

第一章 绪论

1.1 研究背景及意义

自动人脸识别技术是指利用计算机分析人脸图像或视频，从中提取出有效的识别信息，用来辨别人脸身份的一种技术。它始于二十世纪六十年代中后期，迄今已有四十多年的历史，尤其是近十年来更成为模式分类和计算机视觉领域的研究热点。经过几十年的发展，自动人脸识别技术已经取得了长足的进展[7][11][55][101][110][115][153]。目前，最好的自动人脸识别系统在理想的测试条件下已经能够取得可以接受的识别性能。但是，大量的学术研究[23][42][109]和评测结果[95][96]表明：非理想测试条件下的人脸识别技术还远未成熟。要开发出真正鲁棒、实用的自动人脸识别系统，还需要解决大量的关键问题，尤其需要研究人脸识别的核心技术。

人脸识别技术作为一种最直接、自然和容易被人类所接受的生物特征识别技术，在理论研究和实际应用中都有着重要的意义。

1. 人脸识别技术研究的理论价值

人脸识别作为一个图像模式识别的经典问题，为图像分析与理解、计算机视觉、人机交互、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，用来尝试新方法、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，可以极大地促进这些学科的成熟和发展。它作为一个模式识别问题，被认为是最具有挑战性的问题之一：模式种类数目庞大；不同类别模式的差别非常微妙，这一点由于数据采集过程的噪声、成像设备的精度、外界条件的变化以及数据缺损等原因而显得更加突出。从图像处理的角度来看，如何对检测到的人脸图像进行处理以减弱光照、相机抖动等引起的变化，这是一个非常有意义的问题。再如，作为一个计算机视觉问题，如何融合人脸的一般性先验形状信息来准确地恢复特定人脸的三维结构也是一个非常有价值的研究问题。人脸识别技术的研究对于验证心理学和神经科学的一些假设和研究成果有着重要的意义[107]。例如，人类的面像识别机制是否是完全不同于其他一般物体的识别？人类大脑皮层中是否存在专用的人脸识别功能区？这是很多研究人员长期争论不休的焦点问题之一。此外，人脸识别也是智能人机接口领域的重要研究内容之一。人脸识别问题的最终解决可以极大地改善目前呆板、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. 人脸识别技术的应用前景

生物特征识别技术是依托于人类生理或行为特征使用自动化技术来进行身份识别或验证的方法，主要以人身体的一部分直接测量所得到的量度数据为基础。与传统的身

份验证方法相比,生物特征识别技术突出的优势表现在生物特征可以从根本上杜绝伪造和窃取,而且是人类自身拥有的,从而具有更高的可靠性、安全性和可用性。因此,许多生物特征识别技术受到了广泛关注,并在公共安全、金融安全、人机交互等方面有一些应用。目前,常用的生物特征识别技术包括人脸、指纹、掌纹、虹膜、静脉和DNA等识别技术。

与其他生物特征识别技术(如掌纹、虹膜、指纹、DNA)相比,人脸识别技术在可用性方面具有独到的技术优势。这主要体现在:

- 1) 非接触式采集,没有侵犯性,容易被接受;
- 2) 采集设备技术成熟,设备部署成本较低;
- 3) 符合人类的识别习惯,可交互性强;

表 1-1 人脸识别技术的典型应用

应用领域	特定的应用
智能卡	驾照
	移民、身份证、选民证、护照
	福利欺骗
信息安全	个人设备登录、桌面系统登录
	安全贸易终端
	应用安全、数据库安全、文件加密
视频监控	入口控制、事件分析
	嫌犯跟踪、调查
	高级视频监控
娱乐	视频游戏、虚拟现实、训练方案
	人与机器人交互、人与计算机交互

人脸识别技术具有广阔的应用前景。例如,在电子护照及身份证管理方面,许多国家正在计划或者实施用包含生物特征的证件对旅客进行识别和管理。其中,人脸识别技术作为首推的识别模式。再如,在公安、司法和刑侦方面,利用人脸识别技术可以进行走失人口及合法公民的身份认证、嫌疑犯查询等。在企业、住宅安全和管理方面,利用人脸识别技术的门禁考勤系统、人脸识别防盗门进一步提高了安全系数。在娱乐活动方面,人脸识别技术已应用于电子产品(如数码相机)、即时通信工具和虚拟化身的网络游戏等方面。表1-1中汇总了人脸识别技术在智能卡、信息安全、视频监控和娱乐等方面的一些典型应用[153]。

1.2 人脸识别问题描述

简要地讲,人脸识别问题可以描述为如下过程:给定一个场景的静态图像或者视频序列,利用预先存储的人脸图像数据,识别或者确认场景中一个或多个人的身份。对于计算机来讲,要完成自动的人脸识别任务,一般需要经过如下几个步骤,如图 1-1 所示:

1. 人脸检测与跟踪

人脸检测的主要任务是判断输入的图像或者视频序列中是否存在人脸, 如果存在, 则需要确定其所在的位置和大小。人脸跟踪可看作是人脸检测在视频序列上的扩展, 主要目标是将单帧的检测结果聚合为人脸随时间变化的轨迹。

2. 面部特征点定位

面部特征点定位的主要任务是在人脸检测与跟踪所给定的人脸区域(包括位置和大小)基础上, 确定面部显著特征(如眼睛、鼻子、嘴巴等主要器官)的位置甚至人脸轮廓的信息。

3. 特征表示

特征表示是在面部特征点定位的基础上, 从一个人的一幅或多幅人脸图像的面部区域中提取能够最好地描述该人或者最大限度地将此人与其他人区分开的特征。

4. 分类器

分类器的主要任务是根据从人脸图像上得到的特征表示来判定未知人脸的身份: 对于识别任务, 分类器根据注册图像集合来判定测试图像的身份; 对于确认任务, 分类器需要对测试图像声称的身份做出判断(拒绝或者接受)。

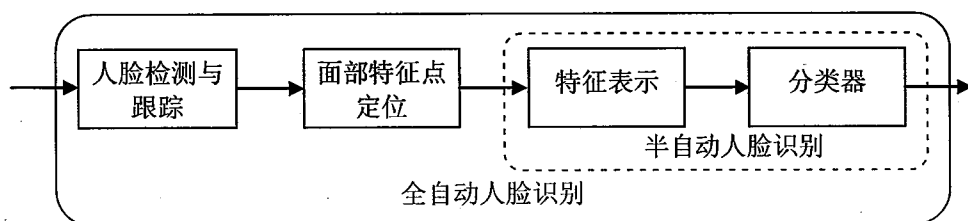


图 1-1 自动人脸识别系统的构成框图

一般来讲, 人脸识别算法的评测主要包括两种模式: 全自动人脸识别和半自动人脸识别。如图 1-1 所示, 全自动人脸识别是指识别算法自动地完成人脸检测与跟踪、面部特征点定位、特征表示和分类器这四个步骤, 不需要人工的辅助; 半自动人脸识别是指在给定面部特征点位置的基础上来完成特征表示和分类器这两个步骤。在实际的应用环境中, 不同识别算法的评测通常是在全自动模式下进行的。另一方面, 自动人脸识别系统的这几个组成部分, 即人脸检测与跟踪[66][137]、面部特征点定位[18][19]和半自动人脸识别[33][76][101][144], 也吸引了研究者们做出大量的研究。

对于半自动模式下的人脸识别算法来讲, 特征表示和分类器是至关重要的两个部分。在每人存在多个样本的情况下, 分类器的设计主要包括如下几类: 第一类方法的基本思路是将多类的人脸样本转换为两类样本, 即: 类内差异构成的正例样本与类间差异构成的反例样本, 然后在这两类样本上学习分类器, 如贝叶斯 (Bayesian) [86][87]分类器, AdaBoost[68][148]分类器和支持向量机 (Support Vector Machine, SVM) [92]; 第二类方法针对每类样本学习一个相应的分类器, 包括一对一 (one-against-one) 法和一对多 (one-against-rest) 法[40], 然后利用一定的结构完成分类, 如二叉树[34]; 第三类方

法通过计算样本到集合或集合到集合之间的距离进行分类,如最近特征线(Nearest Feature Line, NFL) [63]、最近特征子空间(Nearest Feature Space, NFS)[16]、最近流形(Nearest Manifold, NM)[147]和稀疏表示分类(Sparse Representation-based Classification, SRC) [124]等。

然而,人脸识别在应用环境中经常面临的一个问题是,每人只有单个样本。在这种情形下,最近邻分类器得到了广泛的应用。其分类过程如下:对测试图像与注册图像的相似度进行排序并将测试图像判断为最大相似度对应的人脸,或者,如果该最大相似度低于设定的阈值,则拒识该图像。此时,特征表示成为人脸识别算法的关键。本章第1.4节将从全局和局部这两个角度对人脸识别中常用的特征表示方法进行综述。

1.3 人脸识别研究概述

人脸识别作为模式识别和计算机视觉领域中的研究热点,每年都有大量的学术论文发表。以IEEE Xplore数据库为例,在1990年至2009年已发表的期刊和会议文章中,本文检索了“abstract”中包含了关键字“face recognition”的文章并按年份进行了统计,如图1-2所示。总体来看,与人脸识别相关的文章数量随年份的增加呈上升趋势。特别在2008年和2009年,每一年发表的文章数量都超过了1990~2000年这十年发表的文章数目总和。这主要有两个原因[153]:第一,人脸识别技术在商业、金融等领域内大范围的应用推动了它的发展;第二,经过几十年的研究,已经有了一些相对成熟的人脸识别技术。

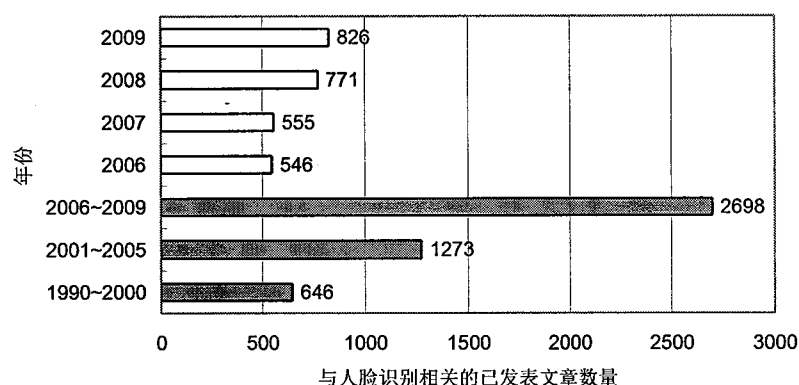


图1-2 根据IEEE Xplore数据库统计的近20年发表的与人脸识别相关的文章数量

本部分概述了人脸识别研究的相关综述及其评测,相关内容安排如下:第1.3.1节主要介绍了人脸识别在可见光图像、红外图像及三维图像等方面的相关综述工作;第1.3.2节从公共人脸识别库和性能评测这两个方面介绍了人脸识别的评测情况,并给出了最近几个评测的一些主要结论。

1.3.1 相关综述

针对人脸识别技术在可见光图像、红外图像及三维图像等方面的研究进展,前人已经发表了大量的综述文章:1995年,Chellappa等[11]从心理学、工程学等领域综述了过去20多年内的人脸识别研究,主要涉及人脸的分割/定位、特征提取和识别方法;2000年,周杰等[156]对此前三、四年内的自动人脸识别的相关工作进行了介绍,主要涉及人脸检测与定位、特征提取与识别这两个方面;2003年,Zhao等[153]对静态图像和视频序列上已有的人脸识别方法进行了分类,并详细介绍了每一类的代表性方法,而且讨论了涉及心理学研究、系统评价以及光照和姿态变化等难点问题;2004年,山[101]以时间为线索将此前人脸识别的研究历史划分为3个阶段,并对每个阶段的特点及代表性方法进行了介绍;2005年,Kong等[55]综述了可见光图像和红外图像方面的最新人脸识别技术,认为融合不同成像模式和光谱分量可以提高人脸识别系统的整体性能;2006年,Bowyer等[7]对三维人脸识别中的方法及其与二维人脸识别方法的结合进行了综述,指出了在传感器、识别算法以及评测等方面面临的一些挑战;2009年,苏[110]在文献[101]的基础上进一步将人脸识别的研究历史划分为四个阶段,并重点对后两个阶段(1998年至今)的情况进行了介绍。

围绕影响人脸识别的一些因素(如光照、姿态等),研究人员开展了大量的工作[17][31][65][160]。这些研究的目的是最大限度地减少这些因素对人脸表现的影响,为识别算法提供尽可能“纯净”的图像,以提高识别算法的性能。此外,头部姿态估计[17]、表情分析[25]、人脸图像的老化合成与年龄估计[26]等已经成为相对独立的研究课题,也受到人们越来越多的关注。

1.3.2 人脸识别评测

客观的评价方法对于促进人脸识别技术的发展与进步有着重要的作用,这主要涉及两个方面:测试数据库与性能评测。具体来讲,测试数据库可以看作某种应用条件下数据的“代表”,对于识别算法的研究与开发、性能的测试与比较是必不可少的;性能评测是在一个共同的测试协议下对不同算法的性能进行比较,在发现问题、指明改进方向等方面起着重要的作用。下面首先介绍了国内外的主要公共人脸识别库,然后简述了近几年人脸识别算法的性能评测情况。

1. 主要公共人脸识别库

在人脸识别技术发展的早期阶段,许多小规模的人脸数据库得到了广泛的应用,如ORL人脸库[100]、Yale人脸数据库(<http://cvc.yale.edu/projects/yalefaces/yalefaces.html>)。然而,随着人脸识别技术的发展,识别算法性能的不断提高,这些小规模人脸数据库已经不能够满足识别算法评测的要求。近几年来,一些大规模、涵盖了多种变化的人脸数据库也相应地被构建并发布出来,如FRGC 2.0[94], Multi-PIE[32]等人脸库。下面将简要介绍国内外的一些公共人脸识别数据库,特别是一些较大规模的人脸数据库。

● AR人脸数据库[81]

由西班牙巴塞罗那计算机视觉中心于1998年创建的, 包括126个人不同光照、表情、遮挡和老化的人脸图像, 超过4,000幅。图像分为两个不同的时间段采集, 中间间隔了2个星期。

● CMU PIE人脸数据库[106]

由美国卡内基梅隆大学于2000年创建, 包含了68个对象的40,000多幅图像。该数据库涵盖了每个对象的13种姿态、43种光照条件和4种表情变化。

● YaleB人脸数据库[31]

由耶鲁大学计算视觉与控制中心创建, 包含了10个人576种视角下的5,760幅单光源图像和每个对象特定姿态、背景光照下的90幅图像, 共计5,850 (5,760+90) 幅图像。其中, 包含姿态和光照变化的图像是在严格控制条件下采集的, 主要用于光照和姿态问题的建模与分析。详情参见: <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>。

● 扩展YaleB人脸数据库[31]

该数据库包含了28个对象在9个姿态和64种光照条件下的16,128幅图像, 所有图像根据手工标定的特征点归一化到大小为168×192个像素。详情参见:

<http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>。

● XM2VTS数据库[84]

该数据库包括295个人在4个月时间内4次录制的人脸和语音数据, 每次采集都包括2个头部旋转视频片段和6种不同语音视频片段。详情参见:

<http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/>。

● FERET人脸数据库[93]

由美国国防部资助的FERET (FacE REcognition technology Test) 项目 (1993年~1997年) 创建, 总共包含14,051幅多姿态的灰度人脸图像。在1998年到2001年之间发布的数据包含了3,737幅图像, 也被称为训练数据。该数据库严格划分了Gallery和测试集, 是人脸识别中应用最广泛的人脸数据库之一。详情参见:

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

● FRGC 2.0人脸数据库[94]

由美国国家标准局 (National Institute of Standards and Technology, NIST) 创建, 包含超过50,000张高分辨率的2D和3D人脸图像。图像在可控和非可控两种环境下采集, 并包含一定的表情变化。此数据库的测试协议定义了6个不同的实验, 分别用来测试不同条件下2D和3D人脸识别算法的性能。详情参见: <http://www.frvt.org/FRGC/>。

● CAS-PEAL人脸数据库[29]

由中科院计算所JDL实验室于2004年创建, 包含了1,040名中国人 (男:595, 女:445), 共99,594幅图像。所有图像在专门的采集环境中采集, 涵盖了姿态、表情、饰物和光照四种主要变化条件, 部分人脸图像具有背景、距离和时间跨度的变化。该数据库定义了

标准的训练集和测试集。目前发布的版本包含了1,040个对象的30,863幅图像。详情参见：<http://www.jdl.ac.cn/peal/index.html>。

- **LFW人脸数据库[42]**

由马萨诸塞大学阿姆赫斯特分校的计算机视觉实验室创建，包含5,749人的13,233张图像，其中的1,680人有两张或两张以上的图像。与绝大多数人脸数据库不同，此数据库中的图像全部来自网络，呈现了在姿态、表情、光照、图像质量等很多方面的变化。详情参见：<http://vis-www.cs.umass.edu/lfw/>。

- **SCface人脸数据库**

由克罗地亚萨格勒布大学电子信息计算系创建，主要包括人脸静态图像集合。图像通过5种不同质量的视频监控摄像头在非可控的室内环境下捕获。该数据库共包括130个对象的4,160幅静态图像（可见光图像与红外图像）。详情参见：<http://www.scface.org/>。

- **Multi-PIE人脸数据库[32]**

该数据库可看作是PIE数据库的扩展，包含了337个对象在15种视角、19种光照条件下在四个时间段内采集得到的图像。详情参见：<http://www.multipie.org/>。

2. 性能评测

近年来，主要的国际级人脸识别算法评测有：FERET (Face Recognition Technology) 测试[93]、FRGC (Face Recognition Grand Challenge) 测试[94]、FRVT(Face Recognition Vendor Test) [95]和 MBGC (Multiple Biometric Grand Challenge) 测试[96]。这里将重点介绍 FRVT 2006 和 MBGC 的评测结果。

- **FRVT 2006**

FRVT 2006 是由美国国家标准局 (NIST) 组织的，其主要目标是衡量 FRVT 2002 以后的原型系统/算法以及商业人脸识别系统的进展。其测试的重点主要集中在高分辨率的静态图像（5 百万~6 百万个像素）、三维人脸图像、每人多个样本以及姿态和光照补偿预处理算法。

FRVT 2006 的主要测试结果包括[95]:

- 相比 FERET 项目初期的结果，FRVT 2006 的识别结果提高了接近两个数量级。在错误接受率 (False Accept Rate, FAR) 等于 0.001 时，FERET 1996, FRVT 2002 和 FRVT 2006 最好方法的错误拒绝率 (False Reject Rate, FRR) 分别是 0.54, 0.20 和 0.01。性能改进的主要原因在于算法设计、传感器以及光照预处理算法等方面的进步。

- 在可控光照条件下，FRVT 2006 在超高分辨率图像上达到了 FRGC 设定的目标，即：在 FAR 等于 0.001 时 FRR 达到 0.02。在高分辨率以及低分辨率的图像集合上，最好方法的 FRR 值都达到或接近 0.02。

- 在三维人脸图像上，在 FAR=0.001 时最好方法的 FRR 值范围在 0.005~0.015，已经达到了 FRGC 的目标。

- 在光照非可控条件下，FRVT 2006 的结果已经比 FRGC 的结果有了显著的提高：

在 2005 年, FRGC 上最好的三种结果 (FRR 当 FAR=0.001) 分别是 0.24, 0.39, 0.56; 在 FRVT 2006 中, 在超高分辨率 and 高分辨率图像上的最好结果 (FRR 当 FAR=0.001) 已分别达到 0.103~0.130, 0.119~0.146。

➤ 在人和计算机识别能力的对比测试中, 性能最好的人脸识别算法已经达到了人类识别性能的级别, 在 FAR 值等于 0.05 时的准确度已经优于人类。此外, 人类对不熟悉人脸的识别能力要远远低于熟悉人脸的识别能力。

● MBGC 测试

MBGC 测试同样是由 NIST 组织, 但是不同于 FRVT, MBGC 测试包含了人脸和虹膜两种数据, 因此也为研究多种生物特征的融合提供了可能。MBGC 的主要目标是通过一系列挑战性的问题来研究、测试和提高人脸与虹膜识别技术在静态图像和视频上的性能。为了促进当前技术的发展以及根据独立的评测得出相应的结论, MBGC 包含了一系列挑战性的问题[96]:

➤ 入口视频的虹膜与人脸识别: 其目标是从人体穿过入口时拍摄的近红外和高分辨率视频序列上开发出身份识别算法。近红外视频上进行虹膜识别, 高分辨率视频上则进行人脸识别。

➤ 静态图像上的人脸识别: 第一个目标是在非可控的室内/室外光照下采集的正面和侧面图像上提高识别算法的性能, 第二个目标是在调整过大小和压缩过的正面人脸图像上提高识别算法的性能。

➤ 视频上的人脸识别: 其目标是从在姿态、光照以及照相机角度等未控制条件下采集的视频序列上识别出人的身份。

在静态图像的人脸识别测试中, MBGC 发布的数据中包括了在光照可控与非可控条件下采集的高分辨率静态图像。因此, 该实验需要在“可控”与“非可控”、“非可控”与“非可控”条件下的图像集合之间进行比对。为了测试图像大小是否满足国际民间航空组织 (International Civil Aviation Organization, ICAO) 电子护照的标准, 该测试还发布了另外两个集合: 在第一个集合中, 图像中两眼间距被缩放到 90 个像素, 并压缩到 8KB; 在第二个集合中, 图像中两眼间距被缩放到 120 个像素, 并压缩至 20KB。这两个集合中的图像均采用 JPEG 2000 的压缩标准。此外, MBGC 还发布了 770 段分别包含正面和侧面人脸的视频片段, 用来评测视频数据上的人脸识别算法。

目前, MBGC 已经完成了前两个阶段的测试。图 1-3 示出了 MBGC 前两个阶段的测试结果。从 2008 年和 2009 年的测试结果来看, 算法性能已经有了比较大的提升。如图中所示, 2009 年的最好的测试结果 (错误接受率=0.1%时的确认率) 能够达到约 98%。此外, 从 2009 年的结果来看, 非可控的测试条件与图像压缩的测试条件对这几种算法的影响不明显。MBGC 的测试报告认为, 跨姿态的人脸识别是一个非常有挑战性的问题。

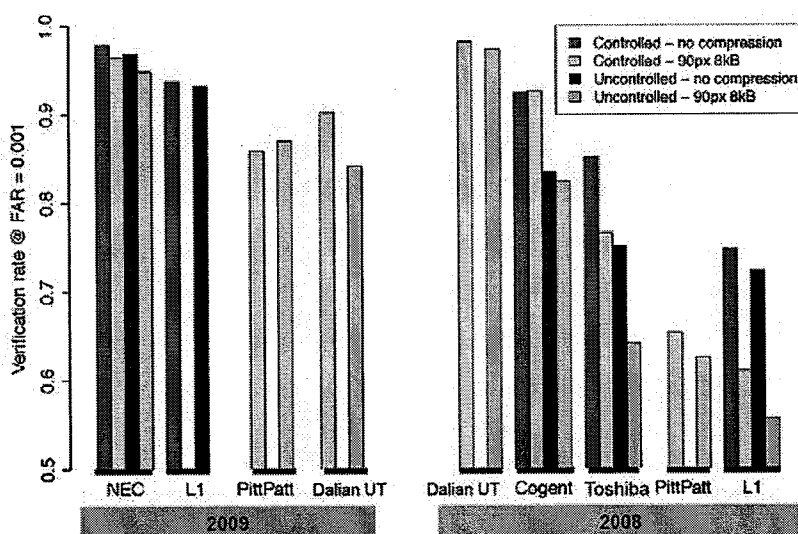


图 1-3* MBGC 测试前两个阶段的测试结果

1.4 特征表示方法综述

总体来讲, 人脸识别中常用的特征表示方法包括两大类: 全局表示与局部表示。全局表示是指其特征向量是从整幅人脸图像(即全部像素)上提取得到, 通常反映了人脸的整体属性; 不同于全局表示, 局部表示的特征向量是从人脸图像的局部区域(即部分像素)中提取得到。当然, 这两种表示之间并不存在严格的界限: 当局部表示中用于提取特征的图像区域变大时, 局部表示的“全局性”逐渐变强; 特别地, 当该区域增大至整幅图像时, 局部表示就变为全局表示。下面主要从“全局”和“局部”两个方面对人脸识别中的表示方法进行介绍。

1.4.1 全局表示方法

概括地讲, 人脸识别中的全局表示方法主要包括子空间表示、流形表示、稀疏表示与全局频域表示四大类。下面将简要介绍每一类表示的代表性方法及其特点。

● 子空间表示

子空间表示[3][76][131][140][135]是人脸识别中研究最为广泛的一类表示方法, 其基本思想是将整幅人脸图像看作高维空间中的样本点, 认为人脸样本在整个图像空间中处于一个维数较低的子空间内, 通过优化求解定义的目标函数得到子空间的基向量, 进而来表示人脸样本。由于子空间的维数常常低于人脸样本的原始维数, 这类方法也被称为特征降维方法。

基于主成分分析(Principal Component Analysis, PCA)的表示是子空间表示的一种代表性方法。主成分分析的主要目的是寻找在均方误差意义下, 最能够代表原始数据的

* 摘自 http://face.nist.gov/mbgc/mbgc_presentations.htm 中静态图像上的测试结果

投影方法。它通过对原始数据构成的协方差矩阵进行特征值分解得到特征向量, 其中, 前 n 个主成分分量依次对应了特征值最大的 n 个特征向量[54][119]。在经典的 Eigenfaces 方法[119]中, 人脸图像表示为该图像在这 n 个主成分分量上的投影系数。由于人脸图像通常具有很高的特征维数, 在训练样本数目比较少的情况下, 这将导致计算得到的协方差矩阵不能准确地反映样本的分布, 从而使求得的子空间不准确。从这一角度出发, 很多研究者提出了相应的改进算法, 如二维主成分分析(Two Dimensional PCA, 2D PCA)[132]、统一主成分分析(Unified Principal Component Analysis, UPCA)[103]。主成分分析对于在低维空间中表示原始数据非常有效, 但是在计算样本的协方差矩阵时并没有区分不同类别样本之间的差异。因此, 在应用于分类任务上, 基于主成分分析的表示并不一定能够很好地区分各类样本。

子空间表示的另一类代表性方法是基于线性判别分析(Fisher's Linear Discriminant Analysis, FLD 或者 LDA)的表示。不同于主成分分析, 线性判别分析的主要目的是寻找能够最大程度区分各类样本的投影方向。为了衡量不同类别样本之间的可分性, 它定义了样本的类内散度矩阵与类间散度矩阵, 然后寻找最大化类间散度并且最小化类内散度的投影方向。通常情况下, LDA 投影向量的求解转化为一个特征值分解过程, 并选择特征值最大的前若干个特征向量作为样本的投影方向。目前, 基于 LDA 的 Fisherfaces[4]方法已经成为人脸识别中的经典算法。围绕 LDA 存在的一些限制或假设, 大量的研究工作也相应地展开, 如针对小样本问题的“PCA+LDA”方法[4][71][82][111][122][154]、正则化方法[20][48]、基于二维图像矩阵的 2D LDA 方法[133][139]以及零空间 LDA 方法[15][123], 降低 LDA 对数据分布要求的边界 Fisher 分析(Marginal Fisher Analysis, MFA)[131]和子类判别分析(Subclass Discriminant Analysis, SDA)[158], 同时利用类内散度矩阵的主子空间与零空间的双子空间 LDA(Dual-Space LDA)方法[123], 令特征距离度量和分类器距离度量保持一致的基于相关相似度(Correlation Similarity)度量的判别分析方法[80], 避免 LDA 投影方向易于合并距离较近类别的子空间选择方法[114]。这些方法改进了原始 LDA 方法中存在的某些限制或不足, 因而在相应的测试条件下取得了比原始 LDA 方法更优的性能。

除主成分分析与线性判别分析之外, 其他的一些子空间表示也受到了关注, 如贝叶斯双子空间表示[86][87]和独立元分析表示[3][140]。此外, 一些研究者也将核技巧(Kernel)引入进来, 提出了子空间表示方法的非线性扩展[52][73]。

● 流形表示

流形表示是借助于流形方法来寻找隐藏在高维样本(如手写体、人脸图像)空间内有意义的低维表示。比较有代表性的流形方法包括局部线性嵌入(Locally Linear Embedding, LLE)[99], ISOMAP(Isometric feature mapping, ISOMAP)[118]和 Laplacian Eigenmaps[5]算法。最近, Wang 等[121]提出了流形-流形距离(Manifold-Manifold Distance, MMD)的理论框架, 并将其应用于基于图像集合的人脸识别问题。Wang 等通过大量的

实验验证表明, 提出的流形-流形距离作为一种通用的图像集合相似性度量, 一致地优于该文献中其他的非判别式算法[120]。然而, 这些方法通常要求每一类的样本数目比较多, 这在实际的应用条件下很难满足; 此外, 它们只能对训练数据完成降维, 无法扩展到训练集之外的测试样本。

为了完成对训练集之外样本的特征降维, 一些研究者提出了利用流形思想的线性降维算法。局部保持投影 (Locality Preserving Projections, LPP) [38] 是其中非常有代表性的一种方法, 其基本思想是寻找一个嵌入来保持人脸样本之间的局部结构信息, 从而得到一个最好地保持人脸流形结构的子空间。He 等[38] 在 Yale 和 PIE 人脸数据库上验证了 LPP 方法提供了更好的表示, 能够取得比 Eigenfaces[91] 和 Fisherfaces[4] 方法更低的错误率。尽管如此, Yang 等[134] 认为 LPP 的优化目标与分类目标无直接关系, 因此提出了无监督判别投影 (Unsupervised Discriminant Projection, UDP) 方法。UDP 方法可看作基于多流形学习框架的一种近似, 其目标是寻找一种最大化非局部散度同时最小化局部散度的投影。此外, 研究者们对其他一些扩展方法也进行了大量的研究, 如邻域保持嵌入 (Neighborhood Preserving Embedding, NPE) [39], 局部判别嵌入 (Local Discriminant Embedding, LDE) [14]。这些方法在保持了数据的某些局部或者全局几何属性的条件下, 可以在一定程度上发现隐含在样本集当中的流形结构, 并能够表示训练集之外的样本。尽管如此, 这些方法大多是从保持数据流形结构的角度出发来寻找特征表示, 因而这些特征未必适合于分类任务。

● 稀疏表示 (Sparse Representation)

稀疏表示来源于近几年在信号处理领域里一项新兴的技术, 压缩传感 (Compressive Sensing)。奈奎斯特采样定理认为, 当采样频率大于信号中最高频率的2倍时, 采样后的信号才能够不失真地重构原始信号。与之相对, 压缩传感理论[10]认为, 只利用信号的少量采样就能够重构某些原始信号。稀疏表示要解决的问题是求解信号在某个过完备词典 (Overcomplete Dictionary) 中元素的最紧致线性表示, 其目标是用尽可能少的元素来降低信号的重构残差[43][124]。它的目标函数主要包括两项: 一项度量了信号的重构残差, 另一项度量了稀疏性。通常引入一个参数 λ ($\lambda > 0$) 在二者之间进行均衡。

Huang 和 Aviyente[43] 将判别方法的判别性和稀疏表示的重构性、稀疏性结合在一起, 根据稀疏表示提出了用于信号分类的一种理论框架, 能够处理含有噪声的信号、信号的缺失等问题。Wright 等[124] 提出了基于稀疏表示的分类 (Sparse Representation-based Classification, SRC) 方法, 其基本思想如下: 将测试样本表示为训练样本的稀疏线性组合, 然后根据稀疏系数计算测试样本在每类训练样本上的重构残差, 最后将测试样本判别为重构残差最小的训练样本所属的类别。Ji 等[47] 对稀疏表示的分类方法作出了如下两点改进, 取得了比 SRC 方法更好的结果: 第一, 在稀疏表示问题中引入了系数非负的约束; 第二, 利用马氏距离 (Mahalanobis distance) 来度量原始样本和重构样本之间的距离。为了进一步提高稀疏表示对遮挡的鲁棒性, Zhou 等[157] 利用马尔科夫随机场将先

验知识引入到稀疏表示的框架中,并验证了该方法能够比SRC方法[123]处理更大比例、更多类型的遮挡。最近,Wright等[125]介绍了稀疏表示在计算机视觉领域中的几种典型应用,认为稀疏表示有效性的关键是如何选择词典,即:在该词典上的稀疏表示能够正确地反映样本的语义特征。

文献[124]中的SRC方法是稀疏表示在人脸识别上的一种典型应用,其主要不足之处在于:第一,该方法通常需要每个人注册有多幅人脸图像而且这些图像的分布需要满足线性子空间假设,这样的假设在很多实际的情况下是无法满足的;第二,计算复杂度比较高,无法满足人脸识别实时应用的需求。

● 全局频域表示

图像的频域特征是指将图像变换到频域后,利用频域里的系数来表示图像。这里主要介绍两种全局频域特征:离散余弦变换(Discrete Cosine Transform, DCT)和离散傅里叶变换(Discrete Fourier Transform, DFT)。这两种频域特征的不同频段反映了图像的不同属性:低频部分对应图像的整体描述,高频部分则对应了图像的细节描述。由于高频部分对于区分不同人的作用不大,因此,大多数的识别算法只利用了低频信息。

Hafed和Levine[37]将DCT作为特征提取工具,只保留了人脸图像DCT系数的一小部分作为特征向量,可以取得与Eigenfaces算法接近的识别性能。Er等[24]根据DCT、Fisher线性判别分析和径向基函数(Radial Basis Function, RBF)神经网络设计了一个快速的人脸识别系统。作者还验证了去除图像DCT系数的前几维,可以减小大范围光照变化的影响。Chen等[12]基于DCT提出了一种光照预处理方法,其基本思想是在图像对数域里应用DCT并去除一部分DCT系数以减少光照变化的影响。Liu和Liu[77]则利用DCT从人脸图像的不同彩色分量上提取低频特征并保留不同大小的模板,然后利用增强的Fisher线性判别模型(Enhanced Fisher linear discriminant Model, EFM)进行特征提取,在大规模的FRGC 2.0人脸库上取得了良好的性能。为了保证实用系统的性能及推广性,Hwang等[44]从多个人脸模板的不同频段中提取混合的傅里叶特征(hybrid Fourier feature)来表示人脸。该方法首先得到由不同的眼睛位置、相同大小的图像构成的多个模板,然后对每个模板提取了不同类型的傅里叶特征,并利用“PCA+LDA”方法分别从频域特征和多个模板图像中提取判别性特征。为了提取人脸图像的全局特征,Su等[109]利用LDA从图像傅里叶变换的低频特征里提取判别性特征,并与局部特征进行融合,在FRGC 2.0人脸库上取得了非常优异的性能。

1.4.2 局部表示方法

粗略地讲,局部表示方法主要包括如下几大类:几何结构特征、局部统计分析表示、局部频域表示以及局部纹理表示。下面将简要介绍这几类表示的代表性方法。

● 几何结构特征(geometric feature)

在早期的人脸识别方法中,几何结构特征[8][49]吸引了大量研究者的目光。其基本思想是将人脸表示为面部主要特征(例如,眼睛、鼻子和嘴巴等)以及它们之间的相对

关系, 利用这些面部特征和面部轮廓的相对位置和大小作为特征参数来形成特征向量, 然后采用适当的距离度量和分类方法进行分类。为了从人脸图像中准确地提取几何结构特征, Yuille提出了可变形模板[141] (Deformable Templates) 方法。该方法根据参数化的几何模型来指定人脸主要特征(如眼睛、嘴巴)的模板, 通过优化的方法来调整参数的数值以达到最优的匹配。这组参数的最终值用来表示面部特征, 可用于人脸的描述与识别。

几何结构特征的优点是所需的存储空间和比对时间比较小, 缺点主要表现在两方面: 第一, 几何结构特征提取的准确度受外界条件的影响非常大, 在复杂的光照条件、姿态变化、表情变化、饰物变化等情况下甚至变得非常困难; 第二, 人脸结构的相似性使得几何结构特征本身对不同人的区分度不高, 这导致了该类方法的识别精度比较低。

● 局部统计分析表示

当人脸图像存在局部变化(如表情和部分遮挡)时, 基于全局统计分析的表示方法(如子空间表示)容易受到较大的影响。为了提取对这些变化鲁棒的特征, 一些研究者提出了基于局部统计分析的表示方法。

一种典型的局部统计分析表示方法是从全局人脸图像的集合中学习出局部的基图像, 然后利用这些基图像来提取人脸图像的局部特征。局部特征分析(Local Feature Analysis, LFA) [90]在本质上是一种基于统计的低维对象描述方法, 其提取的特征是局部的并保留了全局拓扑信息, 具有比PCA表示更好的描述能力。与学习全局表示的PCA和向量量化(Vector Quantization, VQ)方法不同, 非负矩阵分解(Non-negative Matrix Factorization, NMF) [59]通过非负约束来学习物体基于部件的表示。在NMF的基础上, 局部非负矩阵分解(Local Non-negative Matrix Factorization, LNMf) [64]增加了局部性约束来学习视觉模式的空间局部表示。为了获得对应人脸显著区域(如眼睛、眉毛、鼻子和嘴巴)的基图像, 局部显著ICA (Locally Salient ICA, LS-ICA) [51]在计算独立元分析(即ICA)基图像的过程中增加了加性的局部约束。实验结果表明, LS-ICA表示要优于PCA、ICA, LFA和LNMf等表示, 这一点在人脸图像存在部分遮挡和局部形变的情况下尤其明显。

另一种局部统计分析表示方法则在人脸图像的部件或块级别上进行统计分析以获得局部特征表示。文献[91]提出了基于部件的PCA表示(如EigenEyes), 克服了Eigenfaces方法易受图像整体变化的影响, 在实验中取得了更好的识别性能。文献[40]和[41]自动检测出人脸的不同部件, 将这些部件归一化后并将它们的灰度特征组合为一个特征向量, 然后输入到后端的SVM分类器进行分类。为了增强LDA特征的推广能力, 文献[53]将人脸图像划分为一系列区域并在每一个区域内利用LDA提取特征。最近, Kumar等[56]在人脸部件上利用SVM构建simile分类器并将其应用于人脸确认问题, 在极富挑战性的LFW人脸库上取得了优异的性能。该类方法可看作统计分析方法在人脸局部区域上的扩展: 这种扩展一方面降低了原始的特征维数, 一定程度上避免了“维数灾难”问题;

另一方面,由光照、姿态等因素在部件(或块)上产生的变化要小于人脸整体上的变化,因此,这种扩展取得了良好的分类性能。

● 局部频域表示

小波特征作为一种局部的空频变换特征,反映了图像中多尺度、多方向的局部特征,在图像处理领域得到了广泛的应用。根据小波包分解, Garcia等[30]利用包含小波系数的带通滤波图像来表示人脸图像,然后采用一些简单的统计度量来构建人脸图像的紧致特征表示,最后用基于Bhattacharyya距离的概率度量进行分类。Lai等[58]将小波表示与傅里叶变换结合起来,提出了“spectroface”方法。该方法先利用小波变换得到图像的低频段特征,然后利用傅里叶变换提取频域特征并转换到极坐标系下,在对极径应用对数变换后再进行第二次的傅里叶变换,将输出系数作为图像的特征表示。Chien和Wu[16]通过多级的二维离散小波变换(Discrete Wavelet Transform, DWT)获得人脸图像的低维特征向量,并利用LDA方法提取更具判别性的特征表示,取得了比Eigenfaces方法更优的分类性能。

Gabor小波作为一种特殊的小波,其形状类似于哺乳动物视皮层简单细胞的二维感受野,能够捕捉对局部变化(如表情、光照)鲁棒的特征,在模式识别领域得到了广泛应用[21][22][57][104]。图像的Gabor小波表示定义为该图像与一族Gabor小波函数卷积的结果。对于图像中的某个位置来讲,它的Gabor特征包括不同尺度、不同方向的Gabor小波响应,这些响应构成了该位置上的Gabor Jet。在用于人脸识别时,一幅图像可表示为某些预定义特征点的Gabor Jet,也可表示为该图像与不同Gabor小波的卷积结果。

弹性图匹配方法是基于Gabor小波表示的一类最有代表性的方法,如动态连接匹配(Dynamic Link Matching, DLM)[57]和弹性束图匹配(Elastic Bunch Graph Matching, EBGM)[126]。其基本思想是将人脸图像表示为一个属性图:图的顶点代表面部关键特征点,其特征由Gabor Jet表示;图的边反映了不同特征点之间的几何关系,描述了人脸的几何结构信息。对于任意输入的人脸图像,弹性图匹配通过优化搜索策略确定面部关键特征点,并构建与之对应的属性图,最后通过计算其与已知人脸属性图的相似度来完成识别过程。弹性图匹配方法同时保留了人脸的全局几何结构又建模了关键特征点的局部表示,具有比较好的鲁棒性。其不足在于计算复杂度比较高,识别速度比较慢。

由于采用了多个尺度、多个方向的Gabor小波,图像Gabor小波表示的特征维数是非常高的。例如,对于一幅 64×64 的人脸图像,如果采用5个尺度、8个方向Gabor小波进行特征提取,图像的Gabor特征维数高达163,840 ($64 \times 64 \times 5 \times 8$)维。高维的特征不仅带来了巨大的计算代价,而且还会带来“维数灾难”(Curse of Dimensionality)问题。为了得到低维的特征表示, Liu和Wechsler[71]提出了GFC(Gabor Fisher's Classifier, GFC)方法:该方法首先对Gabor特征进行均匀下采样得到低维特征向量,然后利用PCA进行降维,最后通过增强的Fisher线性判别分析得到用于人脸识别的特征。该方法虽然利用均匀下采样来降低特征维数,但在PCA降维之前的特征维数依然比较高。此外,过大的

采样间隔也会丢掉样本很多的表示信息。为了从冗余的高维Gabor特征中得到对分类有利的表示,一些研究者将特征选择的方法应用于Gabor特征,如根据不同区域的重要性来设置不同的采样间隔[75]以及利用Adaboost的学习策略来选择特征的AdaGabor方法[137]。为了从Gabor特征上提取有利于分类的表示, Su等[108]把人脸图像的Gabor特征按照空间位置分组,然后训练多个LDA分类器,最后将分类器进行融合。根据主动视觉中物体识别的前馈模型[98], Meyers和Wolf基于Gabor小波表示来提取S2层的人脸特征,并在FERET、ORL和AR人脸库上验证了该特征的有效性[85]。此外,一些研究者将广义判别分析(Generalized Discriminant Analysis, GDA)[105][113]、独立元分析[72]、核PCA[73]以及多类核的Fisher判别分析[74]等方法应用于Gabor特征以提取更紧致的表示。最近, Yang和Zhang[136]提出了利用Gabor特征的稀疏表示分类方法,同原始SRC方法相比,不仅降低了稀疏表示的计算代价,而且取得了更优的性能。

● 局部纹理表示

局部纹理表示方法主要是对像素所在的局部区域进行建模分析以提取反映其纹理特征的表示。由于利用了局部区域内的信息来表示单个像素的特征,该类方法通常比亮度特征具有更好的表示能力。

局部二值模式(Local Binary Patterns, LBP)表示是该类方法中比较有代表性的一种表示,在纹理分类[88]、人脸检测[36]与识别[1]等领域得到了广泛应用。通过计算图像像素局部邻域内中心像素与邻域像素的亮度差异并二值化, LBP算子将图像上的每个像素赋值为一个二进制串,并将该二进制串转化为十进制数来表示当前像素,称为“模式”(Pattern)。Ojala等[88]对LBP算子提出了两种扩展:处理不同尺度纹理特征的多尺度LBP算子和占图像模式绝大多数的Uniform模式。此外,针对LBP方法的改进工作也得到了广泛的研究,如基于体的LBP[152]、局部三值模式(Local Ternary Patterns, LTP)[116]、基于块的LBP[68][130]及LBP表示结合统计分析的方法[65][102][148]。

受LBP表示的启发,一些研究者从其他的角度来提取图像像素局部邻域内的纹理特征。为了获得人脸专属的模式表示, Meng等[83]提出了局部视觉基元(Local Visual Primitives, LVP)的表示方法:该方法对人脸图像上的局部块进行聚类学习以获得表示人脸的码本,可用于人脸图像的重构与识别。类似地, Cao等[9]提出了基于学习的描述子并将其应用于人脸识别,在LFW人脸库上取得了很好的效果。根据人类的感知与信号变化以及原始信号强度之间存在的关系, Chen等[13]提出了韦伯局部描述子(Weber Local Descriptor, WLD),在纹理分类与人脸检测上表现出了良好的性能。为了利用人脸图像上相邻区域的序关系, Liao等[67]将序特征引入进来并将其应用于人脸识别。为了建模局部邻域内像素间的二阶或高阶关系,利用梯度信息的模式表示[69][143]也被提了出来。这些方法由于利用了更多的信息来提取局部纹理特征,取得了比LBP表示更优的结果[9][13][69][143]。其不足之处在于引入了更多的参数,实现过程比LBP方法复杂。

由于Gabor小波从人脸图像中提取了多尺度、多方向的局部特征,并对噪声有很好

的抑制作用,一些研究者将基于亮度特征的方法扩展到Gabor特征上。Zhang等[150][151]利用LBP算子分别来计算Gabor幅值和相位特征的局部模式,提出了局部Gabor二值模式(Local Gabor Binary Patterns, LGBP)表示方法。受相位特征在掌纹和虹膜中成功应用的启发,Zhang等[142]提出了Gabor相位模式直方图(Histogram of Gabor Phase Patterns, HGPP)表示,并验证了该方法可以取得比LGBP表示更优的性能。与这些方法不同,Zhong等[155]从模式学习的角度,利用聚类算法从Gabor特征上学习得到人脸图像的局部模式。虽然区域协方差矩阵(Region Covariance Matrices, RCM)[120]在人体检测与物体跟踪领域得到了应用,然而,传统RCM表示在人脸识别上的性能并不理想。Pang等[89]利用图像的Gabor Jet特征来提取人脸不同区域内的RCM表示,然后利用基于特征值的距离度量来计算黎曼流形上协方差矩阵的距离。这些表示方法利用Gabor小波作为一种特征增强的手段,具有比较强的表示能力,因此分类性能要优于基于亮度特征的LBP方法。其主要不足在于,多尺度、多方向的Gabor特征导致了比较大的计算代价。

梯度特征作为图像处理的一种基本工具,在人脸识别中也受到了一些关注。边缘线条图(Line Edge Map, LEM)表示[28]在对人脸的边缘图像细化之后应用多边形线拟合得到线段集合来表示人脸图像。根据梯度特征具有对光照变化鲁棒性的这一特点,文献[149]从图像中提取梯度方向特征来描述人脸图像,并分析了该表示在不同光照条件下具有很好的鲁棒性。在物体分类领域,尺度不变的特征变换(Scale Invariant Feature Transform, SIFT)[78]方法受到了广泛的关注。文献[6]从人脸图像中提取SIFT特征,通过计算SIFT特征集合之间的距离来得到图像间的相似度。文献[79]认为,SIFT特征检测到图像上稳定并且显著的局部特征点,可以提取人脸专属(person-specific)的特征表示。此外,也有研究者[2]将梯度方向直方图(Histogram of Oriented Gradient, HOG)特征应用于弹性图匹配方法,并在一些公共人脸库(如FERET)上取得了一定效果。SIFT特征作为梯度幅值和方向信息的一种结合,对刻画人脸图像的一些特异性特征有一定作用,但它在人脸识别上的应用还需要进一步的研究。

直觉上,人脸皮肤上的斑/痣点、胎记等显著性纹理特征对判定人脸的身份起到了重要的作用。从这一角度出发,文献[70]利用了SIFT特征对皮肤上斑痣点特征进行建模表示,并与其他特征进行融合以提高分类的性能。文献[97]提出了一种独立于姿态和光照的皮肤显著点提取框架:该框架首先利用多尺度模板匹配过程来检测斑/痣候选点,然后利用3D形变模型(Morphable Model, MM)对人脸图像进行二值分割以去除非皮肤区域的候选点,最后对每个候选点计算显著性度量,并选择度量值高于设定阈值的候选点。为了准确地检测面部的斑/痣点特征,文献[46]首先利用主动表观模型(Active Appearance Model, AAM)来定位和分割面部主要特征(如眼睛、鼻子和嘴巴),然后利用高斯-拉普拉斯(Laplacian-of-Gaussian, LoG)算子和形态学算子来检测这些特征点。然而,由采集环境、光照、姿态等带来的噪声使得从人脸图像上难以稳定地提取这些显著的纹理特征。此外,在检测到这些特征点之后,如何表示也是一个重要的问题。尽管

如此,在某些条件下(如光照可控、图像分辨率比较高),该种特征与其他类型特征的结合仍可能提高分类精度[46]。

1.4.3 对比讨论

第 1.4.1 节和第 1.4.2 节分别概述了人脸识别中常用的全局和局部表示方法。为了清楚起见,表 1-2 列出了这些表示方法的类别、简要说明以及每一类中的代表性方法。

表 1-2 人脸识别中常用的全局与局部表示方法

	类别	简要说明	代表方法
全局表示	子空间表示	基于主成分分析的表示	Eigenfaces[119], 2D PCA[132]
		基于线性判别分析的表示	Fisherfaces[4], 2D LDA[133][139]
		其他的子空间表示	贝叶斯表示[86][87], ICA[3]
	流形表示	借助于流形方法来寻找隐藏在 高维样本空间内有意义的低维表示	LLE [99], ISOMAP[118]; LPP[38], UDP[134]
	稀疏表示	利用尽可能少的元素来降低信号或 图像的重构残差	SRC[124]
	全局频域表示	利用离散余弦变换特征的表示	DCT 低频系数[37]
		利用傅里叶变换特征的表示	混合的傅里叶特征[44]
局部表示	几何结构特征	利用面部特征和面部轮廓的相对位 置和大小等作为特征	面部特征及相对位置[8][49] 可变形模板[141]
	局部统计分析表示	基于全局图像的局部统计分析	LFA[90], LNMF[64], LS-ICA[51]
		基于图像部件或块的统计分析表示	EigenEyes 等[91], Simile 分类器[56]
	局部频域表示	弹性图匹配	DLM[57], EBGM[126]
		小波表示的特征提取	GFC[71], AdaGabor[138]
	局部纹理表示	局部模式表示	LBP[1], LGBP[150]
		梯度特征及其扩展	SIFT[6][79], HOG[2]
		局部显著性特征(如痣、胎记)	Skin Detail[46][97]

在全局表示方法中,子空间表示和流形表示将整幅人脸图像看作高维样本空间中的一个点来处理。在人脸图像存在局部变化(如光照、表情、部分遮挡)时,这些表示将受到比较大的影响。与之相对,局部表示能够提取对这些变化鲁棒的特征,因而在人脸识别中受到了更多的关注。特别地,Gabor 特征对人脸的光照、表情和部分遮挡等变化具有比较好的鲁棒性,并且,多尺度、多方向的 Gabor 小波滤波增强了人脸图像上的显著性特征。所以,图像的 Gabor 小波表示在人脸识别中得到了广泛的应用。

1.5 问题的提出及本文的主要贡献

1.5.1 问题的提出

虽然图像的 Gabor 小波表示在人脸识别中吸引了大量研究者的目光,然而,大多数

特征表示方法只利用了 Gabor 小波表示的幅值特征,而忽略了相位特征。主要的原因在于, Gabor 相位特征对位置变化非常敏感,这导致了利用 Gabor 相位特征的表示通常取得很低的分类精度。针对 Gabor 相位特征的这一特点,如何设计有效的特征表示方法以提高其在人脸识别中的分类性能;进一步地,如何将该特征表示与幅值特征进行融合以利用二者之间的互补性。这些问题的研究对发掘 Gabor 相位特征在人脸识别上的有效性有着重要的作用。

人脸图像的 Gabor 特征与 LBP 算子结合的局部纹理表示在人脸识别上取得了良好的效果。尽管如此,针对 Gabor 幅值表示反映了人脸图像上不同尺度、不同方向上的显著性特征这一特点,是否存在更有效的局部纹理表示方法?此外,人脸作为结构非常相似的一类物体,应当有不同于其他物体的纹理表示方法。那么,如何获得人脸专属的纹理表示并将其应用于人脸识别。这些问题的研究对改进已有的人脸表示方法有着重要的意义。

1.5.2 本文的主要贡献

本文围绕人脸图像的 Gabor 小波表示,从局部纹理表示的角度,提出了三种局部 Gabor 模式表示方法;进一步地,本文将判别性特征提取方法应用于 Gabor 特征的局部模式表示,以提高识别算法的精度。本文的主要贡献总结如下:

1. 提出了融合 Gabor 幅值和相位局部模式的表示方法。

为了根据人脸图像的 Gabor 相位特征设计有效的表示,该方法提出了“局部 Gabor 异或模式”表示:该表示先将相位特征量化到不同的区间,然后利用异或算子计算空间域内像素的局部模式,并采用直方图来表示人脸图像。进一步地,为了利用 Gabor 幅值和相位特征之间的互补性,该方法采用了基于块的 Fisher 线性判别分析对 Gabor 幅值和相位的局部模式表示进行融合并计算得到图像块之间的相似度,然后通过和规则融合各个块之间的相似度。由于提取了低维的判别特征,该方法不仅减少了图像相似度的计算代价,而且提高了识别精度。此外,局部 Gabor 异或模式是一种有效的基于相位特征的表示,保证了其与幅值局部模式融合的表达方法能够取得更优的识别性能。

2. 提出了 Gabor 特征“体”局部模式的表示方法。

为了高效地建模人脸图像 Gabor 特征的局部模式,该方法将多尺度、多方向的 Gabor 特征组织为三维的“体”:该特征“体”的前两维是图像平面,第三维对应了不同尺度、不同方向的 Gabor 小波。然后,该方法采用了局部二值模式算子来计算 Gabor 幅值特征“体”上的模式。文中进一步提出了该表示的两种扩展方法:利用 Fisher 准则的子块权重学习和判别性特征提取。子块权重学习方法根据 Fisher 准则基于训练集学习得到各个子块的权重,并根据这些权重来融合图像上不同子块之间的相似度。该方法利用了人脸不同区域的判别能力,在训练集与测试集中图像的变化保持一致时能够提高识别的性能。判别性特征提取方法根据 Fisher 线性判别分析从图像块级别上提取低维特征并计算

图像块之间相似度,然后通过和规则融合这些相似度。该方法提取了有利于分类的投影方向,在测试的三个人脸数据库上都表现出了更优的分类性能。

3. 提出了基于学习的局部 Gabor 模式表示方法。

为了学习人脸专属的模式表示,针对每个尺度、每个方向的 Gabor 小波,该方法从人脸图像的 Gabor 特征上采样得到图像块的集合,然后利用聚类算法从该集合上学习相应的码本,并以此表示人脸图像。考虑到人脸的不同区域和不同模式在识别过程中的不同作用,文中进一步提出了利用子块和模式权重的加权方法。此外,针对目前文献中存在的 11 种图像 Gabor 特征的局部模式表示,文中从模式定义流程以及方法参数两个方面对这些方法进行了对比,并在三个大规模人脸库上测试了这些方法及其与判别性特征提取结合方法的结果。大量的对比实验表明,基于学习的局部 Gabor 模式表示是一种有效的人脸表示,特别地,该方法受不同训练集上产生的码本的影响比较小;加权方法同时利用了子块和模式的判别能力,在识别精度上有进一步的提高;结合判别性特征提取的方法提取了低维的判别特征,取得了比基准方法和加权方法更高的识别性能。

1.6 本文的组织结构

本文的组织结构如下:

第一章为绪论部分,介绍了研究的背景及意义,并对人脸识别问题进行了描述,然后概述了与人脸识别研究相关的综述以及人脸识别的评测,之后综述了人脸识别中的特征表示方法,最后提出了本文的研究问题并总结了本文的主要贡献。

第二章提出了融合 Gabor 幅值和相位局部模式的表示方法,系统地测试了不同参数对该方法的影响,并与已有的识别方法进行了对比分析。

第三章提出了 Gabor 特征“体”局部模式的表示方法与它的两种扩展方法,系统地测试了该方法及其两种扩展方法在三个大规模人脸库上的结果,并进行了相应的分析。

第四章提出了基于学习的局部 Gabor 模式表示方法及其加权的方法,对比了不同局部 Gabor 模式表示方法之间的异同,并对提出的基于学习的局部 Gabor 模式方法、加权方法以及不同局部 Gabor 模式表示在三个大规模人脸数据库上的结果进行了对比分析。

第五章是结束语,在对全文工作总结的基础上,讨论了本文工作可能的后续扩展。

图 1-4 绘出了本文主要章节之间的关系图。如图中所示,第二章的 Gabor 幅值和相位特征的局部模式表示方法都是定义在图像块内,第三章提出的局部模式则定义在 Gabor 特征“体”上。这两章内容之间的关系可看作模式定义由 Gabor 特征“块”扩展到“体”上。第二章和第三章中局部模式表示的共同之处在于,它们都属于模式设计方法,即根据人工设计的算子来计算 Gabor 特征上的局部模式。不同于第二章和第三章中的模式设计方法,第四章中的人脸局部 Gabor 模式通过学习得到。第四章与第二、三章内容之间的关系可看作从“模式设计”扩展到“模式学习”。此外,这三章内容之间的一个共同之处在于,它们都采用了判别性特征提取方法从相应局部 Gabor 模式的表示中

提取低维的判别性特征。

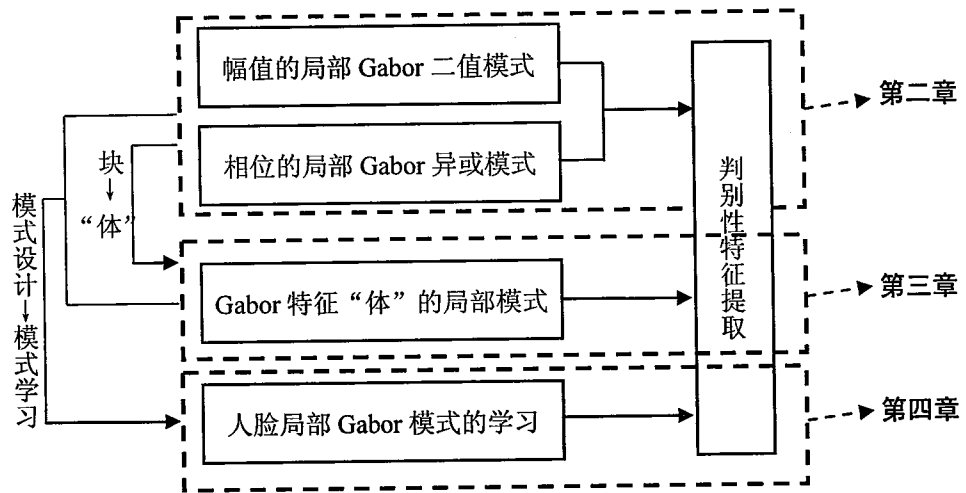


图 1-4 本文各章内容之间的关系

第二章 融合 Gabor 幅值和相位局部模式的表示方法

2.1 引言

近年来,融合不同类型特征的表示方法在人脸识别中受到了关注,并在一些富有挑战性的人脸数据库(如 FRGC 2.0)上取得了良好的性能,如融合频域中不同频段特征的表示[44][77]、融合不同尺度上特征的表示[70]、融合全局和局部特征的表示[109]以及融合 Gabor 特征和局部二值模式(Local Binary Patterns, LBP)特征的表示[117]。这些表示方法利用了不同类型特征之间的互补性,能够取得比利用单一特征的表示方法更好的分类效果。人脸图像的 Gabor 特征作为一种有效的表示,其幅值和相位特征具有不同的特点:幅值特征比较平滑,受位置变化的影响不大,而相位特征对位置变化比较敏感。为了利用这两种特征之间的互补性,一些研究者提出了融合 Gabor 幅值和相位特征的方法,如根据幅值和相位特征分别构建弱分类器进行融合[148]、加权的 Gabor 复小波特征表示[27]和增强的局部 Gabor 模式表示[151]。然而,大多数利用相位特征的方法[27][148]通常取得了很低的分类性能。这其中的原因在于,相位特征对位置变化比较敏感:在图像上相邻的两个位置即使只差了几个像素,它们的相位值会偏差得非常大,从而导致了直接利用相位特征来匹配两幅人脸图像时会产生比较大的误差。

针对 Gabor 相位特征的上述特点,本章提出了基于相位特征的局部 Gabor 异或模式表示。该表示先将图像的相位特征量化到不同的区间,然后利用异或(XOR)算子来计算空间邻域内像素的局部模式,并采用分块的直方图来表示人脸图像。其优点在于,采用的量化策略减弱了局部模式受相位对位置变化敏感性的影响,而多尺度、多方向的直方图特征具有很强的表示能力。进一步地,文中采用了幅值特征的局部 Gabor 二值模式表示[150],并利用判别性特征提取方法在块级别上对这两种局部模式表示进行融合。判别性特征提取方法利用了不同类别之间的信息来获得有利于分类的特征,因而极大地提高了分类精度;而且,它提取特征的维数远远低于 Gabor 特征局部模式表示的维数,从而减少了图像相似度的计算代价。此外, Gabor 幅值和相位局部模式特征之间的互补性对提高识别性能也起到了一定的作用。

本章内容安排如下:第 2.2 节介绍了 Gabor 小波表示的局部模式;第 2.3 节提出了局部 Gabor 异或模式表示,并对比分析了其与以前的局部 Gabor 模式表示方法的异同;第 2.4 节提出了融合 Gabor 幅值和相位局部模式的表示方法,主要从判别性特征提取和块级别上融合两个方面进行介绍;第 2.5 节测试了不同参数对不同方法的影响,并将本章方法与其他方法的结果进行了对比。

2.2 Gabor 小波表示的局部模式

Gabor 小波的实质是以高斯函数作为窗函数的短时傅里叶变换。1985 年, Daugman[21]将 Gabor 小波从一维扩展到二维。二维 Gabor 小波的核函数定义如下:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{(-\|k_{\mu,v}\|^2 \|z\|^2 / 2\sigma^2)} \left[e^{ik_{\mu,v}z} - e^{-\sigma^2/2} \right] \quad (2-1)$$

其中, $\|\cdot\|$ 表示取模操作, 小波向量 $k_{\mu,v}$ 的定义如下式所示:

$$k_{\mu,v} = [k_v \cos \phi_\mu, k_v \sin \phi_\mu] \quad (2-2)$$

这里, $k_v = k_{\max} / f^v$, $\phi_\mu = \pi\mu / 8$, $k_{\max}$ 是最大频率, f 是不同频率核函数之间的间隔。

在式 (2-1) 中, 参数 $k_{\mu,v}$ 控制着高斯窗口的宽度、振荡部分的波长和方向, 参数 σ 决定了窗口宽度和波长之间的比例关系, 亦即高斯包络函数所影响的振荡数目。式 (2-1) 实际上定义了一组在平移、旋转和尺度变化意义下自相似的小波函数。也就是说, 该公式定义的所有核函数均可以通过对某一 Gabor 小波核函数进行适当的平移、旋转和尺度变化得到。图 2-1 示出了 5 个尺度、8 个方向 Gabor 小波核函数的实部和虚部图像。

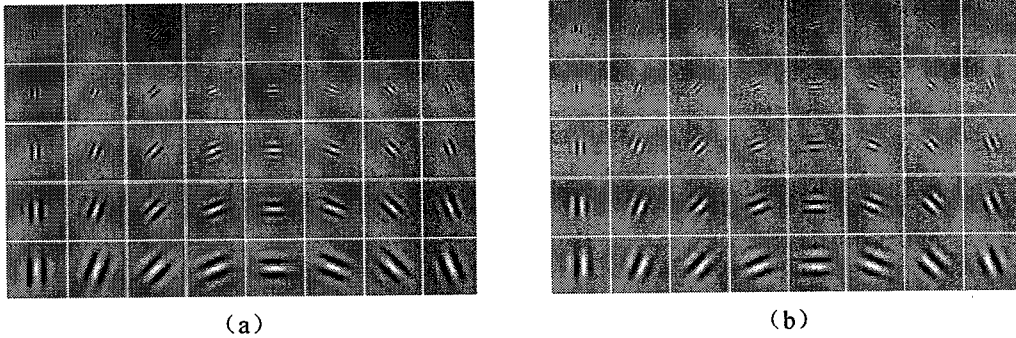


图 2-1 5 个尺度、8 个方向的 Gabor 小波核函数图示 (a) 实部 (b) 虚部

图像的 Gabor 小波表示定义为一幅图像与不同尺度、不同方向 Gabor 小波的卷积, 如式 (2-3) 所示:

$$G_{\mu,v}(z) = I(z) * \psi_{\mu,v}(z) \quad (2-3)$$

其中, $I(\cdot)$ 表示一幅图像, $\psi_{\mu,v}(\cdot)$ 表示式 (2-1) 定义的 Gabor 小波, z 表示图像 $I(\cdot)$ 中的像素位置。由于 Gabor 小波定义为复数形式, 卷积过程会产生一个由实部分量 $\text{Re}_{\mu,v}(\cdot)$ 和虚部分量 $\text{Im}_{\mu,v}(\cdot)$ 构成的复数响应。根据这两部分, 可进一步求得图像的幅值和相位部分, 分别如式 (2-4) 和式 (2-5) 所示:

$$A_{\mu,v}(z) = \sqrt{\text{Im}_{\mu,v}^2(z) + \text{Re}_{\mu,v}^2(z)} \quad (2-4)$$

$$\Phi_{\mu,v}(z) = \arctan(\text{Im}_{\mu,v}(z) / \text{Re}_{\mu,v}(z)) \quad (2-5)$$

图像Gabor小波表示的这几部分具有不同的特点：在物体的边缘附近，实部和虚部会产生振荡，而不是一个平滑的峰值响应，因而不利于识别阶段的匹配；幅值信息反映了图像局部的能量谱，也可以理解为特定方向边缘的强度，对图像的匹配和识别来说具有重要的意义；相位信息则对图像的位置变化敏感，即使相邻位置在图像上仅相差几个像素，其对应的相位特征也会有很大差别，因此直接用于图像匹配时会产生很大的问题。图2-2 (a) ~ (c) 分别示出了一幅人脸图像及其5个尺度、8个方向的Gabor幅值和相位特征图。

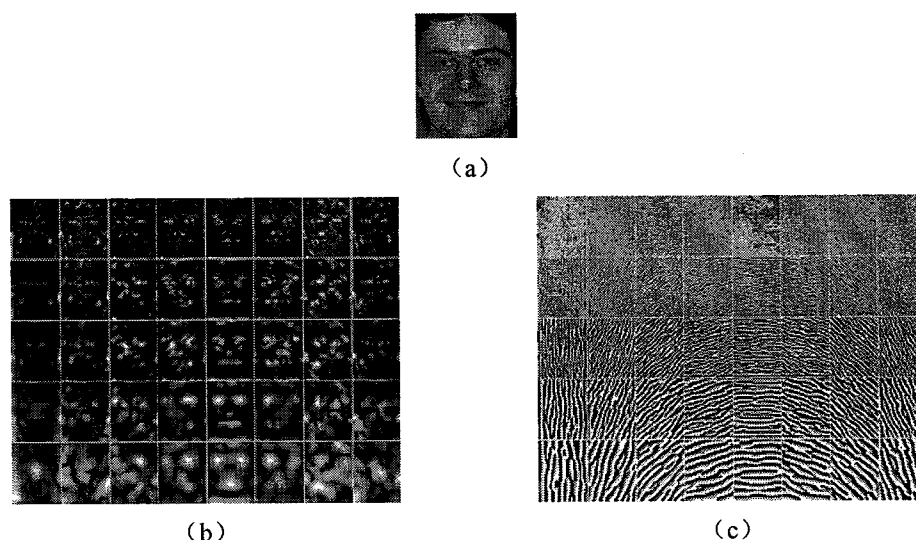


图 2-2 一幅人脸图像与它的 Gabor 小波表示 (a) 人脸图像 (b) 幅值 (c) 相位

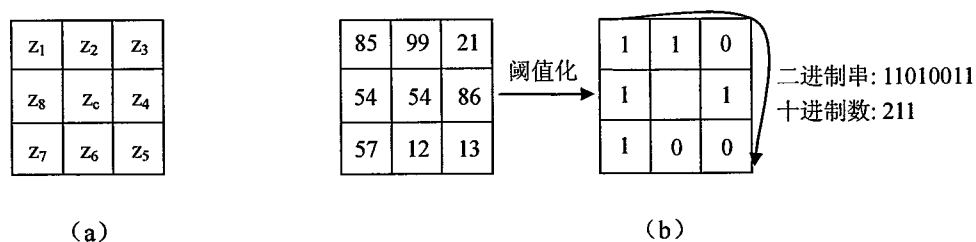


图 2-3 LBP 算子示意图 (a) 3×3 邻域 (b) LBP 算子

近年来，LBP 表示在纹理分类、人脸检测与识别等领域里受到了广泛关注 [1][36][88]。其基本思路如下：在图像中像素所在的 3×3 邻域内，每个像素根据中心像素的亮度值进行阈值化得到一个二进制数（即“0”或“1”），然后将这些“0”、“1”值按一定顺序组合得到一个二进制串来表示中心像素。如图 2-3 所示，“11010011”是中心像素 z_c 的局部模式。一幅人脸图像在应用 LBP 算子之后，得到一幅对应的模式图像；然后，该模式图像被划分为若干个子块并计算每个子块的直方图；最后，所有子块的直

方图连接起来得到整幅图像的特征表示[1]。这里提到“子块”概念的主要原因是区分判别性特征提取方法中“块”的概念（参见本章 2.4.1 节）。

为了从 Gabor 特征上提取有效的模式表示, Zhang 等利用 LBP 算子分别计算幅值和相位特征的局部模式, 提出了幅值的局部 Gabor 二值模式 (Local Gabor Binary Patterns of Magnitude, LGBP_Mag) [150] 和相位的局部 Gabor 二值模式 (Local Gabor Binary Patterns of Phase, LGBP_Pha) [151] 两种表示方法。由于采用了多尺度、多方向的 Gabor 小波, 一幅图像会相应地得到多尺度、多方向的局部 Gabor 模式图像。为了有效地表示图像的特征, Zhang 等采用了图像分块的直方图表示, 即: 每幅模式图像被划分为若干个子块并计算每个子块的直方图, 然后将不同尺度、不同方向上各个子块的直方图连接起来表示人脸图像。受 Gabor 相位特征在虹膜识别[22]和掌纹识别[145]等方面成功应用的启发, Zhang 等[142]提出了实部局部 Gabor 相位模式 (Local Gabor Phase Patterns of Real part, Re_LGPP) 和虚部 Gabor 相位模式 (Local Gabor Phase Patterns of Imaginary part, Im_LGPP)。在这两种表示中, 异或 (XOR) 算子用来分别计算实部和虚部特征上的局部模式。类似地, Zhang 等[142]也采用了图像分块的直方图表示, 并验证了该方法具有比幅值特征的局部 Gabor 二值模式 (即 LGBP_Mag) 更优的表示能力。

2.3 局部 Gabor 异或模式及其与以前局部模式的对比

针对 Gabor 相位特征具有对位置变化敏感性的特点, 局部 Gabor 异或模式 (Local Gabor XOR Patterns, LGXP) 中采用了一种“软” (Soft) 方式来计算相位特征上的局部模式: 如果空间邻域内两个相位角的数值处于同一个区间内, 它们可能反映了相似的局部特征, 其模式赋值为“0”; 否则, 它们可能代表了不同的局部特征, 其模式赋值为“1”。这种策略一定程度上增强了模式对位置变化的鲁棒性: 只要相位角的变化幅度没有超过设定的间隔, 其对应的模式就保持不变。图 2-4 示出了 LGXP 方法在 3×3 邻域上的编码示例。

简要地讲, 在 LGXP 方法中, 相位角首先根据设定的间隔被量化为不同的数值, 然后利用异或 (XOR) 算子对中心像素与其邻域像素的相位量化值进行运算得到二进制位 (即“0”和“1”), 最后将这些二进制位按照一定顺序连接起来得到二进制串来表示中心像素。在 LGXP 表示中, 中心像素 z_c 的局部模式 $LGXP_{\mu,v}(z_c)$ 的二进制和十进制形式定义如下:

$$LGXP_{\mu,v}(z_c) = [LGXP_{\mu,v}^P, LGXP_{\mu,v}^{P-1}, \dots, LGXP_{\mu,v}^1]_{binary} = \left[\sum_{i=1}^P 2^{i-1} \cdot LGXP_{\mu,v}^i \right]_{decimal} \quad (2-6)$$

其中, P 表示空间邻域内像素数目, 其数值大小决定了模式数目; $LGXP_{\mu,v}^i$ ($i=1, 2, \dots, P$)

表示中心像素 z_c 与邻域像素 z_i 通过异或算子运算得到的二进制位, 其定义如式 (2-7)

所示:

$$\text{LGXP}_{\mu,\nu}^i = q(\Phi_{\mu,\nu}(z_c)) \otimes q(\Phi_{\mu,\nu}(z_i)), i = 1, 2, \dots, P \quad (2-7)$$

这里, $\Phi_{\mu,\nu}(\bullet)$ 表示相位角; $\otimes$ 表示异或算子, 定义如式 (2-8) 所示; $q(\bullet)$ 表示量化算子, 根据设定的量化区间来计算相位角的量化码, 如式 (2-9) 所示。

$$c \otimes d = \begin{cases} 0, & \text{if } c = d \\ 1, & \text{else} \end{cases} \quad (2-8)$$

$$q(\Phi_{\mu,\nu}(\bullet)) = i; \text{ if } \frac{360}{b} * i \leq \Phi_{\mu,\nu}(\bullet) < \frac{360}{b} * (i+1), i = 0, 1, \dots, b-1 \quad (2-9)$$

其中, b 表示相位角的量化区间。

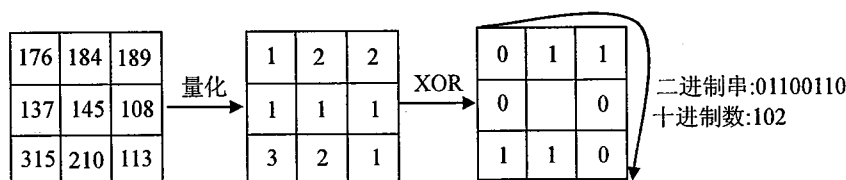


图 2-4 LGXP 方法中的模式编码示例 注: 相位角被量化为 4 个区间

根据上面定义的模式计算方法, 每个 Gabor 小波对应的相位图像都会计算得到一幅相应的模式图像。与 LBP 表示类似, 这里也采用了分块的直方图表示, 即: 每幅模式图像被划分为 m 个不重叠的子块并计算该子块的直方图, 最后将所有尺度、所有方向上各个子块的直方图连接起来得到整幅图像的代表:

$$H = [H_{\mu_0, \nu_0, 1}, \dots, H_{\mu_0, \nu_0, m}; \dots; H_{\mu_{s-1}, \nu_{s-1}, 1}, \dots, H_{\mu_{s-1}, \nu_{s-1}, m}] \quad (2-10)$$

其中, $H_{\mu,\nu,i}$ ($\mu = \mu_0, \dots, \mu_{s-1}$; $\nu = \nu_0, \dots, \nu_{s-1}$; $i = 1, 2, \dots, m$) 表示第 ν 个尺度、第 μ 个方向的模式图像第 i 个子块的直方图。对于两幅人脸图像 I_1 和 I_2 的直方图序列 H^1 和 H^2 , 它们的相似度按照下式计算:

$$S(H^1, H^2) = \sum_{\mu=\mu_0}^{\mu_{s-1}} \sum_{\nu=\nu_0}^{\nu_{s-1}} \sum_{i=1}^m S(H_{\mu,\nu,i}^1, H_{\mu,\nu,i}^2) \quad (2-11)$$

其中, $S(\bullet, \bullet)$ 表示两个直方图之间的相似度。这里采用了直方图交算子, 其定义如下式所示:

$$S(h^1, h^2) = \sum_{i=1}^L \min(h_i^1, h_i^2) \quad (2-12)$$

其中, h^1 和 h^2 分别表示两个直方图, L 表示直方图中 bin 的数目。

显然, 对于 LGXP 表示而言, 量化区间 b 的大小对于刻画相位特征的局部模式至关

重要：如果 b 值太大（例如， $b=360$ ），任意两个不同的相位角就会落入不同的区间，量化操作就失去了本身的意义；如果 b 值太小（例如， $b=1$ ），所有的相位角就会落入同一个区间，模式就失去了对不同粒度特征的区别性。由此可见， b 值的选择在位置变化敏感性与不同粒度特征的区别性之间存在折中。此外，尽管相位随位置的变化比较敏感，采用的量化操作使得局部模式对其影响较小，而多尺度、多方向的直方图特征具有很强的表示能力。因此，LGXP 表示具有比相位特征更强的表示能力。

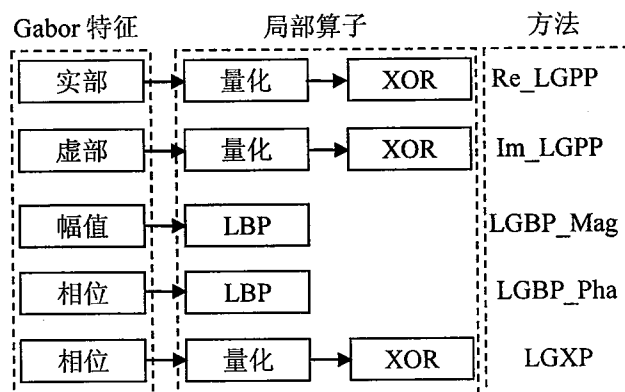


图 2-5 不同局部 Gabor 模式表示的流程对比

为了强调局部 Gabor 异或模式 (LGXP) 与其他局部 Gabor 模式之间的差异，图 2-5 中示出了这些方法的简要流程。如图所示，它们的不同主要体现在两方面：Gabor 特征和局部算子。具体来讲，实部和虚部特征的局部 Gabor 相位模式（即 Re_LGPP 和 Im_LGPP）分别基于 Gabor 的实部和虚部特征并利用了量化（实部或虚部的响应根据其符号量化为“0”或“1”）和 XOR 算子；幅值和相位特征的局部 Gabor 二值模式（即 LGBP_Mag 和 LGBP_Pha）方法利用 LBP 算子来分别计算幅值和相位特征的局部模式；局部 Gabor 异或模式利用了量化和 XOR 算子来计算相位特征的局部模式。

由于 LGXP 表示中采用了量化操作，即使当相邻像素的相位值变化很大时，如果变化的幅度仍未超过设定的量化间隔，它们的量化值仍保持不变。对相位的局部 Gabor 二值模式（即 LGBP_Pha）而言，LBP 算子容易受噪声的影响：如果相邻像素的序关系发生改变，那么计算的模式就会发生改变，而这种情况经常出现在人脸的一些接近平坦的区域[116]。因此，LGXP 表示应当比 LGBP_Pha 表示更适于描述人脸。

在某种意义上，LGXP 表示可看作 Re_LGPP 与 Im_LGPP 方法在编码层上的一种融合。具体来讲，当 $b=4$ 时，4 个相位角区间（即 $[0^\circ, 90^\circ)$, $[90^\circ, 180^\circ)$, $[180^\circ, 270^\circ)$, $[270^\circ, 360^\circ)$ ）对应了 LGPP 方法中实部和虚部符号的 4 种组合（即“00”，“10”，“11”和“01”，这里“0”和“1”分别表示根据实部和虚部符号量化的数值[142]）。此时，LGXP 表示同时利用了实部和虚部特征来编码图像空间邻域内的局部模式。因此，相比 Re_LGPP 和 Im_LGPP 表示，LGXP 表示具有更强的人脸表示能力。

2.4 Gabor 幅值和相位局部模式的融合方法

为了利用 Gabor 幅值和相位特征之间的互补性,一些研究者提出了基于二者融合的认识方法。Zhang 等[148]将多类的人脸识别问题转化为两类问题,然后利用瀑布框架从幅值和相位特征上构建 AdaBoost 分类器:在前一个阶段利用幅值特征构建强分类器,在后一个阶段利用相位特征构建的分类器对未分类好的样本进行分类;Gao 等[27]通过合适的权重将幅值和相位特征结合为复数向量,然后将 PCA 和 LDA 方法扩展到复数向量构成的空间内,并提出了该空间内的余弦相似度进行识别;Zhang 等[151]在利用 LBP 算子分别对幅值和相位特征建模表示后,将二者的直方图特征连接起来作为一种新的特征表示。

尽管文献[27]和[148]中融合幅值和相位特征的方法取得了比利用单一特征的方法更好的效果,但是,相位具有的位置敏感性使得基于相位特征的方法具有很低的分类精度。在文献[151]的方法中,幅值和相位的局部模式表示具有高维的特征,这给融合这两种特征的方法带来了极大的计算代价。由于局部 Gabor 异或模式比相位特征更有效,本部分将对 Gabor 幅值和相位的局部模式特征进行融合。为了降低局部 Gabor 模式表示的特征维数,下面采用了 Fisher 线性判别分析在图像的块级别上对这两种特征进行融合。相关内容安排如下:第 2.4.1 节介绍了判别性特征提取方法;第 2.4.2 节主要介绍了幅值和相位特征在块级别上的融合方法。

2.4.1 判别性特征提取

为了降低局部 Gabor 模式表示的特征维数,本节采用了 Fisher 线性判别分析 (Fisher's Linear Discriminant Analysis, FLD 或 LDA) 方法来提取低维的判别性特征。由于样本的特征维数非常高,直接应用 FLD 方法会带来“小样本”(Small Sample Size, 3S)问题。因此,这里采用了基于块的 Fisher 线性判别分析 (Block-based Fisher's Linear Discriminant Analysis, BFLD) 方法。

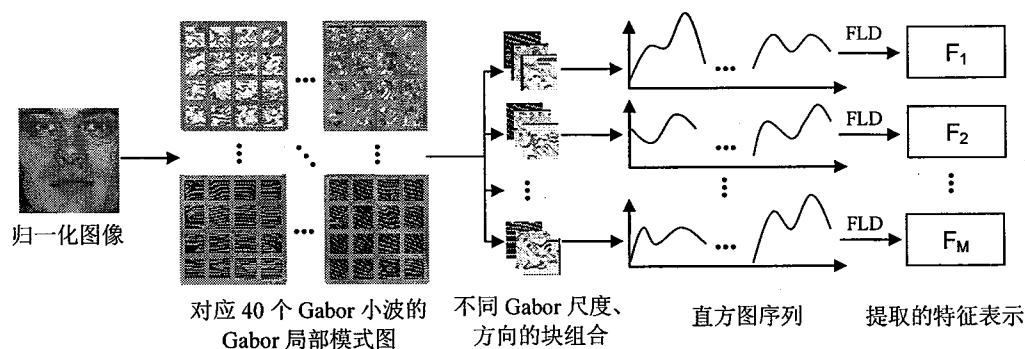


图 2-6 基于块的 Fisher 线性判别分析方法的流程

算法 2.1 FLD矩阵的学习过程

输入:	N 幅人脸图像构成的训练集 T ; 图像块的数目 M ; 每个块的子块数目 K ; 局部Gabor模式空间域内的像素数目 P 。
输出:	M 个FLD矩阵 $W_i (i=1, 2, \dots, M)$ 。
步骤 1	对训练集 T 中每幅图像 I , 根据步骤1.1~1.3计算其局部Gabor模式的分块直方图表示 $H_i (i=1, 2, \dots, M)$:
步骤 1.1	计算图像 I 的局部Gabor模式图像 $\mathfrak{R}_{\mu, \nu} (\mu = \mu_0, \mu_1, \dots, \mu_{o-1}, \nu = \nu_0, \nu_1, \dots, \nu_{s-1})$ 。
步骤 1.2	将 $\mathfrak{R}_{\mu, \nu}$ 划分为 M 个不重叠的块 $\mathfrak{R}_{\mu, \nu, i} (i=1, 2, \dots, M)$, 并将每个块 $\mathfrak{R}_{\mu, \nu, i}$ 划分为 K 个不重叠的子块 $\mathfrak{R}_{\mu, \nu, i}^j (j=1, 2, \dots, K)$ 。对子块 $\mathfrak{R}_{\mu, \nu, i}^j$, 计算它的直方图 $H_{\mu, \nu, i}^j$ 。
步骤 1.3	对每个块 $\mathfrak{R}_i = [\mathfrak{R}_{\mu_0, \nu_0, i}, \dots, \mathfrak{R}_{\mu_0, \nu_{s-1}, i}, \dots, \mathfrak{R}_{\mu_{o-1}, \nu_0, i}, \dots, \mathfrak{R}_{\mu_{o-1}, \nu_{s-1}, i}] (i=1, 2, \dots, M)$, 将它 在所有尺度、所有方向上各个子块的直方图连接起来得到一个直方图序列: $H_i = [H_{\mu_0, \nu_0, i}, \dots, H_{\mu_0, \nu_{s-1}, i}, \dots, H_{\mu_{o-1}, \nu_0, i}, \dots, H_{\mu_{o-1}, \nu_{s-1}, i}]$ 。其中, $H_{\mu, \nu, i} (\mu = \mu_0, \mu_1, \dots, \mu_{o-1}, \nu = \nu_0, \nu_1, \dots, \nu_{s-1}, i=1, 2, \dots, M)$ 表示块 $\mathfrak{R}_i$ 在第 ν 个尺度、第 μ 个方向上 K 个子块的直方图连接起来得到的序列, 即: $H_{\mu, \nu, i} = [H_{\mu, \nu, i}^1, H_{\mu, \nu, i}^2, \dots, H_{\mu, \nu, i}^K]$ 。因此, 图像 I 表示为 M 个直方图序列 $H_i (i=1, 2, \dots, M)$ 。
步骤 2	利用训练集 T 根据步骤2.1~2.2学习每个块的FLD矩阵 $W_i (i=1, 2, \dots, M)$:
步骤 2.1	对 M 个空间块, 将训练集 T 的 N 幅图像组合在一起得到 M 个特征集合 $S_i = \{H_{1,i}, H_{2,i}, \dots, H_{N,i}\} (i=1, 2, \dots, M)$ 。其中, $H_{j,i} (j=1, 2, \dots, N)$ 表示第 j 幅训练图像第 i 个块的直方图。
步骤 2.2	根据特征集合 $S_i (i=1, 2, \dots, M)$ (即第 i 个块的训练集), 按照“PCA+LDA”的方法来学习其相应的FLD投影矩阵 W_i , 即: 直方图特征先投影到一个PCA子空间里, 然后根据投影后的特征来学习FLD投影矩阵。

BFLD 方法的基本思想是采用“分而治之”的策略, 即: 首先将局部 Gabor 模式(如 LGXP) 特征划分为若干个特征段(对应于人脸图像的不同空间块), 然后每一段特征单

独利用 FLD 进行特征提取。通过选择合适的特征段数目, 每个特征段的维数会远远低于原始的特征维数, 从而避免了 3S 问题。图 2-6 示出了 BFLD 方法的简要流程。概括地讲, BFLD 方法需要两个阶段完成特征提取: 第一个阶段是 FLD 矩阵的学习过程, 即: 根据设定的图像块数目, 在训练集合上学习多个 FLD 投影矩阵; 第二个阶段是利用 FLD 矩阵进行特征提取, 即: 根据 FLD 矩阵计算图像上不同块的低维表示。为了清楚起见, 这里将 BFLD 方法的这两个阶段分别在算法 2.1 与算法 2.2 中进行介绍。

算法 2.2 利用FLD矩阵进行特征提取	
输入:	归一化图像 I ; 参数 M, K 和 P 的含义同算法 2.1; FLD 矩阵 $W_i (i = 1, 2, \dots, M)$ 。
输出:	图像 I 的 M 个低维特征向量 $F_i (i = 1, 2, \dots, M)$ 。
步骤 1	对输入图像 I , 根据算法 2.1 中的步骤 1.1~1.3 计算它的分块直方图表示 $H_i (i = 1, 2, \dots, M)$ 。
步骤 2	根据如下等式计算图像 I 的 M 个低维向量 $F_i (i = 1, 2, \dots, M)$: $F_i = (W_i)^T H_i$ 。

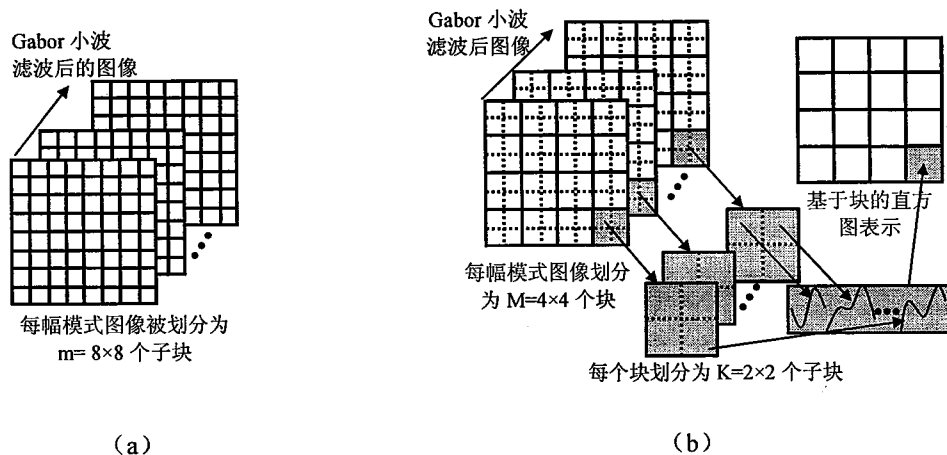


图 2-7 不同方法中的块划分策略 (a) 局部 Gabor 模式 (b) BFLD

值得注意的是, 局部 Gabor 模式表示 (如 LGXP) 与 BFLD 方法中采用了不同的块划分策略。如图 2-7 所示, 在局部 Gabor 模式表示中, 每个模式图被直接划分为 m 个子块; 而在 BFLD 方法中, 每个模式图先被划分为 M 个块, 然后每个块被划分为 K 个子块。直方图表示的一个不足之处在于, 它忽略了被统计特征的空间位置关系, 而只考虑特征出现的频数。因而, 在局部 Gabor 模式表示中, 当子块的尺寸比较小时该方法才能取得比较好的性能[1]。在 BFLD 方法中, 这种块划分策略可以保持属于同一个块的不同子块之间的相对位置关系。因此, 这种表示比每个块只采用一个直方图的表示 (即该块不划分子块) 具有更好的区分性。请注意区分局部 Gabor 模式表示中的“子块”与 BFLD 方法中“块”的关系: 在某种意义上, BFLD 方法中的一个块可看作局部 Gabor 模式表示中若干个子块的组合。当它们取得合适数值的时候, 局部 Gabor 模式的子块数

目 m 与 BFLD 方法的块数目 M 及每个块的子块数目 K 之间存在如下的关系: $m = M \times K$ (参见图 2-7: $m=8 \times 8$, $M=4 \times 4$, $K=2 \times 2$)。为了使子块(或块)的划分方式清楚起见,本章中采用“ $a \times b$ ”表示子块(或块)的数目。其中, a 表示水平方向上子块(或块)的数目, b 表示竖直方向上子块(或块)的数目。此外,为了区分这两种块划分策略,在本文后面的章节(第三章与第四章)中,局部 Gabor 模式表示方法中采用了“子块”进行描述, BFLD 方法中同时采用了“块”和“子块”来进行描述。

在局部 Gabor 模式表示(如 LGXP)同 BFLD 方法结合后,计算代价确实增加了一些:相比局部 Gabor 模式表示, BFLD 方法需要学习 M (即块数目)个投影矩阵。幸运的是,这些矩阵都是离线学习得到,并且在线应用时具有很低的计算复杂度。另一方面,经过判别性特征的提取,每幅图像的特征维数要远远低于局部 Gabor 模式表示的特征维数。在本章的实验部分,一幅人脸图像经 40 个 Gabor 小波滤波后得到 40 幅 LGXP 模式图像,当 $m=64$, $P=4$ 时,其 LGXP 表示的特征维数是 40,960 ($40 \times 64 \times 2^4$) 维;当 $M=16$, $K=4$ 时,每个块的特征维数经 FLD 方法降到 200 维(即 FLD 子空间的维数),该图像最终的特征维数是 3,200 (16×200) 维。由此可见, BFLD 方法极大地降低了样本的特征维数,从而减少了样本间相似度的计算代价。此外, BFLD 方法本身具有的判别性也带来了识别性能上的提高。所以,局部 Gabor 模式与 BFLD 结合的方法是一种更有效的人脸识别方法。

2.4.2 块级别上融合

为了进一步利用 Gabor 幅值和相位特征之间的互补性,本节采用了判别性特征提取方法对幅值特征的局部 Gabor 二值模式(即 LGBP_Mag) [150]和相位特征的局部 Gabor 异或模式(即 LGXP)进行融合。

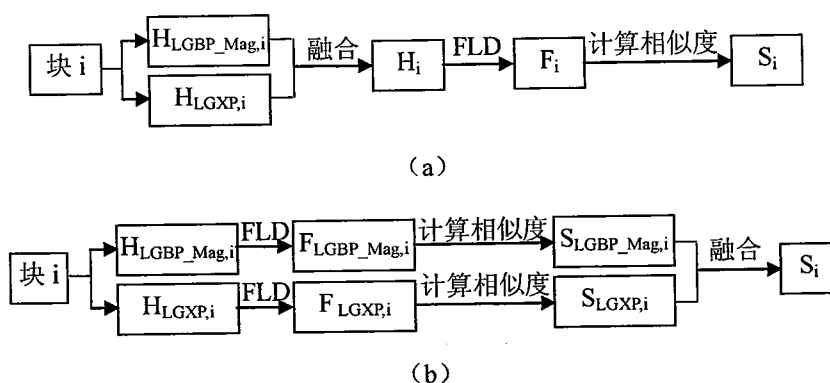


图 2-8 Gabor 幅值和相位局部模式的融合方法 (a) 特征层融合 (b) 相似度层融合

简要地讲,这两种类型的特征可以在图像的块级别上从特征层和相似度层两个方面分别进行融合:在特征层融合方法中,如图 2-8 (a) 所示,每个图像块的 LGBP_Mag 与 LGXP 的直方图特征串接为一个特征向量,然后利用 FLD 进行特征提取,进而根据

提取的特征计算出图像相应块之间的相似度；在相似度层融合方法中，如图 2-8 (b) 所示，每个块单独计算出它的 LGBP_Mag 与 LGXP 直方图特征并利用 FLD 提取相应的判别特征，然后分别根据这两种特征计算出图像块之间的两个相似度，并按照一定的权重将这两个相似度融合在一起。为了得到图像间的相似度，最后采用了和规则来融合各个块上的相似度。

为了清楚起见，这两种融合幅值和相位局部模式的方法可描述为如下步骤：

1. 计算分块直方图表示。对每幅图像，根据算法 2.1 中的步骤 1.1~1.3 分别计算幅值的局部 Gabor 二值模式和相位的局部 Gabor 异或模式的分块直方图表示： $H_{\text{LGBP_Mag},i}$ 和 $H_{\text{LGXP},i}$ ($i=1,2,\dots,M$)。其中， M 表示图像划分的块数目。

2. 融合 LGBP_Mag 与 LGXP 表示。

- i. 特征层融合。如图 2-8 (a) 所示，对图像的第 i 个块，将它的 LGBP_Mag 与 LGXP 直方图特征组合在一起，即： $H_i=[H_{\text{LGBP_Mag},i}, H_{\text{LGXP},i}]$ 。然后，根据算法 2.1 的步骤 2.1~2.2 计算其 FLD 投影矩阵 W_i ，并根据算法 2.2 的步骤 2 计算其低维特征 F_i 。

对注册图像第 i 个块 I_i^s 与测试图像第 i 个块 I_i^p ，它们的相似度由其提取的特征按照下式进行计算：

$$S(I_i^s, I_i^p) = \frac{F_i^s \cdot F_i^p}{\|F_i^s\| \cdot \|F_i^p\|} \quad (2-13)$$

其中， F_i^s 和 F_i^p 分别表示块 I_i^s 与 I_i^p 的低维特征。这里采用了余弦相似度来计算两个特征向量之间的相似度。

- ii. 相似度层融合。如图 2-8(b)所示，根据图像第 i 个块的两个直方图序列： $H_{\text{LGBP_Mag},i}$ 和 $H_{\text{LGXP},i}$ ，按照算法 2.1 的步骤 2.1~2.2 分别得到其对应的投影矩阵，进而依据算法 2.2 的步骤 2 分别得到它们的低维特征： $F_{\text{LGBP_Mag},i}$ 和 $F_{\text{LGXP},i}$ ；然后，对注册图像第 i 个块 I_i^s 与测试图像第 i 个块 I_i^p ，根据提取的低维特征 $F_{\text{LGBP_Mag},i}$ 和 $F_{\text{LGXP},i}$ 按照等式(2-13)分别计算两个相似度： $S_{\text{LGBP_Mag},i}$ 和 $S_{\text{LGXP},i}$ ；最后，相似度 $S_{\text{LGBP_Mag},i}$ 和 $S_{\text{LGXP},i}$ 按照加权和规则融合在一起得到块 I_i^s 与块 I_i^p 之间的最终相似度：

$$S(I_i^s, I_i^p) = w \cdot S_{\text{LGBP_Mag},i} + (1-w) \cdot S_{\text{LGXP},i} \quad (2-14)$$

其中， w 表示相似度 $S_{\text{LGBP_Mag},i}$ 所占的权重，其值的范围是 $[0, 1]$ 。

3. 计算两幅图像之间的相似度。根据步骤 2 计算得到的图像块之间的相似度 $S(I_i^s, I_i^p) (i=1, 2, \dots, M)$, 两幅图像 I^s 和 I^p 的相似度由这些相似度按照和规则融合得到:

$$S(I^s, I^p) = \frac{1}{M} \sum_{i=1}^M S(I_i^s, I_i^p) \quad (2-15)$$

4. 人脸分类。对于识别任务, 人脸图像的身份通常被判别为最相似人脸的身份, 或者, 如果最大相似度小于设定的阈值, 则“拒识”该图像; 对于确认任务, 首先计算出测试图像与注册图像之间的相似度, 然后根据该相似度大于或者小于设定的阈值来确认为同一人或不同人。

原则上, 在这两种融合方法中, 幅值特征的局部 Gabor 二值模式表示 (即 LGBP_Mag) 与相位特征的局部 Gabor 异或模式表示 (即 LGXP) 的参数 (如子块数目和空间域内像素的数目) 可以取不同的数值。然而, 这意味着更大的参数空间去选择合适的参数值。为了简单起见, 这两种表示方法中相同的参数设置为相同数值, 即: $K_{LGBP_Mag} = K_{LGXP}$, $P_{LGBP_Mag} = P_{LGXP}$ 。此外, 考虑到 LGBP_Mag 与 LGXP 表示在识别过程中起到了不同的作用, 因而相似度层融合方法中采用了加权和规则的融合方法。当 $w=0$ 或 $w=1$ 时, 该融合方法变为利用单一的 LGXP 或 LGBP_Mag 特征的识别方法。本章的实验结果也表明, 在二者取得相等权重 (即 $w=0.5$) 时, 融合方法基本上能够取得最优的分类性能。这说明了 LGXP 与 LGBP_Mag 表示在识别过程中起到了同等重要的作用。

除了上面的两种局部 Gabor 模式表示, 其他局部 Gabor 模式表示完全可以在该框架下进行融合, 如幅值的局部 Gabor 二值模式 (即 LGBP_Mag) 与相位的局部 Gabor 二值模式 (即 LGBP_Pha)、实部的局部 Gabor 相位模式 (即 Re_LGPP) 与虚部的局部 Gabor 相位模式 (即 Im_LGPP)。正如第 2.3 节所分析的, 与 LGBP_Pha、Re_LGPP 和 Im_LGPP 等表示相比, 相位的 LGXP 表示是一种更有效的局部模式表示。因此, 它与 LGBP_Mag 表示的融合取得了更高的识别精度。

2.5 实验结果对比及分析

本部分将对相位的局部 Gabor 异或模式 (即 LGXP)、判别性特征提取方法 (即 BFLD) 以及融合 Gabor 幅值和相位局部模式的方法在两个大规模人脸识别库上进行测试: FERET[93] 和 FRGC 2.0[94]。相关内容安排如下: 第 2.5.1 节介绍了这两个人脸数据库的基本情况; 第 2.5.2 节测试了不同参数取值对这三种方法结果的影响并给出了相应的分析; 第 2.5.3 节和第 2.5.4 节分别在 FERET 和 FRGC 2.0 人脸数据库上测试了这三种方法, 并与已有方法的结果进行了对比分析。

2.5.1 测试数据库简介

● FERET 人脸数据库

FERET 人脸数据库是由美国国防部通过 DARPA (Defense Advanced Research Products Agency, DARPA) 资助的 FERET (FacE REcognition technology Test) 计划构建, 用来测试不同识别方法的性能[93]。该人脸识别库总共包含了 14,051 幅多姿态的灰度人脸图像。在 1998 年到 2001 年之间发布的数据包含了 3,737 幅图像, 也被称为训练数据。该数据库严格划分了已知人脸的注册图像集合 (Gallery) 和测试图像集合 (Probe Set)。表 2-1 中列出了该数据库正面图像上各个集合的统计情况。图 2-9 绘出了 FERET 人脸数据库的部分示例图像。此外, 该数据库提供了手工标注的眼睛中心位置。

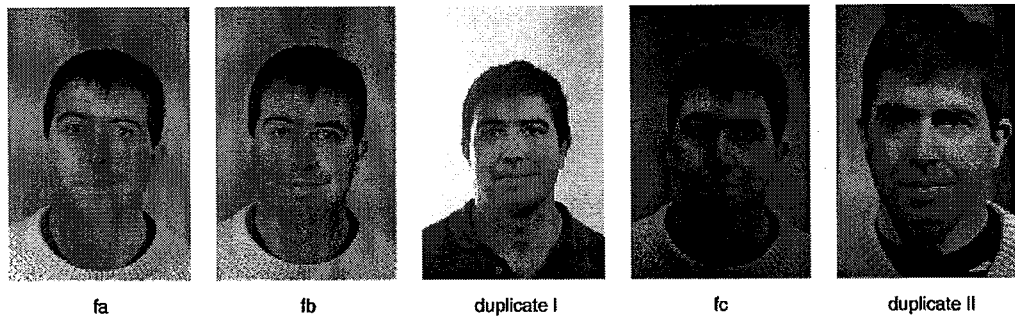


图 2-9 FERET 人脸库的图像示例

表 2-1 FERET 人脸库中各个集合的统计情况

数据集	子集	图像数	人数	说明
注册集	无	1,196	1,196	均匀光照条件
测试集	Fb	1,195	1,195	与注册集图像来自同一个集合, 表情不同
	Fc	194	194	与注册集图像同一天采集, 不同摄像头, 不同光照
	Dup.I	722	243	1,031 天内不定期采集
	Dup.II	234	75	Dup.I 的子集, 与注册集图像的采集间隔超过一年

● FRGC (Face Recognition Grand Challenge) 2.0 人脸数据库

FRGC 2.0[94]人脸数据库是目前已公开的最大规模的人脸数据库, 包含了在可控与非可控条件下采集的三维扫描图像和高分辨率的二维图像, 图像数目超过5万幅。图2-10绘出了一个对象采集的图像示例。FRGC 2.0数据分为训练集 (training partition) 和校验集(validation partition)两个部分。训练集是在2002年~2003年之间采集的, 包括两部分: 一部分是大规模静态图像训练集, 包含了222个对象的12,776幅图像 (可控条件下6,388幅图像与非可控条件下6,388幅图像); 另一部分是三维图像训练集, 包括三维扫描图像和来自943个时期的可控与非可控条件下采集的图像。校验集是在2003年~2004年之间采集的, 包含了466个对象、来自4,007个时期的图像。

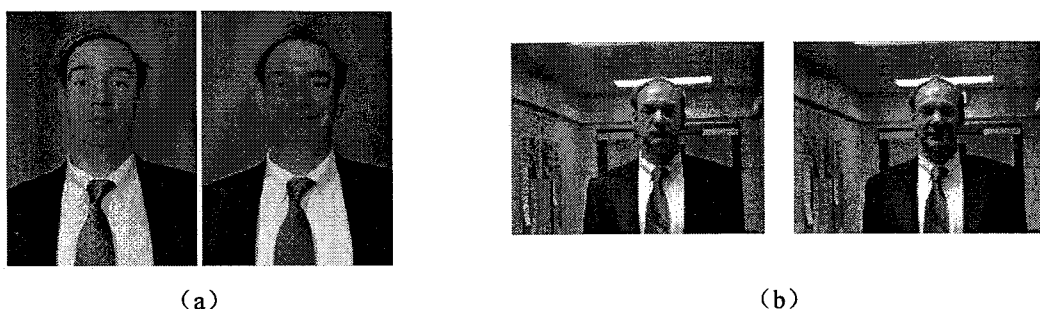


图2-10 FRGC 2.0人脸库的图像示例 (a) 可控条件下 (b) 非可控条件下

FRGC 2.0测试协议中提供了六个实验：实验1、2、4是针对二维图像上识别算法的评测，实验3、5、6是针对三维图像上识别算法的评测。在每个实验中，算法的输入包括两个集合：目标集 (Target Set) 和查询集 (Query Set)。其中，目标集中的图像身份是已知的，查询集中的图像身份是未知的。识别算法的输出是一个相似度矩阵，该矩阵包含了查询集与目标集中的所有图像对。根据该相似度矩阵，可计算得到两个指标：确认率 (Verification Rate, VR) 与错误接受率 (False Accept Rate, FAR)。确认率是根据同一人不同图像之间的比对计算得到的，而错误接受率是根据不同人图像之间的比对计算得到。对于每个实验，FRGC 2.0的测试协议还定义了3条不同的ROC (Receiving Operator Characteristic) 曲线，其区别在于目标集和查询集中图像的采集时间间隔不同。具体来说，ROC1对应同一学期内采集的图像，ROC2对应同一年内采集的图像，而ROC3对应不同学期采集的图像。

2.5.2 参数对结果的影响及分析

下面将在 FERET 人脸数据库中两个较大的测试集 (即 Fb 和 Dup.I) 上来测试相位特征的局部 Gabor 异或模式 (即 LGXP)、判别性特征提取方法 (即 BFLD) 及融合 Gabor 幅值和相位局部模式的方法在不同参数取值条件下的结果变化趋势，并给出相应的分析。根据该数据库提供的眼睛中心位置，所有图像的大小被归一化到 80×88 (宽度×高度) 个像素 (两眼中心分别设置为(21,25)和(59,25))。Gabor 小波的参数设置如下：5 个尺度 $\nu \in \{0, 1, \dots, 4\}$ ，8 个方向 $\mu \in \{0, 1, \dots, 7\}$ ， $\sigma = 2\pi$ ， $k_{\max} = \pi$ ， $f = \sqrt{2}$ 。此外，实验中没有采取任何光照预处理方法，并采用最近邻分类器进行分类。

1. 不同参数对局部 Gabor 异或模式方法 (即 LGXP) 的影响

LGXP 方法主要有三个参数：相位角的量化区间 b 、空间邻域内像素数目 P 和每幅模式图像划分的子块数目 m 。量化区间的大小确定了邻域内两个相位角在多大程度上相似才能被确定为同一种局部特征，正如第 2.3 节所分析的，它的值在增强模式对相位随位置变化的鲁棒性和保持不同粒度特征的区分性之间存在折中；空间邻域内像素数目 P 决定了局部模式的数目 (即 2^P)：当 P 取较大值时，LGXP 方法可以编码更多的局部模式，这也带来更高的特征维数；子块数目 m 影响了人脸不同区域之间的相对位置关系

以及对局部变化的鲁棒性,需要在二者之间进行折中。为了系统地测试参数的不同取值对结果的影响,下面将这三个参数的取值范围分别设置如下:量化区间 b 的范围是 $[2,10]$,空间邻域内像素数目 P 设置为 3 或 4 (模式数目分别是 8 种或 16 种),每幅模式图像划分的子块数目 m 的最大值设置为 10×11 (图像子块的最小尺寸是 8×8 个像素)。

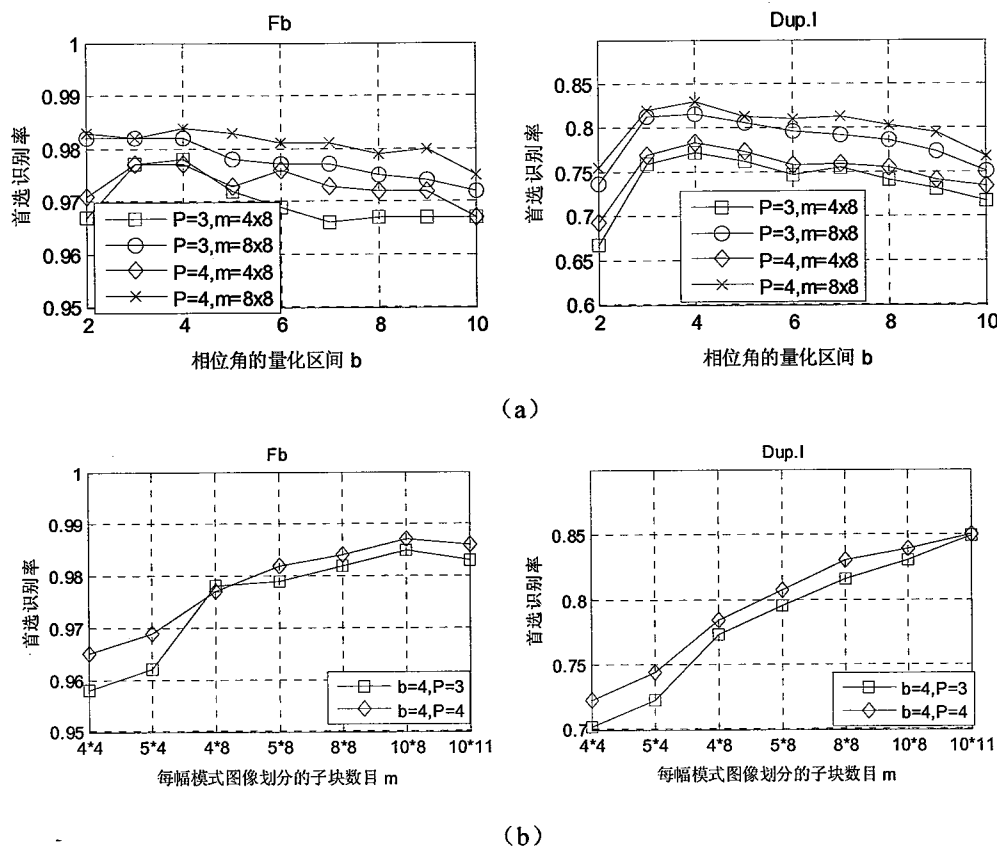


图2-11 不同参数对LGXP方法的影响 (a) 相位角的量化区间 (b) 每幅模式图像划分的子块数目

图 2-11 (a) 和 (b) 分别绘出了相位角的量化区间 b 和每幅模式图像划分的子块数目 m 变化时 LGXP 方法的结果变化曲线。图 (a) 中绘出了空间邻域内像素数目 P 取 3 和 4, 子块数目 m 取 4×8 (即 32) 和 8×8 (即 64) 时, LGXP 方法随量化区间 b 值变化的曲线。可以看到, 当量化区间 b 的值从 2 增加到 10 时, 识别结果呈现了先上升后下降的趋势。特别地, 当量化区间设置为 4 时, LGXP 方法在 Fb 和 Dup.I 两个子集上基本取得了最优的结果。这表明了 LGXP 表示在对相位随位置变化的鲁棒性以及区分不同粒度的特征之间取得了很好的均衡。图 (b) 中绘出了量化区间 b 取 4、空间邻域内像素数目 P 取 3 和 4 时, LGXP 方法随子块数目 m 的变化情况。随着 m 值的增加, Fb 子集上的识别结果先上升到一定数值后略微下降, 而 Dup.I 子集上的识别结果基本呈上升趋势。这与之前的分析基本上保持一致。考虑到计算代价以及识别性能, 在下面实验中, LGXP 方法的参数设置如下: 相位角的量化区间设置为 4, 每幅模式图像划分为 64 (即 8×8) 个子块, 空间邻域内像素数目 P 设置为 3 或 4。

2. 不同参数对判别性特征提取方法（即 BFLD）的影响

判别性特征提取方法主要有两个参数：每幅模式图像划分的块数目 M 和每个块划分的子块数目 K 。块数目 M 决定了原始特征划分的段数目，在降低每一段特征的维数与保持每一段特征的代表能力之间存在折中。类似地，每个块的子块数目 K 需要在保持同一个块不同子块间的相对位置关系与直方图对局部变化的鲁棒性之间进行均衡。此外，由于子块数目 K 与空间域内像素数目 P 共同决定了每个块的特征维数（即 $40 \times K \times 2^P$ ），局部 Gabor 模式中的模式数目不能取太大的数值。这里测试了 LGXP 表示与判别性特征提取相结合的方法（简记为“LGXP+BFLD”）在不同参数取值条件下的结果。这些参数的取值范围设置如下：块数目 M 的值设置为从 2×2 变化到 4×8 ，每个块的子块数目 K 的值设置为从 1×1 变化到 2×4 ，空间邻域内像素数目 P 的值设置为 3 或 4（模式数目分别是 8 种或 16 种），每个块的特征维数经 FLD 降到 200 维。

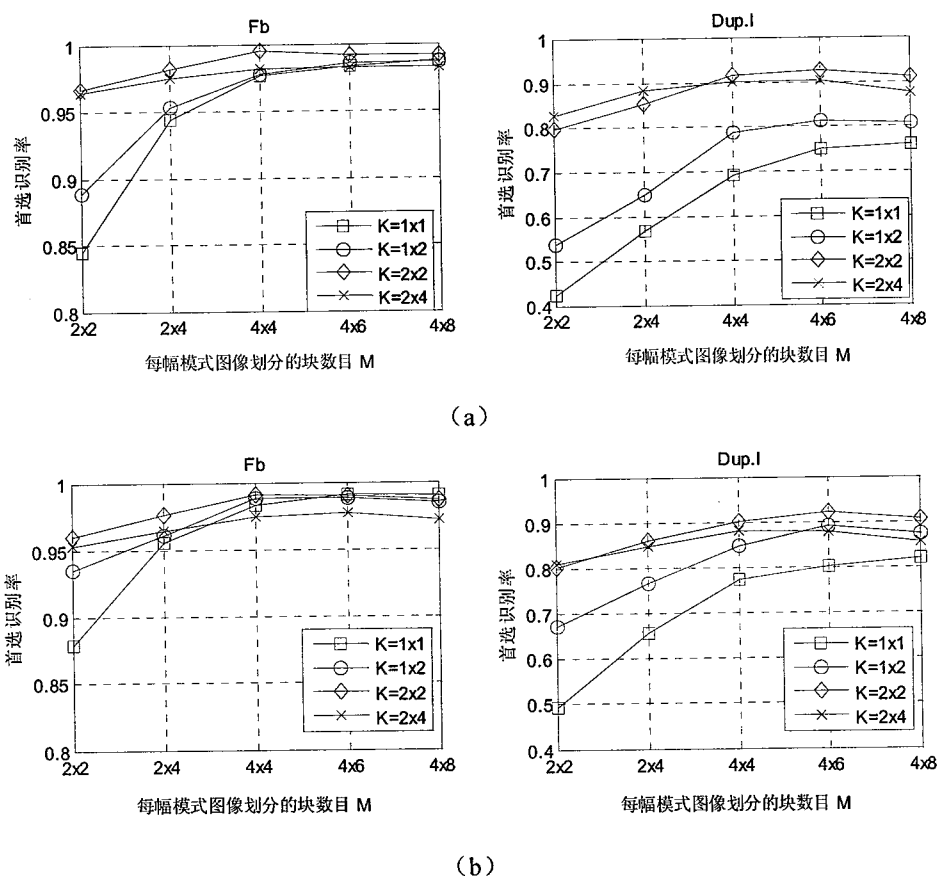


图 2-12 不同参数对“LGXP+BFLD”方法的影响 (a) $P=3$ (b) $P=4$

图 2-12 (a) 和 (b) 分别绘出了当空间邻域内像素数目 P 取 3 和 4 时，LGXP 表示结合判别性特征提取的方法在不同参数取值时的结果变化情况。可以看到，在 P 取 3 和 4 时，实验结果呈现了基本一致的变化趋势：当每个块的子块数目 K 比较小（如 1×1 ）时，随着块数目 M 的增加，结果基本呈上升趋势；当每个块的子块数目 K 比较大（如

2×4) 时, 实验结果则随着块数目 M 的增加呈现了先上升后下降的趋势。原因在于, 当块数目 M 比较小时, 每个块的尺寸比较大, 较少数目的子块导致了每个块的直方图表示会丢掉太多的空间信息, 而更多数目的子块则使每个块的直方图表示保留了更多不同区域间的相对位置关系, 因而具有更强的表示能力; 当块数目 M 比较大时, 每个块的尺寸变小, 过多的子块反而会使每个块的直方图表示对图像的局部变化更为敏感, 因而结果出现了下降现象。从图中也可以看到, 当块数目 M 取 16 (即 4×4), 子块数目 K 取 4 (即 2×2) 时, 该方法基本上取得了最好的结果。

3. 融合 Gabor 幅值和相位特征的方法与利用单一特征方法的结果对比

本小节在 Fb 和 Dup.I 子集上测试了融合 Gabor 幅值和相位特征的方法与利用单一特征的方法。相关参数设置如下: 每幅模式图像的块数目设置为 16 (即 4×4), 每个块的子块数目设置为 4 (即 2×2), 模式数目设置为 8 (即 $P=3$) 种和 16 (即 $P=4$) 种。为了区分不同类别的融合方法, 下面采用 $S[\cdot]$ 和 $F[\cdot]$ 分别表示不同类型特征的相似度层 (Score-level) 与特征层 (Feature-level) 融合方法。图 2-13 中绘出了根据 Gabor 幅值局部模式特征计算的相似度所占的权重 w 从 0.1 变化到 1.0 时 (步长为 0.1) 不同方法的结果。

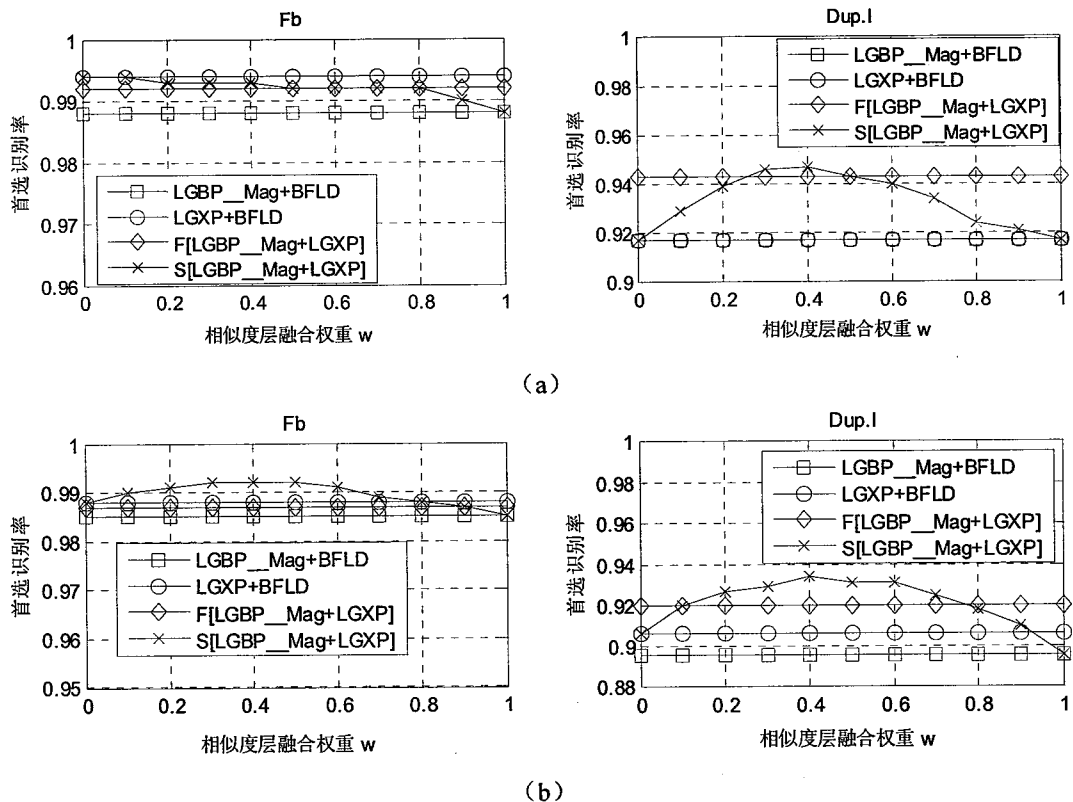


图 2-13 融合幅值和相位局部模式的方法与利用单一特征的方法在 FERET 测试集上的结果对比
(a) $P=3$ (b) $P=4$

从图 2-13 中可以看到, 融合 Gabor 幅值局部模式 (即 LGBP_Mag) 和相位局部模式 (即 LGXP) 的表示方法比只利用单一特征的方法在识别性能上有比较大的提高, 并

且这两种融合方法的结果基本相当。这验证了利用 Gabor 幅值和相位特征之间的互补性来提高分类性能的观点。此外,随着融合权重的增加,相似度层融合方法(即 $S[LGBP_Mag+LGXP]$)的结果呈现了先上升后下降的趋势。基本上,在融合权重取 0.4 或 0.5 时,该方法取得了最好的结果。这说明 Gabor 幅值和相位的局部模式表示在识别过程中具有同等重要的作用。

与只利用单一特征的方法相比,融合多种特征的方法在提高性能的同时,也产生了比较高的计算代价和存储代价:在特征层融合方法中,每个块的原始特征维数是幅值或相位局部模式特征维数的 2 倍,这导致学习 FLD 矩阵时有更高的计算代价;在相似度层融合方法中,FLD 矩阵的数目是只利用幅值或相位特征方法中 FLD 矩阵数目的 2 倍,这导致图像最终的特征维数要高于利用单种特征方法中的特征维数。

2.5.3 FERET 人脸数据库上的结果对比及分析

本节将在 FERET 人脸数据库上测试相位的局部 Gabor 异或模式(即 LGXP)、判别性特征提取方法(即 BFLD)及融合 Gabor 幅值和相位局部模式的方法,并与其他方法的结果进行对比。这里,图像尺寸以及 Gabor 小波参数等的设置与第 2.5.2 节中的设置完全一致。在局部 Gabor 模式表示中,图像划分为 64(即 8×8)个子块,模式数目设置为 16 种。在局部 Gabor 模式与判别性特征提取结合的方法中,每幅模式图像划分为 16(即 4×4)个块,每个块划分为 4(即 2×2)个子块,模式数目设置为 16 种。在相似度层融合方法中,融合权重设置为 0.5。由于相似度层融合方法与特征层融合方法可以取得相当的性能(参见第 2.5.2 节“融合 Gabor 幅值和相位特征的方法与利用单一特征方法的结果对比”),本节将只报告其他 Gabor 特征的相似度层融合方法的结果。

表 2-2 中列出了不同方法在 FERET 测试集上的结果。图 2-14 中绘出了 Gabor 幅值和相位局部模式分别结合判别性特征提取方法以及它们的融合方法在 FERET 人脸库四个子集上的累计识别率曲线(Cumulative Match Curves, CMCs)。表 2-2 中所有的方法都由作者实现,并在完全相同的测试条件下进行测试。从表 2-2 和图 2-14 中可以看到:第一,相位特征的局部 Gabor 异或模式(即 LGXP)方法比以前的局部 Gabor 模式方法表现更好。这验证了该表示方法在人脸识别上的有效性;第二,通过结合判别性特征提取方法(即 BFLD),所有局部 Gabor 模式表示方法的结果都有了很大的提高,特别是在 Dup.I 和 Dup.II 子集上;第三,融合不同 Gabor 特征的方法进一步提高了识别性能,并且,在对比的几种融合不同特征的方法中,融合幅值的局部 Gabor 二值模式(即 $LGBP_Mag$)和相位的局部 Gabor 异或模式(即 LGXP)表示的方法基本上取得了最好的结果。这些现象验证了基于特征融合的表示方法的有效性。特别地,如果可以合理地利用相位特征的特点,也可能设计出比幅值特征更有效的人脸表示方法。

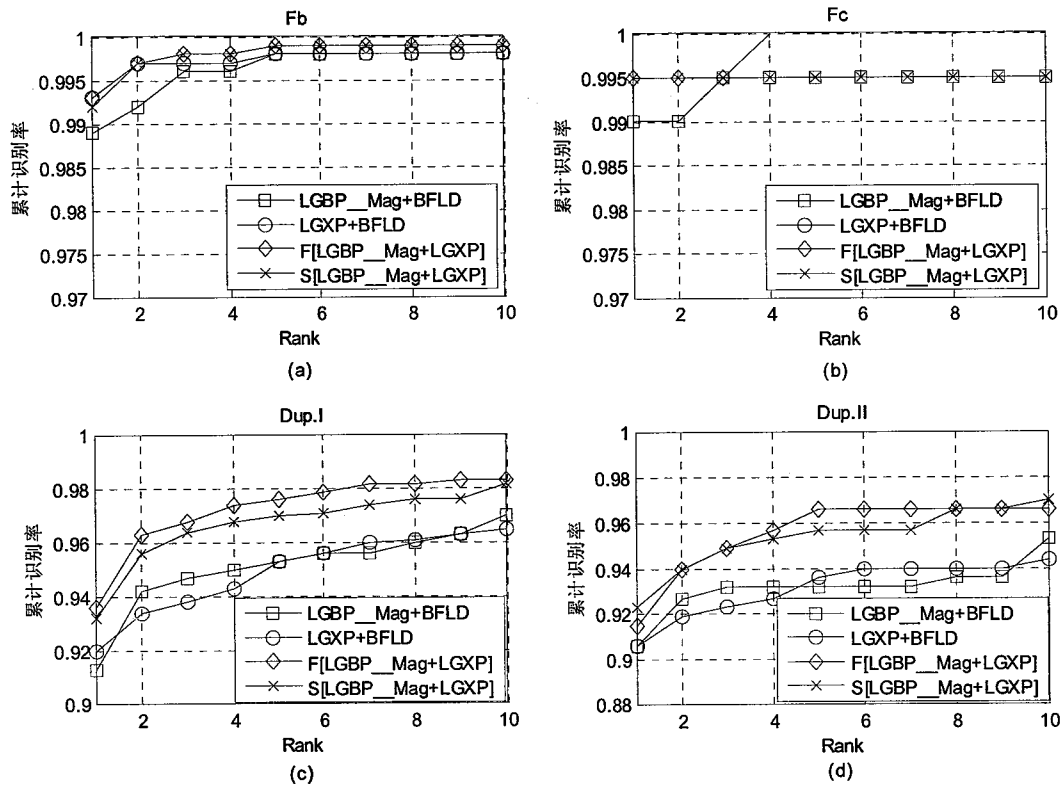


图 2-14 FERET 测试集上不同方法的累计识别率曲线

表 2-2 不同方法在 FERET 人脸库上的首选识别率对比 (%)

方法类别	方法	FERET 测试集			
		Fb	Fc	Dup.I	Dup.II
局部Gabor模式	LGBP_Mag	95.8	95.9	74.0	69.7
	LGBP_Pha	97.0	97.4	79.5	78.6
	Re_LGPP	98.3	99.0	80.1	77.8
	Im_LGPP	98.2	98.5	75.5	72.6
	LGXP	98.4	100	81.9	83.3
局部Gabor模式 与BFLD结合	LGBP_Mag+BFLD	98.9	99.5	91.3	90.6
	LGBP_Pha+BFLD	99.2	99.0	87.8	86.3
	Re_LGPP+BFLD	99.2	100.0	91.3	87.2
	Im_LGPP+BFLD	99.2	100.0	90.0	87.6
	LGXP+BFLD	99.3	100.0	92.0	90.6
不同Gabor特征 的融合	S[LGBP_Mag+LGBP_Pha]	99.2	99.5	92.4	91.9
	S[Re_LGPP+Im_LGPP]	99.2	100.0	91.7	88.9
	S[LGBP_Mag+Re_LGPP]	99.2	99.5	92.9	91.0
	F[LGBP_Mag+LGXP]	98.9	99.5	92.8	91.9
	S[LGBP_Mag+LGXP]	99.3	99.5	93.8	92.7

表 2-3 中列出了本章方法的最好结果与文献中近几年已发表的结果之间的对比。表

中其他方法的结果直接摘自相应的参考文献。文献[27]和[151]的方法均是融合 Gabor 幅值和相位特征的方法;文献[117]提出了利用 PCA 和 KDCV(Kernel Discriminant Common Vector)方法对 Gabor 特征和 LBP 特征进行融合;文献[159]对局部匹配方法的每一步进行了系统的对比,并将每一步的最优选择结合起来实现了一个完整的人脸识别系统。从表中可以看到, Gabor 幅值和相位局部模式的相似度层融合方法取得了更好的结果,特别是在 Dup.I 和 Dup.II 两个子集上。该方法的有效性主要归因于 Gabor 相位特征局部模式表示的有效性以及高效的判别性特征提取方法。

表 2-3 本章方法与其他方法在 FERET 人脸库上的结果对比 (%)

方法	FERET测试集			
	Fb	Fc	Dup.I	Dup.II
Gao等的方法[27]	98.0	83.0	70.0	56.0
Tan和Triggs的方法[117]	98.0	98.0	90.0	85.0
Zhang等的方法[151]	99.0	96.0	78.0	77.0
Zou等的方法[159]	99.5	99.5	85.0	79.5
本章的相似度层融合方法	99.3	99.5	93.8	92.7

2.5.4 FRGC 2.0 人脸数据库上的结果对比及分析

本节在 FRGC 2.0 人脸数据库的实验 1 和实验 4 上对不同的方法进行测试。为了利用人脸图像的高分辨率信息,所有人脸图像的大小根据提供的眼睛中心位置归一化到 128×168 个像素(两眼中心分别设置为(29, 65)和(100, 65))。在局部 Gabor 模式方法中,每幅图像划分为 64(即 8×8)个子块,模式数目设置为 16 种。在局部 Gabor 模式与判别性特征提取(即 BFLD)结合的方法中,每幅图像划分为 16(即 4×4)个块,每个块划分为 8(即 4×2)个子块,模式数目设置为 8 种,根据提供的训练集(222 个人的 12,776 幅图像)来学习多个块的 FLD 矩阵块并将每个块的维数经 FLD 后保留至 200 维。在相似度层融合方法中,权重设置为 0.5。在特征层融合方法中,为了避免每个块产生高维的特征,每个块的子块数目设置为 4(即 2×2)。

表 2-4 中列出了不同方法在 FRGC 2.0 人脸库上的结果对比,同时,图 2-15 绘出了 Gabor 幅值局部模式和相位局部模式分别结合判别性特征提取方法以及其特征层与相似度层融合方法的 ROC 曲线(对应 ROC3)。从这些结果中,可以看到一些有趣的观察:第一,相位特征的局部 Gabor 异或模式(即 LGXP)在相对简单的实验 1 中取得了最好的结果,但在具有挑战性的实验 4 中,它的性能低于相位特征的局部 Gabor 二值模式(LGBP_Pha)的结果。此外,表 2-4 中所有的局部 Gabor 模式方法在实验 4 中都取得了相对较低的结果(大多数低于 20%)。由于这些方法是直接基于单幅图像表示,本身并不具有“判别性”,因而很难处理实验 4 中由复杂光照和图像模糊等引起的变化;第二,通过结合判别性特征提取方法,表中所有的局部 Gabor 模式表示方法都取得了非常大的性能提升,特别是在实验 4 上(从 20%到 80%)。这主要归因于 BFLD 方法的判

别分析本质以及“分而治之”的策略。具体来讲,FLD 方法通过将同一类不同图像之间的外在变化建模为类内差异而将不同类图像之间的身份差异建模为类间差异,其学习的投影方向可以有效地区分不同类别。此外,作为一种“分而治之”的方法,BFLD 方法集成了多个块的决策,这对提高性能也有一定的作用;第三,融合不同 Gabor 特征的方法取得了更优的结果。其中,融合幅值和相位局部模式的方法(即 S[LGBP_Mag+LGXP])基本上取得了最好的结果。从表中可以注意到,并不是融合任意不同特征的方法都带来了性能的提升,如融合幅值和实部的局部模式的方法(即 S[LGBP_Mag+Re_LGPP]),这可能是由于二者的特征表示之间存在较大的相关性。

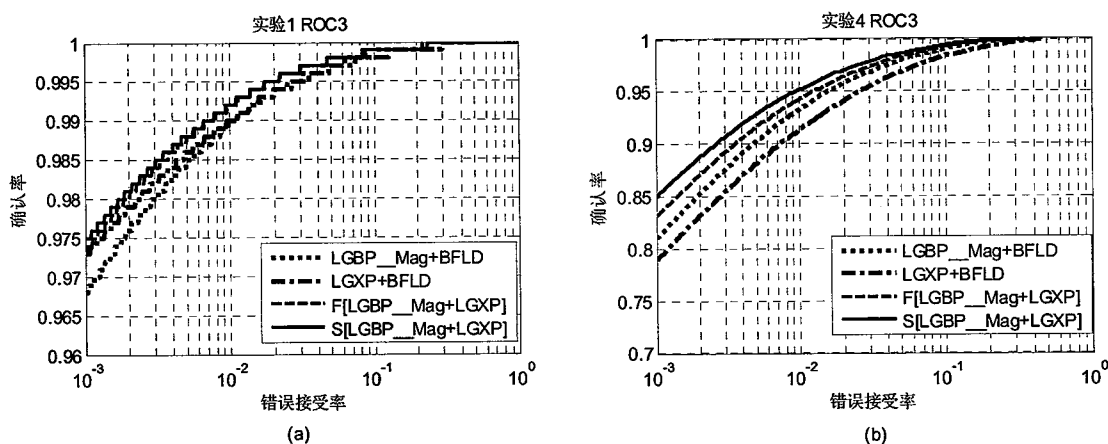


图 2-15 不同方法在 FRGC 2.0 人脸库上的 ROC 曲线 (a) 实验 1 (b) 实验 4

表 2-4 不同方法在 FRGC 2.0 人脸库上的确认率对比(错误接受率=0.1%)(%)

方法类别	方法	实验1			实验4		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
局部Gabor模式	LGBP_Mag	80.3	76.3	71.9	11.3	11.4	11.7
	LGBP_Pha	88.1	84.5	80.4	20.7	19.5	19.4
	Re_LGPP	90.0	86.4	83.0	18.3	17.9	17.4
	Im_LGPP	90.1	86.5	83.0	18.5	18.0	17.6
	LGXP	90.7	87.6	84.3	19.0	18.4	17.8
局部Gabor模式 与BFLD结合	LGBP_Mag+BFLD	98.1	97.5	96.8	79.6	80.3	81.0
	LGBP_Pha+BFLD	97.4	96.5	95.6	71.5	72.0	72.3
	Re_LGPP+BFLD	98.2	97.4	96.6	73.8	74.1	74.4
	Im_LGPP+BFLD	98.2	97.4	96.7	74.1	74.4	74.7
	LGXP+BFLD	98.5	97.9	97.3	78.2	78.6	78.9
不同Gabor特征 的融合	S[LGBP_Mag+LGBP_Pha]	98.3	97.6	97.0	81.5	82.3	83.1
	S[Re_LGPP+Im_LGPP]	98.4	97.7	97.0	77.2	77.7	78.0
	S[LGBP_Mag+Re_LGPP]	97.7	97.1	96.4	79.5	79.4	79.5
	F[LGBP_Mag+LGXP]	98.6	98.0	97.3	81.9	82.6	83.1
	S[LGBP_Mag+LGXP]	98.7	98.1	97.5	83.6	84.3	84.9

考虑到 FRGC 2.0 人脸库上的图像呈现了复杂的光照变化, 这里测试了两种光照预处理方法对实验结果的影响: 直方图均衡化 (Histogram Equalization, HE) 和 PP[116] 预处理。其中, PP 预处理方法在许多具有复杂光照变化的测试集上提高了一些人脸识别算法的性能[77][116]。图 2-16 (a) ~ (c) 分别示出了原始图像、经 HE 和 PP 预处理后的图像。相比原始图像, 直方图均衡化会使图像的整体亮度偏向均匀, 而 PP 预处理更倾向于增强结构性特征明显的区域 (如眼睛、鼻孔、嘴巴)。

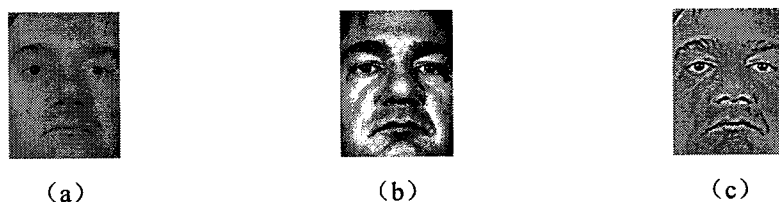


图2-16 不同光照预处理后的人脸图像示例 (a) 无预处理 (b) HE (c) PP

表 2-5 不同光照预处理方法在 FRGC 2.0 人脸库的确认率对比 (错误接受率=0.1%) (%)

方法	光照 预处理	实验1			实验4		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
LGBP_Mag+BFLD	无	98.1	97.5	96.8	79.6	80.3	81.0
	HE	98.2	97.6	97.0	73.4	74.3	74.8
	PP	98.0	97.4	96.6	80.1	80.7	81.3
LGXP+BFLD	无	98.5	97.9	97.3	78.2	78.6	78.9
	HE	98.7	98.0	97.3	70.3	70.5	70.7
	PP	98.4	97.6	96.9	78.0	78.4	78.6
F[LGBP_Mag+LGXP]	无	98.6	98.0	97.3	81.9	82.6	83.1
	HE	98.7	98.1	97.5	75.9	76.6	77.1
	PP	98.6	97.9	97.2	82.7	83.2	83.6
S[LGBP_Mag+LGXP]	无	98.7	98.1	97.5	83.6	84.3	84.9
	HE	98.8	98.3	97.8	78.7	79.4	80.1
	PP	98.6	98.0	97.3	83.9	84.7	85.2

表 2-5 中列出了几种结合判别性特征提取的方法在不同光照预处理方法之后的结果。可以看到, 对表中的四种识别方法, 并不存在一种光照预处理方法能够一致地提高它们的结果。具体来讲, 利用 HE 预处理, 表中所有的方法几乎在实验 1 上取得了比 PP 预处理更好的结果, 而在实验 4 上, 这些方法结合 HE 预处理的结果甚至低于无预处理的结果; 利用 PP 预处理, 除了“LGXP+BFLD”方法外, 表中的其他方法都在实验 4 上的结果要优于 HE 预处理的结果, 但与无预处理的同类方法相比, 结果改进不大。从这里可以看出, 这两种光照预处理方法对提高这些方法的结果并没有起到很大的作用。

对于判别性特征提取方法 (即 BFLD), 一个有趣的实验是测试不同训练集对其结果的影响, 即: 在 FERET 人脸数据库上测试时, 选择 FERET 训练集和 FRGC 2.0 训练

集分别来学习 FLD 投影矩阵, 然后测试其对识别结果的影响; 同样, 在 FRGC 2.0 人脸数据库上测试时, 也选择这两个训练集分别来学习 FLD 投影矩阵。为了简洁起见, 这里只报告了 FERET Dup.I 和 FRGC 2.0 实验 4 (ROC3) 上的结果, 如表 2-6 所示。可以看到, 该判别性特征提取方法受训练集的影响比较大, 特别在 FRGC 2.0 实验 4 上 (性能下降了接近 40%)。这说明该方法很容易“过学习”(Overfitting) 到训练集上: 当测试集与训练集是同质的数据 (如 FERET 测试集 vs FERET 训练集) 时, 该方法可以有效地对测试样本分类; 相反, 如果测试集与训练集是异质的数据 (如 FRGC 2.0 测试集 vs FERET 训练集), 该方法学习得到投影方向不能对测试样本进行准确地分类, 导致分类精度受到了比较大的影响。尽管如此, 这个问题已经超出了本论文的研究范围, 故不再作详细讨论。

表 2-6 不同训练集对 BFLD 方法结果的影响 (%)

BFLD 训练集 方法	FERET Dup.I (首选识别率)		FRGC 2.0 实验 4 (ROC3) (错误接受率=0.1%时的确认率)	
	FERET	FRGC 2.0	FERET	FRGC 2.0
LGBP_Mag+BFLD	91.3	78.7	42.3	81.0
LGXP+BFLD	92.0	83.7	35.4	78.9
F[LGBP_Mag+LGXP]	92.8	84.1	39.7	83.1
S[LGBP_Mag+LGXP]	93.6	84.9	48.7	84.9

表 2-7 本章方法与其他方法在 FRGC 2.0 人脸库上的确认率 (错误接受率=0.1%) (ROC3) 对比 (%)

方法	实验1	实验4
Hwang等的方法[44]	91.5	74.3
Liu的方法[74]	92.0	76.0
Liu和Liu的方法[77]	N/A	81.3
Su等的方法[109]	98.0	86.0
Tan和Triggs的方法[117]	N/A	83.6
本章的相似度融合方法	97.5	84.9

表 2-7 中列出了本章方法的最好结果与文献中近几年已发表结果之间的对比。其中, 表中其他方法的结果直接摘自相应的参考文献。除文献[74]中的方法是基于单种特征 (Gabor 特征) 的方法之外, 其他的方法都是融合了多种特征: 文献[44]中的方法利用了不同人脸模板以及不同频段的傅里叶特征; 文献[77]中的方法利用了彩色图像信息以及不同模板对应的 DCT 特征; 文献[109]中的方法利用了全局傅里叶特征与局部 Gabor 特征; 文献[117]中的方法利用了 Gabor 特征和 LBP 特征; 本章的方法则是融合了 Gabor 幅值和相位特征。可以看到, 融合多种特征的方法比利用单一特征方法在实验 4 上的分类精度要高很多。其主要的不足之处在于多种特征的融合带来了较高的计算代价。此外, 从表中的对比结果来看, 融合 Gabor 幅值和相位局部模式的方法基本上可以达到与 Su

等的方法[109]相当的精度，并且优于表中其他的方法。

2.6 本章小结

本章研究了融合 Gabor 幅值和相位局部模式的表示方法。该方法中首先提出了相位特征的局部 Gabor 异或模式表示，然后采用了基于块的 Fisher 线性判别分析方法对 Gabor 幅值和相位的局部模式特征进行融合。在 FERET 和 FRGC 2.0 这两个人脸数据库上的实验结果表明，相位特征的局部 Gabor 异或模式是一种有效的局部模式表示方法，可以达到与幅值特征的局部模式表示相当甚至更好的结果；融合幅值和相位特征的方法利用了有效的判别性特征提取方法来融合幅值和相位的局部模式表示，在这两个数据库上取得了非常优异的分类性能。

本章的研究结果表明，如果合理地考虑相位特征的特点，完全可能根据相位特征设计出与利用幅值特征的表示相当或更好的方法。此外，虽然大多数局部 Gabor 模式表示在相对简单的测试数据上取得了优异的结果（如 FERET 人脸库上 Fb, Fc, FRGC 2.0 人脸库的实验 1），但它们的性能在具有挑战性的测试数据上受到的影响比较大（如 FRGC 2.0 实验 4）。这主要是由于这些表示本身并不具有“判别性”，当测试集合上存在巨大的变化时，不同类别样本在这些表示上的可分性会变差。幸运的是，基于块的 Fisher 线性判别分析方法有效地提高了这些局部 Gabor 模式表示的分类性能。这主要归因于该方法判别分析的本质和“分而治之”的策略。此外，判别性特征提取方法提取了低维的特征，也极大地降低了图像相似度的计算代价。因此，局部 Gabor 模式表示与判别性特征提取结合的方法是更有效的人脸识别方法。

尽管融合 Gabor 幅值和相位特征的表示方法在人脸识别上取得了良好的识别性能，它还有一些改进的空间：第一，在融合各个分量分类器的时候，只采用了简单的和规则。考虑到人脸不同区域在识别过程中的作用并不相同，加权融合应当是一种更为有效的策略；第二，FLD 方法作为一种典型的统计学习方法，当测试集合与训练集合呈现不同的变化趋势时（如 FERET 训练集 vs FRGC 2.0 测试集），其性能会下降很大（即“过拟合”问题）。但如何提高 FLD 方法的推广能力，这也是一个值得深入研究的课题。

第三章 Gabor 特征“体”局部模式的表示方法

3.1 引言

幅值特征的局部 Gabor 二值模式（即 LGBP_Mag）表示作为 Gabor 特征和 LBP 算子的一种结合，在人脸识别上取得了很好的效果[150]。该方法首先对人脸图像进行 Gabor 小波变换，然后利用 LBP 算子对图像的 Gabor 幅值特征计算得到多尺度、多方向的模式图并将这些模式图划分为多个无交叠的子块，最后将所有尺度、所有方向上各个子块的直方图串接为一个高维向量来表示人脸图像。该表示方法的有效性主要来自于两个方面：第一，多尺度、多方向的 Gabor 小波滤波抑制了噪声，增强了人脸图像上的显著性特征；第二，图像分块的直方图特征是对人脸不同区域内各种模式的频数统计，对表情、部分遮挡以及光照等引起的局部变化鲁棒。此外，不同于传统的 Eigenfaces[119] 和 Fisherfaces[4]方法，该方法不需要离线地学习统计模型，应用更简单。

然而，幅值特征的局部 Gabor 二值模式只刻画了图像块内的局部模式。对于图像的 Gabor 幅值特征来讲，它随图像位置的变化比较缓慢，在物体的边缘附近有很强的响应[57]。因此，对于同一点不同 Gabor 幅值响应构成的 Gabor Jet，它们之间的大小序关系反映了该点处的局部结构与不同 Gabor 小波核函数的匹配程度。因而，这种序关系构成的模式是对图像上微结构的一种刻画，能够用于人脸的表示。由于 LBP 算子是一种基于像素间亮度值大小关系的算子，因而能够用于刻画幅值特征构成的 Jet 上的模式。

为了更有效地建模 Gabor 幅值特征的局部模式，本章将图像的 Gabor 幅值特征组织为一个三维“体”，然后利用 LBP 算子来建模“体”上的模式，并将该特征表示应用于人脸识别。与第二章的局部 Gabor 异或模式不同，本章的局部模式表示将模式的定义从 Gabor 特征块扩展到“体”上。本章也研究了其他 Gabor 特征（如实部、虚部和相位）的局部模式扩展到相应特征“体”上的表示，并提出了 Gabor 特征“体”局部模式表示的两种扩展方法。

本章内容安排如下：第 3.2 节主要介绍了基于“体”的局部 Gabor 模式表示；第 3.3 节从两个方面对 Gabor 特征“体”的局部模式表示方法进行扩展：利用 Fisher 准则的子块权重学习以及判别性特征提取；第 3.4 节是实验部分，分别测试了不同模式定义对结果的影响、不同 Gabor 特征“体”的局部模式表示方法的结果以及两种扩展方法的结果，并给出了相应的分析；第 3.5 节总结了本章的工作。

3.2 基于“体”的局部 Gabor 模式表示

为了高效地建模人脸图像 Gabor 小波表示的局部模式,本部分首先将 Gabor 特征组织为一个三维“体”的结构:该特征“体”的前两维是图像平面,第三维对应了不同尺度、不同方向的 Gabor 小波(参见图 3-1);然后,通过局部算子来计算该特征“体”上的模式,以获得更有效的表示。下面首先介绍了基于“体”的局部 Gabor 二值模式(Volume-based Local Gabor Binary Patterns, V-LGBP)表示并分析其有效性,然后讨论了其他 Gabor 特征的局部模式扩展到相应特征“体”上的方法。

图 3-1 示出了基于“体”的局部 Gabor 二值模式表示中局部模式的计算流程。总体上讲,其模式包括两个部分:图像块内的局部模式和 Jet 上的局部模式。其中,块内的模式由邻域像素与中心像素根据 LBP 算子计算得到,即幅值的局部 Gabor 二值模式(LGBP_Mag);Jet 上的模式根据同一点不同滤波器的响应利用 LBP 算子计算得到,称为“JetLBP”。

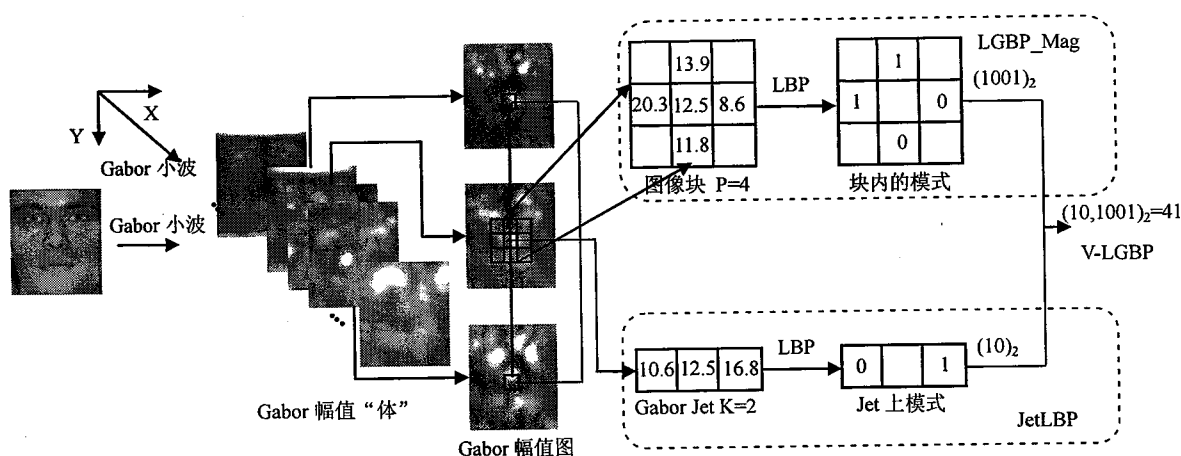


图 3-1 V-LGBP 方法中局部模式的计算流程

在 V-LGBP 方法中,其模式的二进制和十进制形式定义如下:

$$\begin{aligned} \text{V-LGBP}_{\mu,\nu}(z_c) &= [\text{JetLBP}_{\mu,\nu}^L, \dots, \text{JetLBP}_{\mu,\nu}^1, \text{LGBP_Mag}_{\mu,\nu}^P, \dots, \text{LGBP_Mag}_{\mu,\nu}^1]_{\text{binary}} \\ &= \left[\sum_{i=1}^P 2^{i-1} \cdot \text{LGBP_Mag}_{\mu,\nu}^i \right]_{\text{decimal}} + \left[\sum_{i=1}^L 2^{P+i-1} \cdot \text{JetLBP}_{\mu,\nu}^i \right]_{\text{decimal}} \end{aligned} \quad (3-1)$$

其中, $\text{V-LGBP}_{\mu,\nu}(z_c)$ 表示像素 z_c 在 Gabor 特征“体”内的模式; P 和 L 分别表示 3×3 邻域内的像素数目和 Jet 上的邻域大小,这二者共同决定了“体”上的模式数目(即 2^{P+L}); $\text{LGBP_Mag}_{\mu,\nu}^i$ ($1 \leq i \leq P$) 和 $\text{JetLBP}_{\mu,\nu}^i$ ($1 \leq i \leq L$) 分别表示图像块内的模式和 Jet 上的模式,如式 (3-2) 和式 (3-3) 所示:

$$\text{LGBP_Mag}_{\mu,\nu}^i = \begin{cases} 1, & \text{if } A_{\mu,\nu}(z_c) \geq A_{\mu,\nu}(z_i) \\ 0, & \text{otherwise} \end{cases} \quad (3-2)$$