尺度空间中的图像由多少层, 则需要取几个 A_i 。

步骤 4-3、使用 LPP 方法将直方图矩阵 H 通过线性变换 $\bar{H} = AH$ 投影到 \bar{H} 上, 对 H_i 降维, 使得 \bar{H}_i 的 bin 数远小于 H_i , $A = \{A_1, A_2, \dots, A_N\}$ 。

步骤 4-4、将 m 个最大的本征值对应的本征向量构成的空间称为 LBP 模式子空间, \bar{H} 是 H 在 LBP 模式子空间里的投影, 因此称 \bar{H} 为 LBP 子模式。

步骤 5、将各区域的 LBP 子模式特征连接起来得到所求的多尺度 LBP 子模式序列, 称之为多尺度 LBP 子模式特征, 具体方法如下: 使用公式 $\bar{H} = \{\bar{H}_1, \bar{H}_2, \dots, \bar{H}_N\}$ 将 N 层 LBP 子模式结合起来, 形成多尺度 LBP 子模式特征。

步骤 6、将前面 LTP 特征拆分后所形成的两个多尺度 LBP 子模式特征结合为统一的 LTP 子模式特征, $LTP = \{LBP_{\text{正}}, LBP_{\text{负}}\}$ 。

通过以上步骤，就能得到用于识别和区别不同人脸图像的多尺度 LBP 子模式特征，然后将特征送入数据库中进行比对，最后得到识别结果。

需要说明的是：

步骤 2-2 中 LTP 特征拆分为的正负两个部分的 LBP 特征，将在各自完成步骤 5 后再合并在一起，这将比直接使用 LBP 特征能够有更好的噪声抑制能力，对于一些复杂光照条件下的人脸有更高的识别能力。

步骤 3 中每层图像的分块对于不同层的图像可以分为不同数量的子块，这样可以更方便的对全局特征和局部特征进行统计，也避免了一些不必要的计算开销。将 LBP 特征谱分成多少块独立的区域，这是一个需要特别重视的问题，其对识别率的影响较大。但无论固定哪一级特征谱的分块，所取得的识别率都要明显好于各级分块同时变化所取得的识别率。还应该注意过多的分块并不会有助于识别率的提高。

最后我们采用最小欧几里德距离作为人脸相似度的判定标准。

具体各部分采用算法由下表给出：

表 6-1 系统各部分使用算法

图像预处理	Gamma 光照校正，直方图归一化
人脸检测	Adaboost+LBP
人眼检测/人脸校正	ASEF
特征提取	LBP/LTP+LPP
特征比对	最小欧几里德距离，最近邻法

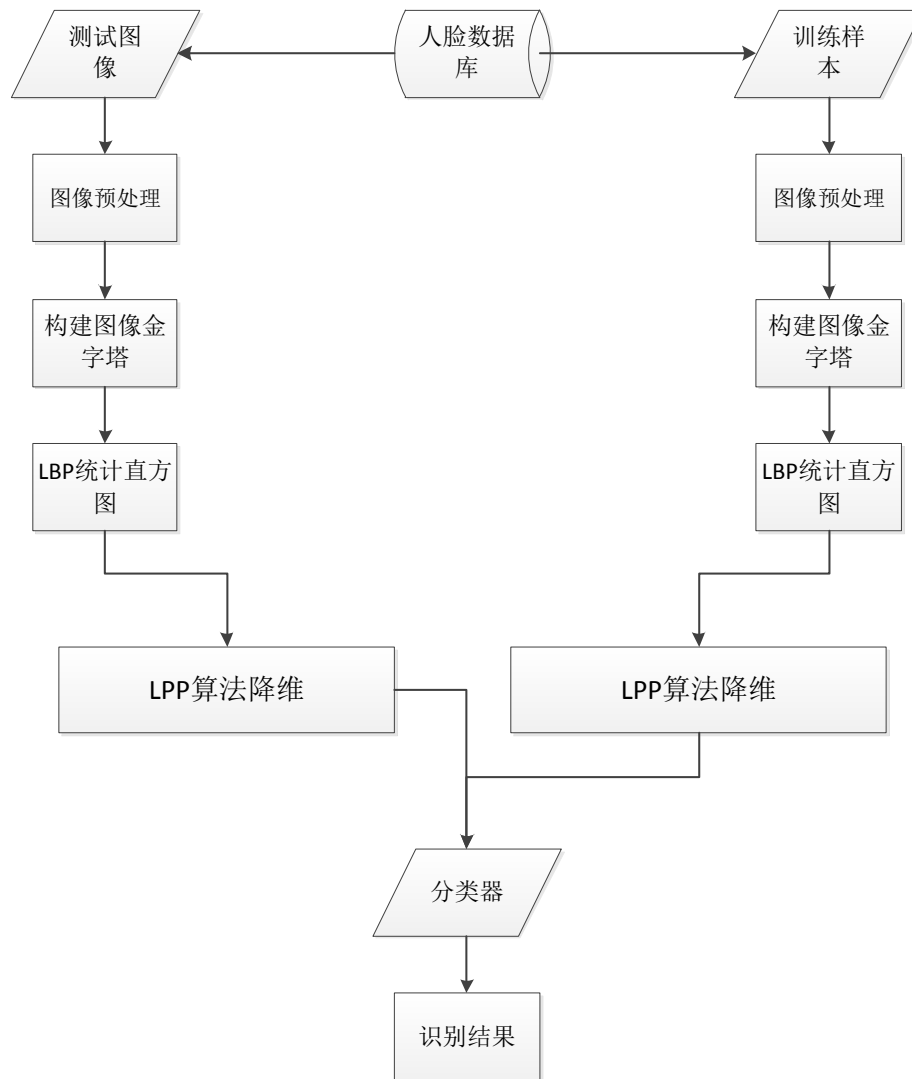


图 6-2 人脸数据库识别算法测试流程

6.2 实验结果与分析

我们在 ORL 标准人脸数据库、FERET 标准人脸数据库和中科院的 CAS-PEAL 数据库中分别进行了测试和比较。我们分别将原始 PCA^[25], PCA+FLDA^[34], LBP^[22], Uniform LBP^[24]同本文的算法做了比较, 在实验中基于 PCA 的方法保留了所有的特征份量; 基于 PCA+FLDA 的人脸识别方法中, 我们用 PCA 将原始数据约减到 80 维, FLDA 保留所有特征份量。

6.2.1 ORL 人脸数据库算法结果

ORL 数据库包含 400 幅人脸图片，其中共有 40 人，每人 10 幅，包括了他们不同表情、不同姿态的图像^[56]。图像尺寸为 92×112 ，我们分别将每人的 1 幅、3 幅、5 幅图像作为训练样本，其余的都作为测试样本，实验结果如表 6-2:

表 6-2 ORL 人脸库中各算法测试数据

	一个训练样本	三个训练样本	五个训练样本
	识别率	识别率	识别率
LBP+LPP+Adaboost	82%	93%	98.2%
$LBP_{8,1}^{riu2}$	78%	86%	93%
LBP_8^1	76%	92%	96%
PCA+FLDA		81%	88%
PCA	68%	82%	90%

6.2.2 FERET 数据库算法结果

由美国 FERET 小组创建的 FERET 人脸数据库被广泛的作为训练和测试数据库。FERET 人脸库分为训练集和测试集，其训练集又分为 fa（正常人脸表情）和 fb（变化的人脸表情），测试集分为 fb，fc，Dup I 和 Dup II。我们采用的是 fb（不同表情人脸图像）和 fc（不同光照条件下的人脸图像）两个测试集^[57]，如表 6-3:

表 6-3 FERET 数据库算法测试结果

算法	fb 中识别率	fc 识别率
LBP+LPP+Adaboost	98%	88%
$LBP_{8,1}^{riu2}$	96%	78%
LBP_8^1	94%	76%
PCA+FLDA	92%	73%
PCA	89%	66%

6.2.3 CAS-PEAL 数据库算法结果

CAS-PEAL 人脸数据库的原始图像为 640×480 图像, 24 位 RGB 色彩的 BMP 位图图像, 我们采用的是其简化库——CAS-PEAL-R1, 其全部在原始图像的基础上做的处理, 图片被剪裁为 360×480 像素, 并且将图像都转为灰度图像, 图像格式同样是 BMP, 其也包括训练集合, 原型图像集合, 测试集合, 测试集总共有 6992 幅图像, 分为表情变化、光照变化、饰物变化、背景变化、距离变化、时间跨度变化等 6 个子集^[58]。数据库图例如图 6-3 所示, 测试结果如表 6-4 所示:

表 6-4 CAS-PEAL 数据库算法测试结果

识别率	LBP+LPP+Adaboost	$LBP_{8,1}^{riu2}$	LBP_8^1	PCA+FLDA	PCA
光照组	86%	64%	72%	62%	58%
表情组	96%	86%	92%	68%	64%
装饰组	96%	92%	94%	71%	66%
背景组	99%	95%	96%	88%	86%

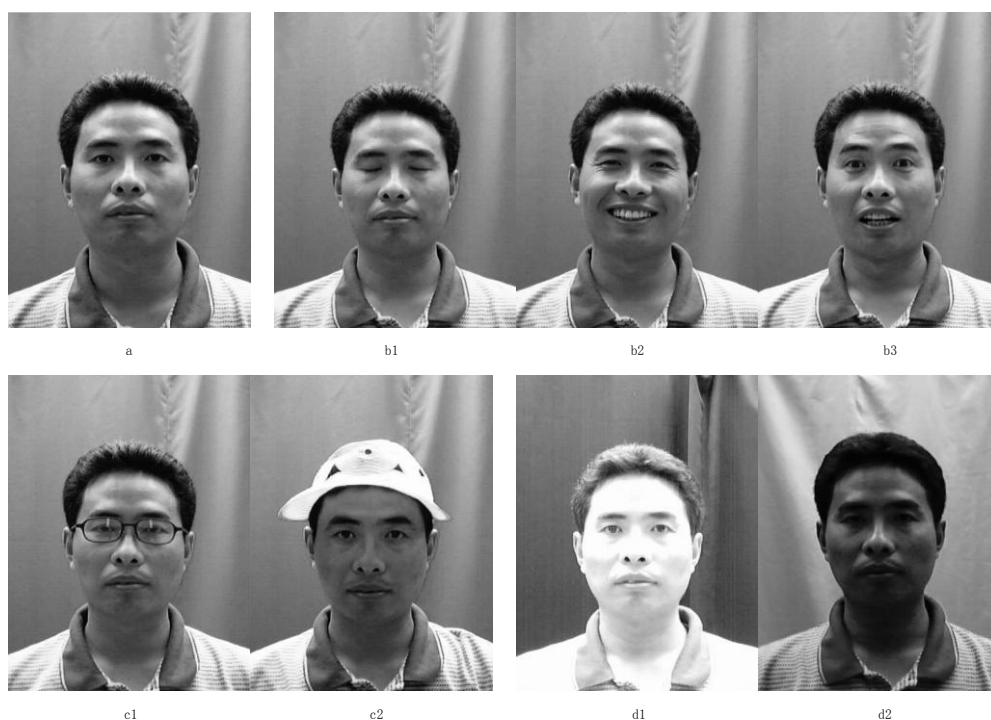


图 6-3 CAS-PEAL 人脸数据库事例

(a 为普通组; b1, b2, b3 为表情组; c1, c2 为装饰组; d1, d2 为背景组)

从上面的几个数据集的测试结果我们可以看出，在光照、表情变化不大的情况下，PCA 和 PCA+FLDA 算法都能得到不错的识别率，主要是因为 PCA 算法非常依赖训练集和测试集的灰度相关性，其要求训练和测试图像外部环境应相识，所以其对遮挡物，表情，光照等变化很敏感，而本文提出的基于 LBP+LPP+Adaboost 算法，结合了 LBP 算子对于光照变化的鲁棒性和 LPP 的流形特点，所以对于光照变化或表情人脸都有很好的识别率。

6.3 基于 OpenCV 的系统原型实现

OpenCV 是很棒的一个开源计算机视觉库。它又 Intel 公司出资赞助的，现在已经更新到了 2.3 版本，其最大的改变是将以前版本中的 C 函数用 C++ 类来代替，具有了更好的封装性和继承性；其将图像处理和计算机视觉中的很多通用算法进行了演示。OpenCV 提供了一系列的 API 以供开发人员能够快速开发出原型系统，而研究人员只需要完善自己的算法，再拿到基于 OpenCV 构建的系统进行算法验证即可^[59]。



图 6-4 ver1.0 人脸识别系统示意图

如图 6-4 所示的是我们设计的人脸识别系统 1.0 版本，其是基于导入静态图片的人脸识别，其主要包括人脸图片训练，待识别图片读取，用户注册，用户识别，退出等五个部分。



图 6-5 ver2.0 人脸识别系统

如图 6-5 所示，2.0 版本中我们设计的是基于视频流的集人脸注册、识别于一体人脸系统。

6.4 本章小结

本章首先介绍了系统整个的算法流程图以及算法实现细则；然后在 ORL、FERET 和 CAS-PEAL 数据库中对于算法进行了测试，并分析了各算法的测试结果；最后大致介绍了我们所设计的人脸识别识别系统。

第七章 总结与展望

7.1 全文总结

人脸不同于其生物特征，其具有直接性、唯一性、方便采集等特点，但是又由于人脸的可塑多变和在采集过程中受诸多外部因素的影响，使得人脸的机器识别这一课题非常具有挑战性。现有算法的显著不足在于对环境的变化鲁棒性较差，尤其是与人脸识别的能力相比而言，自动人脸识别对环境和姿态的变化适应性能低，现有的方法大多数情况下是要求待检测人脸在规定环境下采集，这显然不满足实用性的要求。本文对经典的人脸识别算法进行了研究和改进，在预处理，特征采集，特征约简几个方面做了相应的研究，并且初步实现了一个基于视频的人脸识别系统原型。

本论文所做的主要工作如下：

1. 研究了 LBP 算子的几种扩展形式，包括统一模式，旋转不变的统一模式和改进为三值形式的 LTP 模式。通过研究，可以看到 LBP 算子所具有旋转不变和灰度鲁棒这两个特性对于人脸识别来说很重要，同时 LBP 也具有描述图像局部特征和全局特征的能力，这对于人脸识别来说是也非常有利的；本文还将 LBP 的三值扩展形式 LTP 模式运用到人脸识别中来，其相较于传统 LBP 模式对于局部的光照变化和噪声有更高的鲁棒性。

2. 深入研究了 LBP 原始模式的降维算法。因为 LBP 统一模式本身所存在的一些缺陷，所以我们用降维生成的 LBP 子模式来代替 LBP 统一模式；本文比较了局部保持投影（LPP）、主成份分析（PCA）和线性判别分析（LDA）这三种降维算法在人脸领域的应用，最终选取的是 LPP 模式在构建 LBP 子模式。

3. 研究了人脸图像的预处理方法，特别是人脸校正。本论文方法是：先通过平均精确合成滤波器（ASEF）取得人眼的位置然后再通过仿射变化对人脸进行旋转，最后得到近似标准的正面人脸。ASEF 定为人眼的精度较 Haar-like+Adaboost 方法高很多，并且其对于戴了眼镜或者有遮挡的人脸的鲁棒性也更强，最重要的是其速度非常快，几乎在检测到人脸的同时就能定位人眼。

4. 基于 OpenCV 开源库构造了一个基于视频的人脸识别系统原型。并在上面对于算法进行实际的测试。

7.2 后续工作展望

人脸识别是模式识别邻域的一个热点和难点问题，影响其识别率的因素有很多，本文仅仅是对某一些方面进行了研究和改进，所以后续还有很多的工作需要继续去完成。后续工作主要是在一下方面：

1. 虽然本文提出的 LBP+LPP+Adaboost 算法在各种人脸识别率有一定的提高，但是对于光照和人脸表情变化较大的场合仍然识别率偏低。所以后续的工作是：针对光照变化大的问题可以通过近红外人脸图像来进行彻底的规避，因为近红外人脸图像能够完全消除外部自然光照的影响，其能大大的提高算法的识别率；还有就是通过构造 3D 人脸模型来弥补 2D 人脸识别所无法解决的光照和人脸本身所存在的问题。

2. 进一步完善人脸识别系统。在继续对算法进行改进的基础上，进一步的完善我们所构造的人脸识别系统原型，使得其具有更加丰富的功能。

致 谢

首先，我必须得向我的导师解梅教授致以诚挚地感谢。解老师就好像一盏指路明灯，在学习、生活、科研上都给予了我细致的指导。在撰写最后学位论文的过程中，解老师还不忘提醒我们要时刻保持一颗严谨的心。在我的研究生生涯中，解老师以身作则，教给了我很多东西，这必将成为我在今后人生道路上的宝藏。从解老师这里，我最重要的是学会了如何去学习。在本文的最后，谨向解老师致以我最真诚的谢意和发自心底的敬意。

其次，必须要感谢魏云龙师兄，在我刚刚进入人脸识别课题组时他给予我很多的指导。感谢同课题组的鲍璘和杨占栋在平时枯燥的科研中与我同在，特别还要感谢一下王佳同学在平时生活和科研中对我的鞭策。最后要感谢我的家人，是你们默默的付出，赋予了我现在全部。

感谢所有出现在我生命中人们。

参考文献

- [1] 王映辉. 人脸识别-原理、方法与技术. 北京: 科学出版社, 2010, 3-4
- [2] 田捷, 杨鑫. 生物特征识别技术理论与应用. 北京: 电子工业出版社, 2005, 6-8
- [3] 雷玉堂. 面相识别与其他生物识别技术的比较. 第十一届中国国际公共安全博览会-行业应用, 2007, 128-134
- [4] M.A. Turk, A.P. Pentland. Face recognition using eigenfaces. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991, 586-591
- [5] A.M. Martinez, A.C. Kak. PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 31(5):669-682
- [6] H.A. Rowley, S. Baluja, T. Kanade. Neural network-based face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(1)23-38
- [7] A.V. Nefian, M.H. Hayes. Face detection and recognition using hidden Markov models. International Conference on Image Processing, 1998, Vol. 1:141-145
- [8] E. Osuna, R. Freund, F. Gorosi. Training support vector machines: an application to face detection. IEEE Conference on Computer Vision and Pattern Recognition, 1997, 130-136
- [9] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, 511-518
- [10] B. Moghaddam, T. Jebara, A. Pentland. Bayesian face recognition. Pattern Recognition, 2000, 33(11):1771-1782
- [11] T.F. Cootes, C.J. Taylor. Active shape models-smart snakes. Proceeding of the British Vision Conference, September, 1992, 266-275
- [12] T.F. Cootes, C.J. Taylor. Active shape models-their training and application. [J]. Computer Vision and Image Understanding, 1995, 61(1):38-59
- [13] A. Hill, T.F. Cootes, C.J. Taylor. Active shape models and the shape approximation problem. Image and Vision Computing, 1996, Vol. 14:601-607
- [14] G.J. Edwards, T.F. Cootes, C.J. Taylor. Face recognition using active appearance models. Proceedings of the 5th European Conference on Computer Vision, 1998, Vol. 2:581-595
- [15] T.F. Cootes, G.J. Edwards, C.J. Taylor. Active appearance models [J]. Pattern Analysis and Machine Intelligence, 2001, 23(6):681-685

- [16] T.F. Cootes, C.J. Taylor. Statistical models of appearance for computer vision. World Wide Web Publication, 2001, 69-73
- [17] Xiaoming Liu. Discriminative face alignment. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(11), 1941-1954
- [18] A.S. Georgiades, P.N. Belhumeur, D.J. Keigman. From few to many: illumination cone models for face recognition under variable lighting and poses. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6):643-660
- [19] A. Shashua, T. Riklin Raviv. The quotient image: class based re-rendering and recognition with varying illuminations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2):129-139
- [20] R. Basria, D. Jacobs. Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(2):218-233
- [21] Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary patterns[C]. In: Proc.of 8th European Conference on Computer Vision, Prague, Czech, 2004:469-481.
- [22] T. Ojala, M. Pietikainen, D. Harwood. A comparative study of texture measures with classification based on feature distributions [J]. Pattern Recognition, 1996, 29(1):51-59.24
- [23] X. Qing, Y. Jie, D. Siyi. Texture segmentation using LBP embedded region competition [J].Electronic Letters on Computer Vision and Image Analysis, 2005, 5(1):41-47.
- [24] T. Ojala, M. Pietikainen, T. Maenpaa. Multiresolution gray scale and rotation invariant texture classification with local binary patterns [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, 24(7):971-987
- [25] M. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, 1991.
- [26] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Trans. Pattern Analysis and Machine Intelligence, 1997, 19(7) :711-720.
- [27] Xiaoifei He and Partha Niyogi, "Locality Preserving Projections", Advances in Neural Information Processing Systems, Vancouver, Canada, 2003.
- [28] Xiaoyang Tan; Triggs,B. Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions[C].Proc of Analysis and Modelling of Faces and Gestures Conference.[S.l.]:Springer,2007:168-182.
- [29] T. Maenpaa, M. Pietikainen. Texture analysis with local binary patterns. Handbook of Pattern

- Recognition and Computer Vision [M]. 3rd, World Scientific, 2001:711-736.
- [30] L.Wang and D.He. Texture classification using texture spectrum. *Pattern Recognition*, 23:905–910, 1990.
- [31] T.Ahonen, A.Hadid, and M.Pietikainen. Face recognition with local binary patterns. In *Proc. European Conference on Computer Vision ECCV'04*, pages 469–481, Prague, 2005.
- [32] Xiaofei He and Shuicheng Yan Face Recognition Using Laplacianfaces. *IEEE Trans on Pattern analysis and machine intelligence*, Vol 27, No3, March 2005
- [33] A. U. Batur and M. H. Hayes, “Linear Subspace for Illumination Robust Face Recognition”, *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Hawaii, Dec. 11-13, 2001.
- [34] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, No. 7, 1997, pp. 711-720.
- [35] R. Gross, J. Shi, and J. Cohn, “Where to go with Face Recognition”, *Third Workshop on Empirical Evaluation Methods in Computer Vision*, Kauai, Hawaii, December 11-13, 2001.
- [36] A. Levin and A. Shashua, “Principal Component Analysis Over Continuous Subspaces and Intersection of Half-Spaces”, *Proc. of the European Conference on Computer Vision (ECCV)*, Copenhagen, Denmark, May 2002.
- [37] Y. Chang, C. Hu and M. Turk, “Manifold of Facial Expression”, *Proc. IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, Nice, France, Oct 2003.
- [38] K.-C. Lee, J. Ho, and M.-H. Yang, “Video-Based Face Recognition Using Probabilistic Appearance Manifolds”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 313-320, vol. 1, 2003.
- [39] Sam Roweis, Lawrence Saul and Geoff Hinton, “Global Coordination of Local Linear Models”, *Advances in Neural Information Processing System 14*, 2001.
- [40] A. Shashua, A. Levin and S. Avidan, “Manifold Pursuit: A New Approach to Appearance Based Recognition”, *International Conference on Pattern Recognition (ICPR)*, August 2002, Quebec, Canada.
- [41] J. B. Tenenbaum, Vin de Silva, and Johh C. Langford, “A Global Geometric Framework for Nonlinear Dimensionality Reduction”, *Science*, vol 290, 22 December 2000.
- [42] Sam T. Roweis, and Lawrence K. Saul, “Nonlinear Dimensionality Reduction by Locally Linear Embedding”, *Science*, vol 290, 22 December 2000.
- [43] Lawrence K. Saul and Sam T. Roweis, “Think Globally, Fit Locally: Unsupervised Learning

- of Low Dimensional Manifolds”, Journal of Machine Learning Research, v4, pp. 119-155, 2003.
- [44] M. Belkin and P. Niyogi, “Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering”, Advances in Neural Information Processing System 15, Vancouver, British Columbia, Canada, 2001.
- [45] Xiaofei He and Partha Niyogi, “Locality Preserving Projections”, Advances in Neural Information Processing Systems, Vancouver, Canada, 2003.
- [46] Fan R. K. Chung, Spectral Graph Theory, Regional Conferences Series in Mathematics, number 92, 1997.
- [47] C. Gonzalez and E. Woods, “Digital Image Processing Second Edition” Publishing House of Electronics Industry. 2003
- [48] Gunn S R. Support vector machine for classification and regression. Image Speech and Intelligent Systems Research group. University of Southapton: ISIS Technical Report. ISIS-98, 1998.
- [49] Osuna E, Freund R, Girosi F. Training support vector machines: an application to face detection. In Proc. Of Computer Vision and Pattern Recognition, Puerto Rico, 1997:130~136
- [50] Kapfer M, Bnois-Pineau J. Detection of human faces in color image sequences with arbitrary motions for very low bitrate videophone coding [J]. Pattern Recognition Letters, 1997,18(4):1503~1530
- [51] Schapire R E. The strength of weak learnability. Machine Learning [J]. 1990, 5 (2):197~227
- [52] Valiant L G. A theory of the learnable [C]. Communications of the ACM, 1984, 27(11):1134~1142
- [53] Rainer Lienhart, Jochen Maydt. An extended set of haar-like features for rapid object detection. The International Conference on Image Processing, 2002, Vol. 1:900-903
- [54] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting [C]. Second European Conf. on Computational Learning Theory, Barcelona, Spain, 1995
- [55] Jones M J, Viola P. Face recognition using boosted local feature [C]. International Conference on Computer Vision ICCV Beijing, China 2003
- [56] Samaria F S, Harter A C. Parameterization of a stochastic model for human face identification [A]. In: Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, Sarasoto, Florida, 1994. 245~248

- [57] Phillips P Jonathon, Moon H, Rizvi Syed A, et al. The FERET evaluation methodology for face-recognition algorithms [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1090~1104
- [58] 张晓华, 山世光, 曹波等. CAS-PEAL 大规模中国人脸图像数据库及其基本评测介绍[J] 计算机辅助设计与图形学学报. 2005, 17(1)
- [59] 刘瑞祯, 于仕琪 OpenCV 教程-基础篇. 北京航空航天大学出版社. 北京 2007.
- [60] 段锦 人脸自动机器识别. 科学出版社. 北京 2007

攻读硕士期间取得的科研成果

- [1] 蒋思洋, 解梅, 魏云龙. 使用扩展 LBP 特征进行人脸定位. 电子科技大学电子工程学院 2010 年学术年会, 2010
- [2] Jiang Si-Yang, Xie Mei, Wei Yun-Long. Face Location with an Extended Set of LBP. Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference, Vol. 1:325-328 (EI: 20113914378472)
- [3] 《一种基于局部三值模式的人脸特征提取方法》(专利号 201110136807.4)
- [4] 《基于生存指数熵的特征提取方法》(专利号 201110138932.9)