

摘要

人脸识别是模式识别、图像处理和人工智能领域的一个重要研究课题，它具有广泛的应用背景并日益受到学术界、企业界、政府和军事部分的高度重视。人脸识别一般分为三个部分：预处理、人脸特征提取和分类器设计。人脸特征提取是在低维特征空间内对原高维空间的人脸模式进行描述。特征提取是人脸识别的关键环节，有效的人脸特征提取方法不仅能够设计出性能比较好的分类器，而且能够提高识别率。但由于人脸的复杂性和多变性，在视角、光照和表情等条件变化下人脸图像的类内散度远大于类间散度，导致很难有效地提取分类特征。因此，人脸特征提取也是人脸识别的主要难点所在。

本文对传统 LDB 方法进行改进，提出了一种改进的基于特定小波包分解系数矩特性的人脸识别方法，本文的主要工作有以下几个方面：

1. 对可分度定义进行了改进。传统的 LDB 对人脸图像小波包分解后选择部分系数作为特征矢量，然而这一准则的计算不满足可分度的线性可加性。本文针对该问题，给出了用系数对应的平均类间距离与平均类内距离的差作为特征系数的差可分度定义，以确定小波包分解子带。

2. 通过子带选择改进识别效果。并不是所有的小波子带和所有的小波分解系数都有助于分类识别，因此只考虑选定子带上的特征系数的一、二阶原点矩作为人脸特征。

3. 在分类识别方法上，本文提出了一种加权子带分类特征向量距离分类法，用来取代 LDB 中常用的最近邻分类识别法。

4. 将本文提出的方法应用到数字媒体信息处理系统研究项目中，设计并实现了人脸检测与识别模块，并通过集成实现了项目所需要的功能。

人脸识别研究的目标主要有两个，一是提高识别正确率，二是降低训练与识别时间。在此基础上，本文提出的这种新的人脸识别方法，既减少了计算复杂度，降低训练与识别时间，保证实时性，又能够更好的描述对分类有用的人脸特征，提高识别正确率。实验结果显示，我们的算法具有更好的识别效果，而且训练和识别的时间比原始 LDB 方法分别减少了两个数量级。

关键词：人脸识别，特征提取，小波包，矩特性，LDB

ABSTRACT

Face recognition is an important embranchment problem in the fields of pattern recognition, image processing and artificial intelligence. It has been widely applied and highly valued by academic field, enterprise field, government and the military day by day. Generally, face recognition contains three steps: face image preprocessing, feature extraction and classifier design. Face feature extraction is a key step for face recognition. The goal is to represent high dimension face patterns in low dimension feature space so as to extract face feature with discriminant power for classification. Effective face representation methods will not only to design a better classifier, but also to improve recognition rate. However, since face patterns are complicated and multiform, the within-class scatters of face images are larger than the between-class scatters under various conditions, such as changing illumination, pose and facial expression, resulting that it is very difficult to represent face effectively. Therefore, face representation is also as the primary difficulty in face recognition.

In this paper, the traditional LDB method is improved to be a new face recognition method, which is improved LDB method of face recognition based on moment property. The major contributions of this thesis are summarized as follows.

1. Improve the definition of separability. The traditional LDB chooses the parts of coefficients after the face image decomposition with Wavelet packet as eigenvector, but this criterion computation does not satisfy with the separability's linear additive property, this paper improves the definition of separability, it is called difference separability, which is used the average between-class distance difference with the coefficients correspondence average within-class distance, as the characteristic coefficient to certain the subbands of Wavelet packet decomposition.

2. Improve the recognition effect by choosing the subbands. Because of not all subbands of wavelet and all decomposition coefficients are helpful for recognition, only the first and second origin moments of chosen coefficients in chosen subbands are used as the feature of faces;

3. Improve the classification method. In the classification recognition method, this paper proposes a new classification method with the distance of weighted subband classification eigenvector, to substitute the nearest neighbor classification in the LDB.

4. Using the method proposed in the ethesis, the module of face detection and recognition are designed and carried out which is an important composing of the digital media information processing system.

There are two targets in face recognition: enhance recognition correctness and lower training and recognition time. In this paper, the traditional LDB method is improved to be a new face recognition method. This new method reduces the

calculation complexity, lowers the training and recognition time and achieves real time recognition, while the useful face features are better described for classification and recognition correctness are enhanced. The experiments show that the proposed method in this paper has better recognition effect, the training and the recognition time separately reduced two magnitudes compared to the LDB method.

Key words: face recognition, feature extraction, wavelet packet, moment property, LDB

表 目 录

表 1	人脸识别实验方案设计因子、水平表.....	43
表 2	各种算法实验结果.....	46
表 3	各种算法实验相对结果.....	47

图目录

图 1 Mallat 分解示意图.....	18
图 2 Mallat 重构示意图.....	18
图 3 图像的 2 层小波分解示意图.....	19
图 4 Lena 原图.....	20
图 5 Lena 的 1 层小波分解.....	20
图 6 Lena 的 2 层小波分解.....	20
图 7 由小波包组成的二元树.....	23
图 8 计算出所有子空间的可分性.....	35
图 9 把最下层的小波包基标上*号.....	35
图 10 确定最优基.....	35
图 11 人脸图像 2 层小波包分解示意图.....	40
图 12 选取人脸区域示意图.....	40
图 13 ORL 图像库人物示意图.....	44
图 14 CPOEP 图像库.....	44
图 15 视频中截取的人物图像.....	45
图 16 人脸识别实验流程图.....	45
图 17 识别率图.....	48
图 18 误识率图.....	48
图 19 训练时间图(单位:毫秒/M).....	49
图 20 部分训练时间图(单位:毫秒/M).....	50
图 21 测试时间图(单位:毫秒/M).....	50
图 22 支持基于内容检索的媒体特征分析平台体系结构图.....	52
图 23 人脸识别处理系统截图.....	53

独创性声明

本人声明所呈交的学位论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，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和撰写过的研究成果，也不包含为获得国防科学技术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。

学位论文题目：基于小波包矩特征的改进 LDB 人脸识别方法及其应用研究

学位论文作者签名：周蕾 日期：2007年12月1日

学位论文版权使用授权书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；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。

(保密学位论文在解密后适用本授权书。)

学位论文题目：基于小波包矩特征的改进 LDB 人脸识别方法及其应用研究

学位论文作者签名：周蕾 日期：2007年12月1日

作者指导教师签名：吴玲达 日期：2007年12月1日

第一章 绪 论

§ 1.1 研究背景

在过去的几十年间,人脸识别技术作为图像分析、理解研究的成功应用之一,正逐渐引起计算机视觉领域科研人员和信息安全产业界人士的极大关注。由于问题本身的特殊性和复杂性,人脸识别已不再局限于计算机科学研究单一范畴,越来越多的心理学家和神经科学家也对该技术表现出了浓厚的兴趣。人脸识别作为一个跨学科的新兴研究领域,为模式识别、图像处理、计算机视觉、人工智能、机器学习、计算机图形学、生理学、心理学等多种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,尝试新方法,验证新理论,解释新现象。因此,人脸识别问题的深入研究和最终解决,可以极大的促进这些学科的发展和成熟。

人脸是一种极为复杂的、多维的模式,也是一种典型的非刚性模式。人的面部特征十分丰富,除了形状、表情之外,还有五官的特征及分布。通过对这些特征的研究,可以理解人的情绪等状况。与其他利用视网膜识别(无法确保对人体的安全性)及指纹识别(涉及个人隐私)等人体生物特征进行身份验证相比,人脸识别的应用具有直观、友好及方便等特点,正越来越受到国际学术界、企业界、政府及国防军事部门的高度关注,具有广泛的应用前景^[1]。

人脸因人而异,绝对不会有完全相同的两张脸出现。虽然人类的人脸识别能力很强,能够记住上千个不同人脸,并且可以在表情、年龄^[2]或发型等发生巨大变化的情况下,可以毫不困难地由人脸而辨别出某一个人,可是这对于计算机而言,要建立一个能够完全自动进行人脸识别的系统却是非常困难的。主要原因包括人脸表情丰富;人脸随年龄增长而变化;人脸受光照,成像角度及成像距离的影响;从二维图形重建三维图像是一个艰难的过程,目前还没有很好的描述人脸的三维模型等。另外,人脸识别还牵涉到模式识别、图像处理、统计分类、神经网络、生理学、心理学以及认知科学等方面的诸多知识,这诸多因素使人脸识别成为一项极富挑战性的课题。与指纹、视网膜、虹膜、基因等其他人体生物特征识别系统相比,人脸识别系统更加直接、友好、易于为用户接受,并且通过人脸的表情/姿态分析,还能获得其他识别系统难以得到的一些信息。

人脸识别是人类视觉最杰出的能力之一,人脸所包含的视觉信息占据了主导地位,它是区别人与人之间差别的最重要的特征,因而人脸识别技术成为当今研究的一个热点^[3]。尽管对人类自己毫不费力,但实际上,由于人脸之间存在很大的相似性以及人脸的高度可变性,所以它的内在的机制是非常复杂的。因此人脸识

别不但具有非常重要的意义，还非常具有挑战性。

人脸识别技术在国家安全、军事安全和公共安全，智能门禁、视频监控、身份验证等领域得到了广泛的应用^[4]；在民事和经济领域，各类银行卡、信用卡、储蓄卡的持卡人的身份验证等具有重要的应用价值。

1.1.1 公共场所安全监控

美国 9.11 事件之后，世界各国普遍意识到公共场所存在的安全隐患，这促使各国对机场、码头、娱乐场所等人流较为密集的场所的安全问题更加关注。一些国家已经开始在机场安装人脸识别系统，这些人脸识别系统能够在很短的时间内拍下路过乘客的面部图像，分析处理后，将图像信息与恐怖分子和在逃犯人的面部图像数据库进行比较，进行人物识别。

1.1.2 智能监控

人脸识别作为一种智能技术，可以满足政府部门、住宅小区和企事业单位出入的自动化管理的要求，通过对各种门的出入控制，可以方便内部人员出入，杜绝陌生人的随意进出，增强了安全防范水平。

1.1.3 银行和金融系统

在一些国家，每年约有上亿美元的福利款被人以假冒的身份领取；每年利用丢失和被盗的信用卡进行犯罪的金额也有上亿的美元；由于使用盗窃来的身份识别码，造成移动电话通讯的损失高达十亿美元。利用有效的方法进行人物的身份识别可大大减少损失的金额。据估计，利用可靠的方法鉴别自动提款机持卡人的身份可以使每年由自动提款机诈骗造成的损失减少数十亿美元。因此，将人脸识别技术应用于柜台提款、自动提款机、远程交易身份确认、各种社会福利受益人的身份确认将起到不可忽视的安全保障作用。

§ 1.2 人脸识别研究的发展

人脸识别技术开始于本世纪六十年代，但由于受技术条件的限制，发展缓慢，因而在最初的二三十年里关于这个课题只有少量的论文出现^[5]。八十年代开始，随着计算机技术以及图像处理和模式识别技术的发展，人脸识别技术受到很大重视并得到了进一步的发展。

1.2.1 人脸识别的研究历程

人脸识别的研究已经有很长的历史,最初的人脸识别研究可以追溯到 1872 年,但从 20 世纪中期开始才有了较系统的研究。人脸识别的输入图像通常有三种情况:正面、侧面和倾斜^[6]。由于人脸正面图像包含了人脸更明显的特征,所以对人脸正面模式的研究最多,它的发展可以分为三个阶段^[7]:

第一阶段是从二十世纪 50 到 60 年代,以 Bertillon, Allen 和 Parke 为代表,主要研究人脸识别所需要的面部特征。在 Bertillon 系统中,提供了一个较强的识别系统,他用一个简单的语句与数据库中某一张人脸相联系。为了提高脸部识别率,Allen 为待识别脸设计了一种有效逼真的摹写,Parke 则用计算机实现了这一思想,并且产生了较高质量的人脸灰度图模型。总的来说,这一阶段的工作都需要利用人的某些先验知识,摆脱不了人的干预,还不能完成自动人脸识别工作。

第二阶段是人机交互识别阶段,时间是二十世纪 70 年代,人脸识别研究刚刚起步^{[8][9]},经典的模式分类方法,即采用正面脸部图像或者侧面脸部图像中那些可以计算的,比较精确的属性来识别人脸。这一时期的代表性工作:Goldstion, Harmon 和 Lesk 用几何特征。Goldstion 等人用几何特征参数来表示人脸正面图像,采用 21 维特征矢量表示人脸面部特征,并设计了基于这一特征表示法的识别系统。Kara 和 Kobayashi 使用基于统计的识别方法,用欧氏距离表征人脸。T.Kanade(M.Nagao)设计了一个高速而且有一定知识引导的半自动回溯识别系统^[10],创造性地运用积分投影法从单幅图像上计算出一组人脸特征参数,再利用模式分类技术与标准人脸匹配。Kanade 的系统实现了快速、实时的处理,是一个很大的进步。相比之下,Baron^[11]所做的工作较少为人所知:他先将图像灰度归一化,再用掩模(眼、鼻、嘴及眉毛以下的整个脸部)表示人脸,然后分别计算 4 个掩模与数据库中每幅标准图像的相应掩模之间的互相关系数,以此作为判别依据。总的来说,这类方法需要利用操作员的某些先验知识,仍然摆脱不了人的干预。

第三阶段是真正的机器自动人脸识别阶段。近几十年来,随着高速度高性能计算机的出现,人脸识别方法有了重大突破,提出了多种机器全自动识别系统,进入了真正的机器自动识别阶段,人脸识别研究也吸引了更多科学家的注意,使发展更具有鲁棒性的人脸识别方法称为时代的必然。于是基于整体的识别方法应运而生,如特征脸方法^[12]和弹性图匹配方法^[13]。国外有很多大学在此方面取得了很大进展,他们研究涉及的领域很广,其中有从感知和心理学角度探索人类识别人脸的机理,如美国 Texas at Dallas 大学的 Abdi 和 Toole 小组,主要研究人类感知人脸的规律,如漫画效应,性别识别与人脸识别的关系,种族效应等;由 Stirling 大学的 Bruce 教授和 Glasgow 大学的 Burton 教授合作领导的小组,主要是研究人类大脑在人脸认知中的作用,并在此基础上建立了人脸认知的两个功能模型,他

们对熟悉和陌生人脸的识别规律以及图像序列的人脸识别规律也进行了研究。也有从视觉机理角度进行研究的,如英国 Aberdeen 大学的 Craw 小组,主要研究人脸视觉表征方法,他们对空间频率在人脸识别中的作用也进行了分析;荷兰 Groningen 大学的 Perkov 小组,主要研究人类视觉系统的神经生理学机理并在此基础上发展了并行模式识别方法。不仅如此,随着计算机技术以及图像处理和模式识别技术的发展,人脸识别技术已应用到商用领域,代表性的商用人脸识别软件有 Facelt, Viisage system, TrueFace 等。

1.2.2 人脸识别的研究现状

人脸识别的研究历史是比较悠久的, Galton 早在 1888 年和 1910 年就分别在《Nature》杂志发表了两篇关于利用人脸进行身份识别的文章,对人类自身的人脸识别能力进行了分析,但当时还不可能涉及到人脸的自动识别问题。最早的人脸自动识别(Automatic Face Recognition, AFR)的研究论文见于 1965 年 Chan & Bledsoe 在 Panoramic Research Inc.发表的技术报告^[14],到现在已有三十多年的历史。人脸识别试验所采用的人脸库通常不大,最常见的人脸库一般包括 100 幅左右的人脸图像,如 MIT 库、Yale 库、CMU 库、ORL 库。近年来,人脸识别研究得到了诸多研究人员的青睐,涌现出了诸多技术方法。尤其是 1990 年以来,人脸识别更得到了长足的发展,每年都有大量的学术论文发表。

我国人脸识别技术的研究起步较晚,但发展较快。国内关于人脸识别的研究始于二十世纪 80 年代,主要是在国际流行的方法基础上作了发展性工作^{[15][16][17][18][19]}。

中国科技大学杨光正^[20]等提出一种基于镶嵌图的人脸自动识别方法,采用基于知识的三级金字塔结构对人脸进行分割和定位,前两级建立在不同分辨率的镶嵌图基础上,用于对人脸进行基本定位,第三级用一种改进的边缘检测方法进一步检测眼睛和嘴。基于这些器官的匹配就可进行人脸识别。

清华大学张长水^[21]等对特征脸的方法做了进一步发展,提出来用类间散度矩阵作为产生矩阵,进一步降低了产生矩阵的维数,在保持识别率的情况下大大降低了运算量。

南京理工大学杨静宇^{[22][23]}等主要是采用奇异值分解方法进行人脸识别研究,如用 Daubechies 正交小波变换对人脸图像作预处理,得到它在不同频带上的 4 个子图像,对它们分别提取奇异值,然后用最近邻方法进行分类,同时设计一种适用于多分类结果融合的群体决策算法,并且对分类结果有选择的进行融合。他们还研究了基于 Fisher 最佳鉴别矢量的人脸识别方法,并对神经网络用于人脸识别也进行了研究。

与此同时,中科院自动化所、公安部等单位也进行了人脸识别方面的研究,

并取得了一定的成果。

人脸识别的大多数算法及理论主要集中解决两个方面的问题：一是如何提取特征以表示人脸；二是对未知人脸用已有的特征表示出来并进行识别分类。事实上，特征提取是许多方法及理论中研究的重中之重，因为对于某一特定的分类方法而言，特征提取几乎决定了最终的识别效果。总的来说，这些方法与理论的目的就是找出一种在较低维空间中具有较大判别能力的人脸特征表示技术。从目前来看，人脸识别技术中最为成功的都是将人脸直接看成一个二维（2-D）的整体图案来处理，从而避免了对于三维（3-D）模型中的其它物体边界的检测^[24]。

§ 1.3 人脸识别的主要方法

计算机人脸识别方法的研究主要有两大方向：一是基于人脸图像局部特征的识别方法；二是基于人脸图像整体特征的识别方法。

基于人脸图像局部特征的识别通常抽取人脸器官如眼睛，眉毛，鼻子和嘴等器官的位置，尺度以及彼此间的比率作为特征。进一步地可以用几何形状拟合人脸器官，从而以几何参数作为描述人脸的特征。由于此类方法通常要精确地抽取人脸器官位置，尺度，比率或几何参数作为描述人脸的特征，因此这类方法对人脸图像的表情变化比较敏感，同时，人脸器官分割的精确度也对人脸特征的提取有一定的影响。另外，该类方法并没有充分利用到人脸图像本身具有的灰度信息，这类方法已经不是人脸识别技术发展的主流方向。

基于人脸图像整体特征的人脸识别方法由于不需要提取人脸图像中器官的具体信息，而且充分利用到人脸图像本身具有的灰度信息，因此可获得更高的识别性能。基于人脸图像整体特征的人脸识别方法主要有特征脸法，最佳鉴别矢量集法，贝叶斯法，基于傅里叶变换特征法，弹性图匹配法，线性子空间法，可变形模型法和基于人工神经网络的方法等等。其中弹性图匹配法和傅里叶不变特征法侧重于表述人脸图像；最佳鉴别矢量集法，贝叶斯法，基于人工神经网络的方法侧重于分类；特征脸法和线性子空间法等侧重于人脸图像的重构。下面分别介绍几种主流的人脸识别方法。

1.3.1 基于几何特征的人脸识别方法

人脸由眼睛、鼻子、嘴巴、下巴等部位构成，正因为这些人脸面部具有代表性的部位形状、大小和结构上的各种差异才使得世界上每张人脸千差万别，因此对这些部件的形状和结构关系的集合描述，可以作为人脸识别的重要特征。

这种识别方法首先将人脸用一个几何特征矢量表示，用模式识别中的层次聚类思想设计分类器来对人脸进行识别。它要求该矢量具有一定独特性，可以反映

不同人面部特征的差别。由于该方法对脸部朝向的改变非常敏感，要求有一定的弹性，以消除时间跨度和光照的影响。识别工作的流程大体如下：首先检测出面部特征点，通过测量这些关键点之间的相对距离，得到描述每张脸的特征矢量，比如眼睛、鼻子和嘴的位置和宽度，眉毛的厚度和弯曲程度等，以及这些特征之间的关系。用这些特征来表示人脸。比较未知脸和图像库中已知人脸的这些特征矢量，从而决定最佳匹配。

常用的方法有灰度的水平或垂直投影、基于边缘的特征提取等。由于对人脸器官的关键点进行水平或垂直投影后对应波峰或波谷，因此可以用来定位脸部主要器官。这种方法比较简单，但精度较差。边缘检测是图像处理中的一个经典的方法，它是利用图像的灰度变化来反映图像的特征，分别检测每个像素的邻域并对灰度变化率进行量化，通常也包括方向的确定。根据检测的领域和权值系数的不同，可将边缘算子分为：梯度边缘算子、Laplace 边缘算子、Roberts 边缘算子、Sobel 边缘算子、Canny 边缘算子等。基于边缘的特征提取往往不是把提取出的边缘作为最终结果，而是在这个基础上进一步处理，如利用一些先验知识和规律进行边缘的连接、拟合等。边缘的连接一般要考虑物体的实际形状以及当前点的邻域情况，而曲线的拟合则可能需要用分段线性和高阶样条曲线来拟合这些点，通常采用抛物线、椭圆等来拟合物体。Hough 变换对直线和圆进行拟合时能达到较好的效果，在人脸识别中通常用它来对眼球进行检测，进而完成对眼睛的定位。

总体来说，人脸的几何特征属于直观性特征，简单却容易受到人脸表情、光照条件、噪音的因素的影响，如人脸表情变化会直接影响提取的人脸五官的轮廓、形态和位置的很大变化，因此虽然基于几何特征的方法内存要求少，识别速度要比基于模板的方法高，但稳定性不高，较难实现特征的准确提取。

1.3.2 基于模板匹配的人脸识别方法

基于模板匹配^[25]的方法是模式识别中最简单的一种模式分类方法。主要代表是 Harvard 大学 Smith-Kettlewell 眼睛研究中心的 Yuille, Cohn 和 Hallinen, 他们利用可变形模板提取面部特征。这些模板可以平移、旋转和变形以便能够最佳地表示它们在图像中的形状。他们对图像进行了预处理，定义了各种能量函数，并在实验中发现，模板的起始位置对于确定眼睛的准确位置至关重要。当模板从眉毛上开始时，算法不能够区分眼睛和眉毛。算法的另一个缺点是计算复杂。

人脸识别和它的 3-D 目标识别一样，当投影不同或形状发生改变时，如何保持视觉的拓扑结构是非常重要的。这一问题的一个有效的解决途径是动态连接结构法或 DLA(Dynamic Link Architecture)^[26]。在通常的执行中，DLA 可以视为一个网格，它的节点包含对局部空间频率的一种多分辨率描述。每一个节点包含多个基于改进的 Gabor 小波变换特征检测器。这些特征检测器用高的精度描述

局部灰度分布,用较低的精度描述更全局的灰度分布。网格的节点同弹性连接相连。这些连接将特征分组成更高级别的排列,这些排列对可视目标编码。弹性使得适应目标视角投影的扭曲和变化成为可能。一个新的人脸可以通过手工,用网格覆盖在脸上来注册。一个人可以存储多幅图片以适应不同的面部表情。识别时,将测试图像变换成矢量网格,也称为图像域。识别过程就是将图像域和存储的所有原型图进行匹配,目标是 minimized 单个节点对之间的价格函数。如果一个人的原型图的匹配程度远大于其它图的匹配程度,就认为已被识别。图的匹配分为两个阶段。开始,用不变形的网格来定位图像中的人脸。网格一旦覆盖人脸,其结构按照每个节点在它的邻域内使价格函数达到极小值的方式变形。最终的匹配价格是网格每个节点的匹配价格和网格变形量价格的加权合成。

1.3.3 基于特征脸的人脸识别方法

特征脸方法是从主成分分析(Principal Component Analysis)导出的一种人脸识别和描述技术。PCA 实质上是 K-L 展开的递推实现, K-L 变换是数字图像压缩中的一种最优正交变换^[27],它以样本的最优重建为目的,通过 K-L 变换,可以把图像在高维空间表示转换到低维空间表示,而由低维空间恢复的图像和原图像具有最小的均方误差,从而可以以图像在低维空间的变换系数作为人脸图像的描述特征。

K-L 变换用于人脸识别的前提是人脸图像处于低维空间,并且不同人脸是线性可分的。通常情况下, K-L 变换的变换矩阵由训练样本类间散度矩阵的特征矢量生成,由类间散度矩阵得到的特征矢量类似于人脸,故将其成为特征脸。将变换矩阵的特征矢量按特征值的大小进行排列,人脸图像排在前面的特征矢量上的投影具有较大的能量,称为主分量;在排在后面的特征矢量上的投影具有较小的能量,称为次分量。当舍弃部分次分量时,称为主成分分析法 (PCA)。

K-L 变换从压缩角度看是最优的,但从分类角度来看却不是最优的。虽然它考虑了人脸图像的所有差异(从压缩角度),但没有考虑这些差异是类内差异(如光照变化、表情变化或几何变化)还是类间差异(从分类角度)。

Sirovich 和 Kirby^[28]首先采用主成分分析法 (PCA) 将 K-L 变换用于人脸图像的最优表征。Turk 和 Pentland^[29]具体将 K-L 变换应用于人脸识别,提出了经典的“特征脸”(Eigenfaces)方法。该方法的主要思想是,对于一幅由 N 个像素组成的图像,可以看作是一个 N 维矢量,或是一个 N 维空间中的一点。Eigenfaces 方法假设人脸图像只占据这个高维图像空间的一个很小的子区域,因此可以利用 PCA 原理来得到一个人脸图像的优化坐标系统。即是对这个人脸子区域的坐标进行降维,使得每个人脸图像可以用很少几个参数来表示,这就大大地降低了人脸识别的计算复杂度。

传统的特征脸存在的种种不足之处，促使研究人员在特征脸方法基础上发展了许多改进方案，如将特征脸与线性判别函数^[30]相结合的方法，特征半脸方法^[31]等。

1.3.4 基于神经网络的人脸识别方法

基于神经网络^{[32][33]}的方法应用比较广泛，目前常用的人工神经网络方法是 BP 神经网络^[34]、自组织神经网络 (SOM)、径向基函数神经网络^[35]。人工神经网络 (Artificial Neural Network, ANN) 是由大量简单的处理单元 (神经元) 相互连接而构成的网络系统。人工神经网络目前正越来越被频繁地应用于人脸识别，人工神经网络方法的自学习功能非常强大，由于对人脸识别的许多规律和规则进行显式描述是非常困难的，神经网络则可以通过反复学习的过程获得对这些规律或规则的隐式表达，因此在提取特征后进行分类识别时具有其得天独厚的优势。利用神经网络进行人脸识别时，网络输入层的每一个节点对应样本的一个特征，而输出层的节点对人脸的类别进行编码。可以直接将人脸图像的灰度值输入到神经网络的输入层，但这样由于输入层节点数目远大于输出层节点数目，训练时网络的收敛速度慢，通常情况下，先对图像进行特征降维，再利用神经网络进行分类识别。Fleming 和 Cottrell^[36]提出了一种采用两个网络识别人脸的方法，第一个网络为特征提取网络，用于人脸图像的特征降维，这个网络采用自联想神经网络，网络的输出层和输入层有相同的节点数，训练时网络的输出和输入均为原始图像，其隐层节点的输出作为图像的降维特征，隐层节点数决定特征的维数；第二个网络为分类识别网络，将降维后的特征输入到分类识别网络，其输出决定人脸图像的类别。Shang-Hung Li 等人^[37]提出基于概率决策神经网络 (Probabilistic Decision Based) 的人脸识别算法，将基于概率决策的神经网络分别应用于人脸定位，特征提取和人脸识别，通过 FERET 和 ORL 人脸库的测试，算法具有良好的识别效果。Ricanek.K^[38]将基于 Hough 网络的最优能量方程应用于模式识别与人脸识别，算法具有较高的识别效率和良好的鲁棒性。CMU 的 Rowley 等^[39]使用了多个 ANN 检测多姿态的人脸：1 个位姿检测器用于估计输入窗口中人脸的位姿、3 个检测器分别检测正面、半侧面和侧面的人脸。使用经过对准和预处理的“人脸”样本以及采用“自举” (bootstrap) 方法收集分类器错分的样本作为“非人脸”样本训练各个 ANN，进一步修正分类器。基于神经网络的方法便于建模，鲁棒性好，但是计算量大，运算时间较长，并可能陷入局部最优，难于实现实时应用。

1.3.5 基于弹性图匹配的人脸识别方法

基于弹性匹配的人脸识别方法，其基本思想体现在所谓动态连接结构

(Dynamic Link Architecture, DLA)^[26]上。可以认为,弹性图匹配是一种考虑到识别目标局部特征点之间拓扑结构的、具有适应性的局部特征匹配方法。弹性匹配的理论基础是图匹配。弹性图匹配用图来描述人脸,图的顶点表示面部特殊的局部特征点,边则表示面部特征之间的拓扑连接关系。不仅可以利用整幅人脸图像的认识信息,而且还允许局部特征在一定程度上的变形,在很大程度上克服了人脸姿态变化,表情变化和发型变化对人脸识别的影响。

基于弹性图匹配的人脸识别方法采用属性拓扑图来表述人脸模式,通常情况下,属性拓扑图为二维稀疏网格。属性拓扑图上的每一个顶点均包含一特征矢量,它记录了人脸在该顶点位置的分布信息,可以采用各种描述局部信息的特征,如小波特征,形态特征和统计特征等。

从二维稀疏网格到实际人脸图像的映射数目相当庞大,因此应用弹性匹配进行人脸识别的时间也是非常长的。应用不同尺度和方向的小波可以提取属性拓扑图顶点处的人脸特征。较小尺度的小波用来描述局部细节,较大尺度的小波用来描述较大范围内的概貌,不同方向的小波对方向信息敏感。在弹性匹配中,应用小波变换抽取特征对于细微表情变化具有一定的不变性,同时保留了空间位置信息,在一定程度上容忍人脸从三维到二维投影引起的变化。

本文以上介绍的是人脸识别中几类最重要的算法,是当前人脸识别算法发展的主流方向。然而人脸识别问题是非常复杂的,单单凭借一种方法都难以达到理想的识别效果,因此,在实际的应用中,往往是把多个算法相结合,以发挥最大的功效,提高识别效率。

§ 1.4 人脸识别研究的难点

人脸识别是指将采集到的人脸图像的特征信息,与已知的人脸特征信息进行比较,从而判断待识别人脸是否为一个已知人脸的过程。人脸是具有复杂结构的三维非刚性生物体,人脸图像的模式受到多种因素的影响。人脸识别是一项极具挑战性的任务,影响识别效果的相关因素主要表现为以下几个方面:

1. 姿态:从三维人脸到二维人脸图像的投影成像过程中,当头部姿态发生变化时,产生了人脸图像的外观变化;

2. 光照:当光照的角度、强度和颜色发生变化时,人脸图像的模式也会随之变化,同一个人不同的光照条件下采集到的图像差别甚至大于不同的被测对象的图像之间的差别;

3. 表情:人的心理和面部表情会直接影响人的外观;

4. 遮挡:人脸在某些情况下会被遮挡,而使脸部特征缺损或丢失,会给识别带来严重的后果。例如眼镜、胡须、和其他一些饰物等都会直接影响到人脸的

成像效果。

另外，图像的采集环境、图像的分辨率，采集设备都影响着人脸图像模式。这些变化通常都具有非常大的不确定性，因此增加了人脸识别的难度。

§ 1.5 本文的主要工作

本文在参考了大量的人脸识别中外文献的基础上，介绍了当前国际上几种主流的人脸识别算法，并进一步研究了基于统计的人脸识别方法，对目前较为盛行的几种算法作了详尽的论述。然而，值得注意的是，在人脸自动识别系统研究中，这些方法在对信号图像高频部分进行处理时效果并不理想。小波变换被誉为“数学显微镜”，它是信号分析发展史上的里程碑式的进展。小波变换在信号分析、图像识别等方面具有重要的应用价值。小波变换虽然具有频率越高相应的时间（或空间）分辨率越高的优点，但其在频率域上的分辨率却相应降低。小波包在一定程度上弥补小波的这一缺陷。

1.5.1 本文的基本方法

本文借鉴了主成分分析（PCA）和线性判别分析（LDA）的算法思想，对传统的基于小波包的 LDB（Local Discriminant Basis）方法进行改进，提出了一种改进的基于特定小波包分解系数矩特性的人脸识别方法（LDB_DS_OM），综合 PCA、LDA 和 LDB 三种方法，从三个方面作了改进：

1. 以差可分度（Difference Separability, DS）代替 LDB 算法中的平均类间类内距离，用来选择小波包分解系数和最优基；
2. 以选定子带内选定系数的一、二阶原点矩（Origin Moments, OM）作为人脸特征；
3. 以 Bhattacharyya 距离分量作为子带分类特征向量距离度量，以各子带差可分度为权重对子带分类特征向量距离线性加权，以加权子带分类特征向量距离分类法取代 LDB 常用的最近邻分类识别法。

为了检验本文提出的改进算法的效果，我们分别选取了来自 Olivette 研究实验室的 ORL 人脸数据库、彩色的正面人物图像 CPOEP 和从新闻视频中截选的人物图像 PON。其中 ORL 人脸图像库包括从 1992 年 4 月到 1994 年 4 月拍摄的一系列人脸图像，由 40 个人的 400 幅灰度图像组成，每人 10 幅图像，图像尺寸为 92×112 ，图像背景为黑色。其中人脸面部表情和细节均有变化，例如笑与不笑，睁眼或闭眼，戴或不戴眼镜等，人脸姿态也有变化，其深度旋转和平面旋转可达 20 度，人脸的尺寸也有最多 10% 的变化。这是目前使用最广泛的人脸数据库。另外两类图像需要进行人脸探测后再进行人脸识别。实验结果表明，本文所提出的算法具有

很好的识别效果，与原始的 LDB 算法相比，训练时间和识别时间均减少了两个数量级。

1.5.2 本文结构

全文共分六章：

第一章 研究的背景，介绍了人脸识别的研究历史，以及目前的研究水平和取得的成果，并对几种重要的人脸识别方法进行了详细介绍。

第二章 介绍了小波与小波包的理论基础和一些相关的基本概念以及性质。本文主要是基于小波包进行的特征提取的研究，对特征提取的研究内容涉及到很多小波和小波包的基础知识，因此在本章对小波和小波包的内容作了详尽的阐述。

第三章 本文提出的方法借鉴了主成分分析法（PCA）和线性判别分析方法（LDA），这两种方法是当前比较流行的算法，本章从这两种方法的基本原理入手，具体的介绍了它们在人脸识别中的应用和基本原理。

第四章 在本章中首先对原始的人脸识别 LDB 方法作了论述，然后提出了一种新的特征提取的方法，对基于小波包特征提取的改进方法，同时也从对可分度定义和分类距离两个方面作了改进。

第五章 实验分析与应用研究，对人脸识别的各种方法作了大量的实验研究，对各种方法的实验结果进行对比，实验结果显示，本文提出的算法具有更好的识别效果，与原始 LDB 方法相比，训练和识别的时间也都大大的缩短了。介绍了人脸识别算法的实现和人脸探测与识别软件模块的构建方法，阐述了本文在数字媒体信息处理系统研究中的作用。

第六章 总结全文，指出本文在理论与实践上的贡献，并讨论研究还有待改进的地方和指出未来研究的方向。

第二章 小波与小波包分析的基本理论

在工程和数学上, Fourier 级数可以将一个信号表示成正弦函数和余弦函数的叠加, 它的物理意义是: 每个周期信号都是具有简单频率的简谐振动的叠加。Fourier 分析的诞生, 给工程界带来了一次深刻的革命, 同时它是一个强有力的数学工具, 用它来刻画函数空间, 求解微分方程, 进行数值计算及信号处理等等。200 年来, Fourier 分析具有十分旺盛的生命力。

不过, 虽然 Fourier 分析在信号分析中长期占据突出位置, 但它也具有不可忽视的缺点。这是因为正弦函数和余弦函数是永不衰减地分布在整个时间域上的, 因而 Fourier 系数是信号在整个时间域上的加数平均, 所以, 要反映信号的局部性是不可能的。

小波分析是传统 Fourier 分析发展史上的一个里程碑, 近年来成为众多学科共同关注的热点。一方面, 小波分析被看成是调和与分析这一数学领域半个世纪以来工作的结晶。另一方面, 它已经广泛应用于图像纹理分析、图像编码、布朗分形维数计算、语音识别等科技领域。在许多使用传统 Fourier 分析的地方都可以用小波分析取代。小波分析优于 Fourier 分析的地方是: 小波分析突破了傅氏分析在应用中的平稳性假设、周期延拓等限制, 具有时频定位的能力, 并在算法实施过程中实现了多分辨率分析的功能, 它在时域和频域同时具有良好局部化性质, 而且由于对高频成分采用逐渐精细的时域或空间域取样步长, 从而可以聚焦到对象的任意细节。小波分析的这一特点被誉为数学显微镜。这使得小波分析迅速成为数字信号处理的强大工具, 在图像处理、语音识别等领域取得令人瞩目的成绩, 成为国际研究的一个热点。

小波变换有连续和离散两种情形, 由于小波变换在信号处理中一般是在计算机上实现, 为了使得叙述更直观易懂, 本文将从多分辨率分析的角度简要讲述离散小波变换, 并给出离散小波变换的分解与重构算法(Mallat 快速算法), 以及小波包变换的分解与重构算法。

§ 2.1 小波变换概述

小波分析方法的提出, 可以追溯到 1910 年 Haar 提出的小波规则正交基及 1938 年 Littlewood-Paley 对傅立叶级数建立的 L-P 理论, 即按二进制频率成分分组傅立叶变换的相位变化在本质上不影响函数的形状和大小。其后, Caldern 于 1975 年用其早期发现的再生公式给出抛物线空间 H' 上的原子分解, 这个公式后来成了许多函数分解的出发点, 他的离散形式已接近小波展开, 只是还无法得到组成一正交

系的结论。1981年 Stromberg 对 Haar 系进行了改造,证明了小波函数的存在性。值得注意的是,1984年法国地球物理学家 Morlet 在分析地震的局部性质时,发现传统的变换难以达到要求,因此,他引入小波概念于信号分析中,对信号进行分解。随后,理论物理学家 Grossman 尝试对一个确定函数 $\Psi(x)$ 进行了伸缩、平移成为 $\left\{ |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) : a, b \in R, a \neq 0 \right\}$ 形式,他对 Morlet 的这种信号依此方法进行展开的可行性进行了研究,这无疑为小波分析的形成开创了先河。

小波变换实质上是对信号用一种多分辨率的带通滤波器进行滤波,将信号分解到不同的频带上再进行分析处理,小波变换的优点在于良好的时间和频率特性,应用范围很广。用小波来分解图像的原因是:

1. 用小波变换将图像分解之后,子带图像的分辨率降低了,相应的计算复杂度也大大降低了;
2. 小波分解提供了很好的空间和频率局部信息,并能逐步聚焦到分析对象的任何细节;
3. 人脸图像信息中低频部分保留了图像中的主要信息,即图像的整体,而图像的高频部分则保存着图像的细节信息,人脸的光照、遮掩、旋转扭曲和面部表情等因素只影响图像中的高频部分;
4. 小波的分解和重构都具有快速算法。

§ 2.2 Fourier 分析

Fourier 分析这门学科是数学分析中最古老的的学科之一,它对数学家和工程师都是相当重要的。Fourier 分析通常是指(积分)Fourier 变换和 Fourier 级数。Fourier 变换是在实直线 \mathbb{R} 上定义的某个函数 f 的 Fourier 积分。当 f 看作是一个模拟信号时,它的定义域 \mathbb{R} 就称为连续时域。在此情况下, f 的 Fourier 变换 \hat{f} 描述信号 f 的谱特性。因为谱信息用频率给出,所以 Fourier 变换 \hat{f} 的定义域还是 \mathbb{R} ,它称为频域。另一个方面,一个 Fourier 级数是双无限序列到周期函数的一种变换。因此,当一个双无限序列看作是一个数字信号时,它的定义域是整数集合 \mathbb{Z} ,称为离散时域。这时,它的 Fourier 级数再次描述数字信号的谱特性,一个 Fourier 级数的周期是 2π ,在此情况下,频域 \mathbb{R} 常用单位圆等同。

Fourier 变换与 Fourier 级数的重要性不仅由它们的物理解释的重要性,如信号的时间-频率分析,而且还由于 Fourier 分析技术是极其有力的。例如,在小波分析研究中, Poisson 求和公式、级数与积分的 Parseval 恒等式、Gaussian 的 Fourier 变换等等都是经常遇到的。

Fourier 分析理论提出了一种将能量有限的信号分解成一系列正弦波或余弦波的

信号之叠加的 Fourier 变换(简记为 FT)方法,其中能量有限是指 $\int_{\mathbb{R}} (f(t))^2 dt < +\infty$ 。

至 1965 年, Cooley 和 Tukey 提出了一种 Fourier 变换的快速算法(简记为 FFT),使得 Fourier 分析在工程技术领域得到广泛的应用。

一个能量有限的信号,可视为实数集 \mathbb{R} 上平方可积函数空间 $L^2(\mathbb{R})$ 中的一个函数 $f(t)$ 。对信号 $f(t) \in L^2(\mathbb{R})$, 其 Fourier 变换定义为

$$F(\omega) \triangleq \hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-i\omega t} dt \quad (2.1)$$

而 Fourier 逆变换定义为

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega)e^{i\omega t} d\omega \quad (2.2)$$

经典的傅氏变换在信号的分析与处理中发挥了重要的作用。然而,用(2.1)与(2.2)式对信号实施分解与重构,有如下不足之处:

- (1) 为了从 $f(t)$ 中提取频谱信息 $F(\omega)$, 要用到 $f(t)$ 的无限时区的信息;
- (2) (2.1)式没有反应出随时间变化的频率,也就是说 Fourier 变换不能较好的刻画信号在时域上的任何局部信息。

(3) 没有灵活可变的时间-频率窗,使在“中心频率”增大时自动变窄,而在“中心频率”降低时自动变宽,即 Fourier 分析无法作局部分析。

因此,傅氏变换对于大多数应用来说是不够的。

在实际应用中,人们需要确定时间间隔,使在任何希望的频率范围(或频带)上产生频谱信息。由于信号的频率与它的周期成正比,因此,对于高频谱的信息,时间间隔变小,从而给出较好的精度;对于低频谱的信息,时间间隔变大,从而给出完全的信息,也就是,需要一个可变的时间-频率窗,使得在高中心频率的时间窗自动变窄,而在低中心频率的时间窗自动变宽。小波变换具有这种移近和远离的伸缩能力^[40]。

§ 2.3 多分辨率分析

小波分析能够提供 $L^2(\mathbb{R})$ 中具有良好局部化性质的正交基,把 $L^2(\mathbb{R})$ 中的函数与 $L^2(\mathbb{Z})$ 中的数列等同起来,从而把分析问题转化为代数问题来解决。多分辨率分析的思想就是,先在能量有限函数空间 $L^2(\mathbb{R})$ 的某个子空间中先建立基底,然后利用简单的伸缩与平移变换,把子空间的基底扩充到 $L^2(\mathbb{R})$ 中。

定义 1^[42]

设 $\{V_j, j \in \mathbb{Z}\}$ (\mathbb{Z} 为整数集) 为 $L^2(\mathbb{R})$ 中的闭子空间列, 如果满足

$$(1) \quad \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots;$$

- (2) $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}), \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\};$
- (3) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, j \in \mathbb{Z};$
- (4) $f(x) \in V_j \Leftrightarrow f(x+n) \in V_j, n \in \mathbb{Z};$
- (5) 存在 $\varphi(x)$, 使得 $\{\varphi(x-k), k \in \mathbb{Z}\}$ 是 V_0 的一个 Riesz 基,

则称 $\{V_j, j \in \mathbb{Z}\}$ 是 $L^2(\mathbb{R})$ 的一个多分辨率分析, φ 被称为尺度函数。

定义 1 中的 $\{\varphi(x-k), k \in \mathbb{Z}\}$ 称为 V_0 的 Riesz 基是指: $\overline{\text{span}\{\varphi(x-k) | k \in \mathbb{Z}\}} = V_0$, 其中 $\overline{\text{span}\{\varphi(x-k) | k \in \mathbb{Z}\}} = V_0$ 表示由 $\{\varphi(x-k) | k \in \mathbb{Z}\}$ 所张成的闭线性空间的闭包, 且存在正常数 A 与 $B, 0 < A \leq B < +\infty$, 使得对于 l^2 中任意成员 $\{c_k\}$, 有

$$A \sum_{k \in \mathbb{Z}} |c_k|^2 \leq \left\| \sum_{k \in \mathbb{Z}} c_k \varphi(x-k) \right\|_2^2 \leq B \sum_{k \in \mathbb{Z}} |c_k|^2 \quad (2.3)$$

定义 1 中的性质 (1) 称为一致单调性, 表明子空间列 $\{V_j, j \in \mathbb{Z}\}$ 是嵌套的; 性质 (2) 称为渐近完全性, 表明通过增大 $j, L^2(\mathbb{R})$ 中的每一个函数 f 能够用它在 V_j 中的投影 $P_j f$ 非常地接近希望的逼近; 反之, 通过减小 j , 投影 $P_j f$ 能够具有任意小的能量; 性质 (3) 称为伸缩规则性, 表明空间列 $\{V_j, j \in \mathbb{Z}\}$ 中任一空间 V_j 的基可由其中另一个空间 V_j 的基经过伸缩变换得到; 性质 (4) 称为平移不变性, 表明空间列 $\{V_j, j \in \mathbb{Z}\}$ 中任一空间 V_j 的基可由其中另一个空间 V_j 的基经过平移变换得到; (5) 称为 Riesz 基的存在性。

由定义 1 知, 如果 $\{\varphi(x-n)\}$ 是 V_0 的一个 Riesz 基, 则 $\varphi(x-n), n \in \mathbb{Z}$ 是线性无关的, 从而 $\{\varphi(x-n)\}$ 是 V_0 的一个基。由定义 1 中的性质 (3) 知

$$\varphi_{k,n}(x) = 2^{\frac{k}{2}} \varphi(2^k x - n), n \in \mathbb{Z} \quad (2.4)$$

是 V_k 的基。方程(2.4) 称为尺度函数 φ 的双尺度方程。

利用多分辨率分析, 我们可以构造出更多的尺度函数 φ 和小波函数 $\psi(t)$, 其中由 Daubechies 在 1998 年成功构造出了具有紧支集的光滑小波函数, 这是一个小波发展史上的一个里程碑。

§ 2.4 小波与 Mallat 算法

1981 年, Morlet 在分析地质数据时, 基于群论, 首先提出了小波分析的概念, 并成功地应用于数值分析。1985 年左右, Y.Meyer 提出了光滑正交小波基。1989 年, S.Mallat 在文献^[41]中提出了多分辨率分析的概念, 给出了构造小波基的一般方法, 并提出了对信号进行小波分解与重构的快速算法——Mallat 算法。这一算法相

当于 Fourier 分析中的 FFT 算法, 使得小波分析从理论研究走入诸如数据压缩、时一频分析、流体力学、神经网络等许多工程技术领域。

2.4.1 小波分析

由文献^{[42][43]}知, 存在 $\varphi \in V_0$, 使得 $\{\varphi(x-n), n \in \mathbb{Z}\}$ 是 V_0 的规范正交基, 从而由 (2.4) 式确定的 $\{\varphi_{k,n}(x), n \in \mathbb{Z}\}$ 是 V_k 的规范正交基。由定义 1 的 (3) 及 (2.3) 式, 因 $\frac{1}{\sqrt{2}}\varphi(\frac{x}{2}) \in V_{-1} \subset V_0$, 故存在序列 $\{h_n\} \in l^2$, 使得下列称之为双尺度方程的式子成立,

$$\frac{1}{\sqrt{2}}\varphi(2^{-1}x) = \sum_{n \in \mathbb{Z}} h_n \varphi(x-n) \quad (2.5)$$

由 $\{\varphi(x-n), n \in \mathbb{Z}\}$ 的规范正交性得

$$h_n = \langle \frac{1}{\sqrt{2}}\varphi(2^{-1}x), \varphi(x-n) \rangle = \int_{\mathbb{R}} \frac{1}{\sqrt{2}}\varphi(2^{-1}x) \overline{\varphi(x-n)} dx \quad (2.6)$$

设 $\{V_k\}$ 为 $L^2(\mathbb{R})$ 的一个多分辨率分析, $\varphi(x)$ 是其尺度函数, 且使得 $\{\varphi(x-n)\}$ 是 V_0 的规范正交基, $\{h_n\} \in l^2$ 使得双尺度方程 (2.4) 成立, 令

$$g_n = (-1)^n \bar{h}_{1-n} \quad (2.7)$$

$$\psi(2^{-1}x) = 2^{\frac{1}{2}} \sum_{n \in \mathbb{Z}} g_n \varphi(x-n) \quad (2.8)$$

记 $\psi_{k,n}(x) = 2^{\frac{k}{2}}\psi(2^k x - n), k, n \in \mathbb{Z}$, 则 V_k 在 V_{k+1} 中的正交补

$$W_k = \text{Close}_{L^2(\mathbb{R})} \text{span}\{\psi_{k,n}(x) = 2^{k/2}\psi(2^k x - n), n \in \mathbb{Z}\}$$

且 $\{\psi_{k,n}(x), n \in \mathbb{Z}\}$ 是 W_k 的规范正交基。其中 $\text{Close}_{L^2(\mathbb{R})}$ 表示在 $L^2(\mathbb{R})$ 中取闭包, $\text{span}\{\psi_{k,n}(x), n \in \mathbb{Z}\}$ 表示由 $\{\psi_{k,n}(x), n \in \mathbb{Z}\}$ 张成的线性子空间。该定理的证明见文献^[44]。

定义 2 小波^[42]

定义在 $t \in \mathbb{R} := (-\infty, \infty)$ 实值函数 $\psi(t)$ 被称为小波, 当它满足下列条件:

- 均方可积: 即 $\psi(t) \in L^2(\mathbb{R})$;
- 正规化条件: $\int |\psi(t)|^2 dt = 1$ (or $\int |\Psi(\omega)|^2 d\omega = 2\pi$);
- 容许条件: $c_\psi := \int |\omega|^{-1} |\Psi(\omega)|^2 d\omega < \infty$ 。

其中 $\Psi(\omega) = \mathcal{F}\psi$ 是 $\psi(t)$ 的傅立叶变换:

$$\Psi(\omega) = \int \psi(t) e^{-i\omega t} dt \quad (2.9)$$

在信号处理中，一般采用离散小波变换（DWT）。

定义 3 离散小波变换 DWT^{[42][43]}

设 $Y(n)$ 为离散信号，则 $\hat{W}_l[b^l k], l = 0, 1, 2, \dots, k = 0, 1, 2, \dots$ 构成信号的 DWT。其中：

$$\hat{W}_l[k] = \sum_{n=0}^{N-1} Y[n] \psi_l(n-k) \quad (2.10)$$

在数字化实现中，离散小波变换是利用 Mallat 分解与重构算法来实现的。小波分解算法的优点不仅在于其具有尺度特性，而且它的计算复杂度为 $O(N)$ 对于数据长度 N 是线性的，从而可以实现对于高速计算机网络数据流的在线、实时处理。

2.4.2 Mallat 算法

多分辨率分析理论为人们讨论信号的局部信息提供了一个相当直观的框架。这一点在非平稳信号中的作用尤为重要。对于图像而言，任何一幅图像都可以分解为低频部分和高频部分。为了将信息的低频和高频部分分开处理，Mallat 系统提出了信号的塔式多分辨分解与重构的著名算法，俗称 Mallat 算法。

定义 4 Mallat 算法

由多分辨率分析的嵌套结构，可以得到一个重要结论，即小波分解系数 $c_{k,n}, d_{k,n}$ 可以由序列 $c_{k,n}$ 与离散时间滤波器系数 h, g 的离散时间卷积得到，即利用(2.7)式与(2.8)式，可得 $f_N(x)$ 的分解如下^[42]：

$$\begin{aligned} c_{k,n} &= \sum_{l \in \mathbb{Z}} c_{k+1,l} \overline{h_{l-2k}} \\ d_{k,n} &= \sum_{l \in \mathbb{Z}} c_{k+1,l} \overline{g_{l-2k}} \end{aligned} \quad (2.11)$$

重构算法：

$$c_{k+1,n} = \sum_{l \in \mathbb{Z}} c_{k,l} h_{k-2l} + \sum_{l \in \mathbb{Z}} d_{k,l} g_{k-2l} \quad (2.12)$$

其中， $\{h_n\} \in l^2(\mathbb{Z})$ 可看成低通滤波器系数， $\{g_n\} \in l^2(\mathbb{Z})$ 可看成是高通滤波器系数。 $c_{k,n}$ 是 $c_{k+1,l}$ 的近似信号， $d_{k,n}$ 是 $c_{k+1,l}$ 的细节信号。

根据 Mallat 算法，一维信号的小波分解与重构过程可由两组滤波器的级联滤波而产生，设 $\{c_{m,n}\}$ 为输入序列， $\{c_{m-i,n}\}$ 为经 i 次低通滤波而得到的输出， $\{d_{m-i,n}\}$ 为经 i 次高通滤波而得到的输出。称这种分解方式为塔式分解，其中，每次对低频分量(即近似信号)进行分解。上述递归分解过程可以得到离散小波变换的多层滤波金字塔算法-Mallat 算法(如图 1)，重构过程如图 2 所示。这一算法比传统的快速 Fourier 变换计算量更少。在上述算法中，滤波器系数 h_1, g_1 分别由小波函数与尺度

函数 ψ_0, φ_0 导出。通常情况下, 为进行分解, 我们需要给出过程的初始值 $c_{0,n}$, 它是过程 Y 在空间 V_0 上的分解系数。在实际计算过程中, 我们通常把 Y 的初始值 $Y[n]$ 作为小波分解初始值的近似。

事实上, Mallat 分解算法的结果相当于, 先把输入信号通过滤波器 $\{\overline{h}_n\}$ (或 $\{\overline{g}_n\}$), 再对滤波器输出进行抽样, 取其偶数部分。因此, 结果信号的分辨率是原信号的二分之一。

而 Mallat 重构算法的结果相当于, 先对近似信号 (或细节信号) 进行插值, 相邻两个数值之间插一个 0, 再通过滤波器 $\{h_n\}$ (或 $\{g_n\}$), 最后把两个滤波器的输出相加。

根据 Mallat 算法, 一维信号的小波分解与重构过程可由两组滤波器的级联滤波而产生, 设 $\{c_{m,n}\}$ 为输入序列, $\{c_{m-i,n}\}$ 为经 i 次低通滤波而得到的输出, $\{d_{m-i,n}\}$ 为经 i 次高通滤波而得到的输出, 信号的分解与重构过程如图 2 所示。称这种分解方式为塔式分解, 其中, 每次对低频分量 (即近似信号) 进行分解。

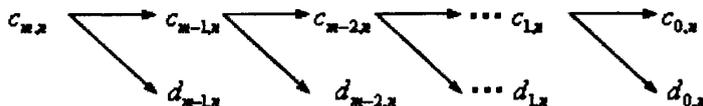


图1 Mallat 分解示意图

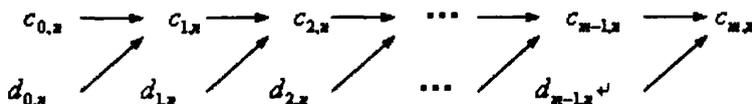


图2 Mallat 重构示意图

对于静态图像这样的二维信源, 需要使用二维滤波器进行处理。这时, 我们要考虑 $L^2(\mathbb{R}^2)$ 中的二维多分辨率分析。考虑到小波函数的可分离性, $L^2(\mathbb{R}^2)$ 中的二维多分辨率分析可由 $L^2(\mathbb{R}^2)$ 中的多分辨率分析推广生成, 进而可得如下二维 Mallat 分解与重构算法:

定义 5 二维 Mallat 分解与重构算法

对于双正交小波, 设 $(\{h_k\}_{k \in \mathbb{Z}}, \{g_k\}_{k \in \mathbb{Z}})$ 与 $(\{\widehat{h}_k\}_{k \in \mathbb{Z}}, \{\widehat{g}_k\}_{k \in \mathbb{Z}})$ 是双正交小波导出的一组对偶滤波器, 则图像的双正交小波分解和重构公式为

$$\left\{ \begin{array}{l} C_{j,k,m} = \sum_{l,n} \overline{\hat{h}_{l-2k}} \overline{\hat{h}_{n-2m}} C_{j+1,l,n} \\ D_{j,k,m}^1 = \sum_{l,n} \overline{\hat{h}_{l-2k}} \overline{\hat{g}_{n-2m}} C_{j+1,l,n} \\ D_{j,k,m}^2 = \sum_{l,n} \overline{\hat{g}_{l-2k}} \overline{\hat{h}_{n-2m}} C_{j+1,l,n} \\ D_{j,k,m}^3 = \sum_{l,n} \overline{\hat{g}_{l-2k}} \overline{\hat{g}_{n-2m}} C_{j+1,l,n} \end{array} \right. \quad (2.13)$$

$$\begin{aligned} C_{j+1,k,m} = & \sum_{l,n} \overline{\hat{h}_{k-2l}} \overline{\hat{h}_{m-2n}} C_{j,l,n} + \sum_{l,n} \overline{\hat{h}_{k-2l}} \overline{\hat{g}_{m-2n}} D_{j,l,n}^1 \\ & + \sum_{l,n} \overline{\hat{g}_{k-2l}} \overline{\hat{h}_{m-2n}} D_{j,l,n}^2 + \sum_{l,n} \overline{\hat{g}_{k-2l}} \overline{\hat{g}_{m-2n}} D_{j,l,n}^3 \end{aligned} \quad (2.14)$$

事实上，二维小波分解与重构算法是一维算法的推广，其中既有行变换，又有列变换。利用二维 Mallat 分解算法，我们可以把一幅图像 $C_{j+1} = \{C_{j+1,k,m}\}_{k,m}$ 分解为一个低频子图(子带) C_j 和水平、垂直与对角线三个方向的高频子图 D_j^1 、 D_j^2 与 D_j^3 。同样，可以把低频子图 C_j 分解为 C_{j-1} 和水平、垂直与对角线三个方向的高频子图 D_{j-1}^1 、 D_{j-1}^2 与 D_{j-1}^3 。二维小波包分解除了对低频子图进行分解外，对高频子图也进一步作二维分解。

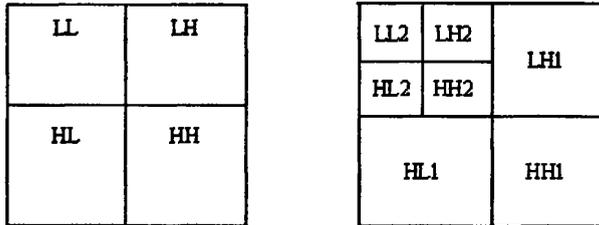


图3 图像的2层小波分解示意图

其中，L表示低通滤波器，H表示高通滤波器。图4是灰度图象Lena，像素为512×512，图5，图6分别是Lena的进行1层和2层小波分解的示意图。



图4 Lena 原图



图5 Lena 的 1 层小波分解



图6 Lena 的 2 层小波分解

§ 2.5 小波包

小波变换虽然具有频率越高相应的时间（或空间）分辨率越高的优点，但其在频率域上的分辨率却相应降低。小波包在一定程度上弥补了小波的这一缺陷。小波包变换是小波变换的推广，它能够提供一种更加精细的分析方法，它对小波变换没有细分的高频部分进行进一步的分解，并能够根据被分析信号的特征自适应的选择相应的频带，使之与信号频谱相匹配，从而提高了时频分辨率，因此小波包具有更广泛的应用价值。

2.5.1 小波包定义

对于尺度子空间 V_j 和小波子空间 W_j ，若令 $U_j^0 = V_j$ ， $U_j^1 = W_j$ ，则 $V_{j+1} = V_j \oplus W_j$ 可用 U_j^n 的统一表示为

$$U_{j+1}^0 = U_j^0 \oplus U_j^1 \quad (2.15)$$

定义子空间 U_j^n 是函数 $u_n(t)$ 的闭包空间，而 $U_j^{2^n}$ 是函数 $u_{2^n}(t)$ 的闭包空间，并令 $u_n(t)$ 满足下面的双尺度方程：

$$\begin{cases} u_{2n}(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k u_n(2t-k) \\ u_{2n+1}(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k u_n(2t-k) \end{cases} \quad (2.16)$$

其中 $g_k = (-1)^k h_{-k}$ ，特别地，当 $n=0$ 时，由以上两式直接得到 $u_0(t)$ 和 $u_1(t)$ 分别为尺度函数 $\varphi(t)$ 和小波函数 $\psi(t)$ 。 $\{h_k\}$ 和 $\{g_k\}$ 本身所要满足条件和小波滤波器系数一样，如公式 $\sum_n h_n \overline{h_{n-2k}} = 2\delta_{k,0}$ 和 $\sum_k h_{2k} = \sum_k h_{2k+1} = 1$ 。这时有

$$U_{j+1}^n = U_j^{2^n} \oplus U_j^{2^{n+1}} \quad (2.17)$$

这样就可以对小波子空间再进行二进细分。

定义 6 小波包^[45]

由(2.16)构造的序列 $\{u_n(t)\}_{n=0}^{\infty}$ 称为由基函数 $u_0(t) = \varphi(t)$ 确定的小波包。

由于 $\varphi(t)$ 由 h_k 唯一确定，所以又称 $\{u_n(t)\} (n \in \mathbb{Z})$ 为关于序列 $\{h_k\}$ 的正交小波包。

对于任意的非负整数 n ， n 有唯一的二进制表示： $n = \sum_{j=1}^{\infty} \varepsilon_j 2^{j-1}$ ， $\varepsilon_j = 0$ 或 1 ，

则小波包函数 $u_n(t)$ 的 Fourier 变换为： $\hat{u}_n(\omega) = \prod_{i=1}^{\infty} m_{\varepsilon_i}(\omega/2^i)$ 。其中

$$m_0(\omega) = \frac{1}{2} \sum_{k=-\infty}^{+\infty} h_k e^{-ik\omega}, \quad m_1(\omega) = \frac{1}{2} \sum_{k=-\infty}^{+\infty} g_k e^{-ik\omega}.$$

集合 $\left\{2^{j/2} u_n(2^j t - k)\right\}_{j,k,n \in \mathbb{Z}}$ 称为一个小波包库, 从小波包库中选择能构成 $L^2(\mathbb{R})$

的一个基函数系称为 $L^2(\mathbb{R})$ 的一个小波包基。显然在小波包库中能组成一组基的选择有很多, 可以进行比小波变换更细致的分析。设 $\{U_j^n, (j,n) \in \Omega\}$ 可以覆盖整个 $L^2(\mathbb{R})$ (为了方便叙述, 这时也称 $\{U_j^n, (j,n) \in \Omega\}$ 构成 $L^2(\mathbb{R})$ 的一组基), 则任意的 $f(t) \in L^2(\mathbb{R})$, 在 $\{U_j^n, (j,n) \in \Omega\}$ 中都可以表示成

$$f(t) = \sum_{(j,n) \in \Omega} \sum_k d_k^{(j,n)} \cdot 2^{j/2} u_n(2^j t - k) \quad (2.18)$$

其中 $d_k^{(j,n)} = \left\langle f(t), 2^{j/2} u_n(2^j t - k) \right\rangle$ 称为小波包系数。

小波包有尺度指标、位置指标和频率指标三个参数, 而小波只有离散尺度和离散平移两个参数, 正是因为小波包增加了频率参数, 使得小波包克服了当小波时间分辨率越高时, 而频率分辨率却越低的缺陷。

2.5.2 小波包分解与重构

类似于小波, 对于公式(2.18)的小波包系数 $d_k^{(j,n)}$, 有如下小波包分解算法:

$$\begin{cases} d_k^{(j+1,2n)} = \frac{1}{\sqrt{2}} \sum_i \overline{h_{i-2k}} d_i^{(j,n)} \\ d_k^{(j+1,2n+1)} = \frac{1}{\sqrt{2}} \sum_i \overline{g_{i-2k}} d_i^{(j,n)} \end{cases} \quad (2.19)$$

以及下波包重构算法:

$$d_k^{(j,n)} = \sum_i \left(h_{k-2i} d_i^{(j+1,2n)} + g_{k-2i} d_i^{(j+1,2n+1)} \right) \quad (2.20)$$

和小波一样, 小波包也很容易推广到二维的情况。

2.5.3 最优基的选择

小波库中包含多组 $L^2(\mathbb{R})$ 的规范正交基, 那么对于具体信号 $f(t)$ 来说, 我们可以把信号 $f(t)$ 在正交基下进行展开, 从而 $f(t)$ 与系数序列等同起来。因此, 衡量基的好坏需要定义一个反映系数序列本质的函数 $M(\{x_k\})$, 其中 $\{x_k\}$ 是 $f(t)$ 在某组基下的系数序列。 $M(\{x_k\})$ 称为代价函数, 小波库中使代价函数最小的基底称为最优基。

当系数序列 $\{x_k\}$ 的能量集中在少数几个系数上, 而大多数的系数可以忽略不计时, 我们认为这组基比较好, 代价函数 M 的值应该比较小; 当系数序列 $\{x_k\}$ 分布比较均匀时, 我们认为这组基不好, M 的值应该比较大; 代价函数 M 具有可加性, 即 $M(0) = 0, M(\{x_k\}) = \sum_k M(x_k)$ 。

代价函数的定义很多, 但目前用得最多的是把 Shannon 熵作为序列 $x = \{x_k\}$ 的代价函数, 即

$$M(x) = -\sum_k p_k \log p_k \quad (2.21)$$

其中, $p_k = \frac{|x_k|^2}{\|x\|^2}$, 当 $p = 0$ 时, 令 $p \log p = 0$ 。上式定义的代价函数不满足可

加性, 只满足半可加性。用熵来定义代价函数是由于在一定均方误差条件下, $\exp M(x)$ 正比于表示信号所需的系数数目。

一个正交基库称为一个二元树, 如果它满足:

1. 基向量组成的子集等价于非负整数集 Z_+ 中一个具有如下形式的子集合:

$$I_{j,n} = \{2^j n, \dots, 2^j(n+1) - 1, j \in \mathbb{Z}, n \in \mathbb{Z}_+\}$$

2. 库中的每一个基对应于 Z_+ 的一个由 $I_{j,n}$ 组成的不相交的覆盖;

3. 若 $U_{j,n}$ 是与 $I_{j,n}$ 等价的子空间, 则 $U_{j,n} = U_{j-1,2n} \oplus U_{j-1,2n+1}$ 。

则小波库可以组织成一个二元树, 其树结构如图 7 所示。

给定代价函数 M , x 是子空间 U 中的一个信号, 记 B 为从小波库中选取的一组正交基, Bx 是在基 B 下的系数。

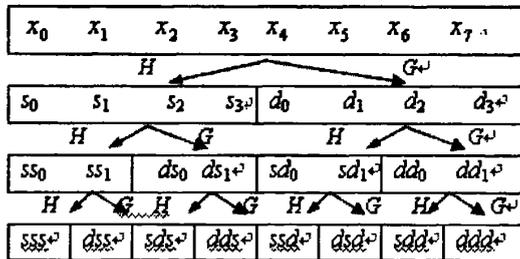


图7 由小波包组成的二元树

定义 7

设 $x \in U$, M 给定, 如果 $M(Bx)$ 是最小的, 则 B 为最优基。

在实际应用中, 只要对信号做有限次分解即可, 可用归纳法找到 $f(t)$ 的最优基。对于 $f(t)$, 首先, 采用小波包算法可计算出 $f(t)$ 在各子空间上的系数, 假设分解三层, 如图 7 所示; 其次, 计算出在各层上系数的代价函数值, 令上层框称为父框, 其下一层框称为子框 (每个父框对应于两个子框); 最后, 采用如下算法

找到 $f(t)$ 的最优基:

a) 从最底层的框开始, 把每个代价函数值全标上“*”号;

b) 把两个子框中的值之和与其父框中的值比较, 如果两子框中的值之和大于其父框中的值, 就把父框中的值标上“*”号; 否则, 就用这个值代替父框中的值, 把父框中的值用括号括起来, 并把其两子框中的值之和写在外面;

c) 只考虑括号外的值, 按步骤 b) 进行到最顶层;

d) 从最顶层开始, 每一层只选择第一次遇到的标“*”号的框, 该框以下的子框就不再考虑了。

这样选出的带“*”号的框, 恰好对应于的一组正交分解, 由此可选出对应于 $f(t)$ 的最优基 (相对代价函数 M)。

第三章 人脸识别的算法研究

§ 3.1 特征提取概述

在模式识别中,由被识别的对象产生一组基本特征,这些特征可以是计算出来的(当识别对象是波形或数字图像时),也可以是用仪表或传感器测量出来的(当识别对象是实物或某种过程时),这样产生出来的特征称为原始特征。原始特征的数量可能很大,或者说样本处于一个高维空间中,通过映射(或变换)的方法可以用低维空间来表示样本,这个过程叫做特征提取。映射后的特征叫二次特征,他们是原始特征的某种组合(通常是线性组合)。

特征提取是模式识别或分类中的核心问题,它直接影响到分类器的设计和性能。如果不同类别的特征差别较大,就比较容易设计出性能较好的分类器。所以对识别或分类来说关键不在于完整地描述模式,而是提取模式中有效的分类特征。所谓有效分类特征就是不同模式类差别较大的特征。但这些特征在原始特征域通常不易被观察或检测。特征提取就是通过变换的方法,使这些重要的特征在变换域显示出来,去掉对分类无意义的信息,这样把原始的高维空间变为低维的特征空间。

然而人脸图像是一种非平稳的信号,通常包含长时低频(面部)和短时高频(边缘)不同尺度的特征,用于分类的特征往往包含在这些局部的时频信息中,比如面部、眼睛、嘴巴等特征部位的轮廓大小和位置。由于人脸图像的数据量非常的大,为了有效地进行分类识别,就要对原始图像进行压缩,并得到最能反应分类本质的特征,这就是人脸识别的特征提取过程。考虑到人脸非常相似的特性,寻找稳定和有效的识别特征成为解决人脸识别问题的关键。在模式识别理论中,特征提取的一般原则是所抽取的特征之间相关性越小越好,最好是提取不相关的特征。

§ 3.2 人脸识别的相关统计特征

由于人脸图像一般处于高位的图像空间,且含有大量由背景、表情、光照、姿态、位置及尺寸等大幅度变化所引起的噪声,不利于分类,并且计算复杂,因此无论是从计算复杂度,还是从分类性能的角度上来讲,直接在高维的原始图像空间进行分类是不可取的。现阶段,人们往往通过某种映射(或变换) W 的方法将图像从高维空间投影至某一低维子空间,使得人脸图像在该子空间的分类更为紧凑,另外计算的复杂度也大为降低。这一过程从广义上来说也可称为特征提取。

根据不同的性能要求，得到的映射 W 是不一样的，即提取的特征是不同的。

基于统计的人脸识别将人脸图像看作一个的高维空间的随机矢量，在高维空间中对分布信号进行变换以获取某种统计特征作为人脸的基本特征，在特征空间实现人脸检测和归类。

把图像看成是 d 维的随机变量，并服从多为正态分布。设有 K 类 d 维样本空间 $X = \{w_1, w_2, \dots, w_K\}$ ，其中 $w_k = \{x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,N_k)}\}$ 是第 k 类样本集，含有 N_k 个样本。 $x^{(k,i)} = (x_1^{(k,i)}, x_2^{(k,i)}, \dots, x_d^{(k,i)})^T$ 是第 k 类的第 i 个样本向量， d 为样本向量维数（图像像素数）。

$N = \sum_{k=1}^K N_k$ 为样本总数，则每类样本的先验概率 $\hat{P}(w_k)$ 由各类样本个数

来估计， $\hat{P}(w_k) = \frac{\#(w_k)}{\#(X)}$ ，第 k 类样本均值 μ_k 的估计 $\hat{\mu}_k = m_k = \frac{1}{N_k} \sum_{x \in w_k} x$ ，总体均值

μ 的估计 $\hat{\mu} = m = \frac{1}{N} \sum_{k=1}^K \sum_{x \in w_k} x$ 。

人脸识别的目的就是对于任意一副已经选区出来的人脸图像 x （样本向量），需要确定 x 属于样本空间中的哪一类样本集。

定义 8 类内散度矩阵 S_w

$$S_w = \sum_{k=1}^K \frac{N_k}{N} \sum_{x \in w_k} \frac{1}{N_k} (x - m_k)(x - m_k)^T = \frac{1}{N} \sum_{k=1}^K \sum_{x \in w_k} (x - m_k)(x - m_k)^T \quad (3.1)$$

且有类内协方差 $\Sigma_w = \sum_{k=1}^K P(w_k) \Sigma_{w_k} = \sum_{k=1}^K P(w_k) E[(x - \mu_k)(x - \mu_k)^T]$ 的估计为

$$\hat{\Sigma}_w = S_w。$$

定义 9 类间散度矩阵 S_b

$$S_b = \sum_{k=1}^K \frac{N_k}{N} (m_k - m)(m_k - m)^T \quad (3.2)$$

且有类间协方差 $\Sigma_b = E[(\mu_k - u)(\mu_k - u)^T]$ 的估计 $\hat{\Sigma}_b = S_b$ 。

定义 10 总体散度矩阵 S_t

$$S_t = \frac{1}{N} \sum_{k=1}^K \sum_{x \in w_k} (x - m)(x - m)^T \quad (3.3)$$

此时有总体协方差 $\Sigma_t = E[(x - u)(x - u)^T]$ 的估计 $\hat{\Sigma}_t = S_t$ ，且 $S_t = S_w + S_b$ 。

定义 11 总体向量平均距离

$$tr(S_t) = \frac{1}{2} \sum_{k_1=1}^K \frac{N_{k_1}}{N} \sum_{k_2=1}^K \frac{N_{k_2}}{N} \frac{1}{N_{k_1} N_{k_2}} \sum_{i_1=1}^{N_{k_1}} \sum_{i_2=1}^{N_{k_2}} \|x^{(k_1, i_1)} - x^{(k_2, i_2)}\|^2 \quad (3.4)$$

定义 12 平均类内距离

$$\text{tr}(S_w) = \frac{1}{N} \sum_{k=1}^K \sum_{i=1}^{N_k} \|x^{(k,i)} - m_k\|^2 \quad (3.5)$$

定义 13 平均类间距离

$$\text{tr}(S_t) = \sum_{k=1}^K \frac{N_k}{N} \|m_k - m\|^2 \quad (3.6)$$

目前有两种经典的寻找线性变换的方法,事实上这两种方法就是经典的统计模式识别方法,即特征脸方法和 Fisher 脸方法。特征脸方法是从 PCA 方法导出的,而 Fisher 脸方法是从 LDA 方法导出的。PCA 方法的目的是寻找在最小均方差意义下能够代表原始数据的投影方法。Turk^[46]等首次将其应用于人脸识别当中,提出了经典的特征脸(Eigenfaces)方法,形成了子空间分析方法的基础。另一种方法为基于 Fisher 准则的线性鉴别分析(LDA)^[24],其目的是以样本的可分性为目标,寻找在最小均方差意义下,最能够分开各类数据的投影方法。下面首先介绍这两种方法。

§ 3.3 主成分分析方法 (PCA)

主成分分析方法 (Principal Component Analysis. PCA) 在人脸识别领域成功应用的一个重要理论基础是较好的解决了协方差矩阵的特征向量的求解问题。人脸识别是一个典型的高维小样本问题,即人脸图像向量的维数一般较高,比如,实验用的 ORL 人脸库的图像大小为 92×112 的人脸图像,其对应的图像向量特征空间高达 10304 维,在如此高维的图像空间内,按照通常的算法,计算样本的协方差矩阵的特征向量是非常耗时的。由于人脸识别训练样本的数目一般较小,训练样本的总数远远小于人脸图像向量的维数。因此,针对高维小样本的情况,求解特征向量所采取算法的基本思想是,将高维的问题转化为低维的问题加以解决。

3.3.1 主成分分析的概述

主成分分析(Principal Component Analysis. PCA)是一项在计算机视觉中用于降低维数、提取特征的技术,它被认为是人脸识别的有效方法之一,由 Kirby 和 Sirovich^[28]首先引入到人脸识别领域,提出应用 Karhunen-loeve 变换表征人脸,即人脸由被称之为特征脸向量的带权特征向量的线性组合表征。1991 年, Turk 和 Pentland^[29]开发了基于 PCA (K-L 变换的延伸)的人脸识别系统。

主成分分析法 (PCA) 是模式识别判别分析中最常用的一种线性映射方法,该方法是根据样本点在多维模式空间的位置分布,以样本点在空间中变化最大方向,即方差最大的方向,作为判别矢量来实现数据的特征提取与数据压缩的。从概率统计观点可知,一个随机变量的方差越大,该随机变量所包含的信息就越多,

如当一个变量的方差为零时, 该变量为一常数, 不含任何信息。所谓主成分就是原始数据的 m 个变量经线性组合 (或映射) 后得到的变量, 该变化使得其变换后的变量方差为最大 (第一主成分) 的部分。各个主成分之间是相互线性无关 (正交的), 从第一主成分往后, 主成分按方差大小的顺序排列 (对应特征值按大小顺序排列)。对于特征值为 λ_i 的主成分, λ_i 也是该主成分的方差, 该值表示样本点在该主成分方向上的离散程度, 主成分 λ_i 的贡献率 η_i 可表示为 $\eta_i = \lambda_i / (\lambda_1 + \lambda_2 + \dots + \lambda_p)$ 。

主成分中方差较小或 λ_i 较小的主成分被认为包含的是噪声, 在分析时可去除这部分变量, 减少分析的主成分, 从而达到降维的目的。主成分中任两个, 可构成判别分析平面, 因此可实现高维空间向二维平面及其他维平面映射的目的。通常取方差较大的几个主成分构成判别分析平面。

3.3.2 主成分分析的原理

设图像数据库图像大小为 $Width \times Height$, 令 $d = Width \times Height$ 。把图像看成是 d 维的随机变量, 并服从多为正态分布。

PCA 方法的核心就是寻找正交变换矩阵 $W = (u_1, u_2, \dots, u_L) \in \mathbb{R}^{d \times L}$, $WW^T = 1$, 使得 $J(W) = tr(W^T S_i W)$ 达到最大。也就是在 $y = W^T x$ 的正交变换后使得总体矩阵 $W^T S_i W$ 的迹最大, 即各个样本分离的尽量远, 将样本的方差尽量的保留下来, 和原样本的均方差尽量的小。实际上, $W = (u_1, u_2, \dots, u_L)$ 就是对应于矩阵 S_i 的前 L 个最大特征值对应的特征向量组成的矩阵, 即 u_i 是方程 $S_i u_i = \lambda_i u_i$ 的解。

对于图像高维数据, 总体散度矩阵 S_i 是一个很大的矩阵, 计算其特征值和特征向量是很困难的。而根据奇异值分解定理 (SVD) [47], $d \times N$ 维的矩阵 Φ_i 存在两个正交矩阵 U 和 V , 使得 $\Phi_i = U \Lambda^{\frac{1}{2}} V^T$, 其中 U 和 V 的列向量分别是 $\Phi_i \Phi_i^T$ ($d \times d$ 维) 和 $\Phi_i^T \Phi_i$ ($N \times N$ 维) 的特征向量, Λ 是相应的特征值组成的对角矩阵, 并且有 $U = \Phi_i V \Lambda^{\frac{1}{2}}$ 。因为 $S_i = \frac{1}{N} \Phi_i \Phi_i^T \in \mathbb{R}^{d \times d}$ (通常 $d \gg N$), 因此可以先求出矩阵 $R = \frac{1}{N} \Phi_i^T \Phi_i \in \mathbb{R}^{N \times N}$ 的特征向量组成的矩阵 V 和相应的特征值 Λ , 根据 SVD 定理计算 S_i 的特征向量 U , 这样就可以大大减少计算量了。

3.3.3 基于主成分分析的人脸识别

用 PCA 降维并用于人脸识别主要基于以下三点: (1) 压缩功能。在低维空间内比较图像将提高计算效率; (2) 人脸样本的分布近似正态分布。方差大的维可能与

有用信号相关，而方差小的维可能对应噪声，因此去掉小方差对应的维将有利于提高识别精确率；(3)因为每幅图像都被减去均值，且被放缩成单位向量，两幅图像之间的相关性与特征空间中投影之间的距离成反比，因此特征空间中的最近邻匹配是图像相关性的有效近似。

特征脸方法是从 PCA 方法导出的一种人脸识别和描述技术。特征脸方法就是将包含人脸的图像区域看作是一种随机向量，因此可以采用 K-L 变换获得其正交 K-L 基底。对应其中较大的特征值的基底具有与人脸相似的形状，因此又称为特征脸。利用这些基底的线性组合可以描述、表达和逼近人脸图像，因此可以用来进行人脸识别。识别过程就是将人脸图像映射到由特征脸张成的子空间上，与已知人脸在特征空间中的位置相比较，具体步骤为：

1. 获得人脸图像的训练集，计算特征脸并定义为人脸空间，存储在模板库中，以便进行人脸识别；
2. 输入新的人脸图像，将其映射到特征脸空间，得到一组关于该人脸的特征数据；
3. 通过图像与人脸空间的距离，判断该图像是否为人脸；
4. 若为人脸，根据权值模式判断它是否为数据库中的某个人，并做出具体的操作。

首先，计算特征脸。设人脸图像用二维 $Width \times Height$ 向量 X 表示。人脸图像训练集为 $\{X_k\}$ ，其中 $k=1, \dots, M$ (M 为训练集中图像总数)。这 M 幅图像的平均向量为：

$$\mu = \frac{1}{M} \sum_{k=1}^M X_k \quad (3.7)$$

每个人脸 X_k 与平均人脸 μ 的差值向量：

$$\phi_i = (X_k - \mu) \quad (i=1, \dots, M) \quad (3.8)$$

训练图像的协方差矩阵可表示为：

$$C = AA^T \quad A = [\phi_1, \dots, \phi_M] \quad (3.9)$$

特征脸有协方差矩阵 C 的正交特征向量组成。对于 $Width \times Height$ 人脸图像，协方差矩阵 C 的大小为 $Width^2 \times Height^2$ ，对它进行求解特征值和特征向量是很困难的，用一种新的方法来替代求解。令：

$$L = A^T A$$

此矩阵为协方差矩阵 C 的转置矩阵，是 $M \times M$ 的一个较小的矩阵。首先计算 $M \times M$ 矩阵 L 的特征向量 $v_l (l=1, \dots, M)$ ，则矩阵 C 的特征向量 $u_l (l=1, \dots, M)$ 由差值图像 $\phi_i (i=1, \dots, M)$ 与 $v_l (l=1, \dots, M)$ 线性组合得到：

$$U = [u_1, \dots, u_M] = \left[[\mu_1, \dots, \mu_M]^T \right]^{-1} [v_1, \dots, v_M]$$

实际上, $m(m < M)$ 个特征值对于人脸识别来说已经足够了, 所以, 仅取 L 的前 m 个最大特征值的特征向量计算特征脸。

基于特征脸的人脸识别过程由两个阶段组成: 训练阶段和识别阶段。

在训练阶段, 对于二维 $Width \times Height$ 的人脸图像, 每个已知人脸 X_k , 映射到由特征脸张成的子空间上, 得到 $Width$ 维向量 Y_k :

$$Y_k = W^T (X_k - \mu) \quad (3.10)$$

其中 $k = 1, \dots, N_c$ (N_c 为已知人数), $W = (u_1, u_2, \dots, u_L)$ 为人脸投影空间。

人脸之间距离的门限值定义为: $\theta = \frac{1}{2} \max_{j,k} \{\|Y_j - Y_k\|\} (j, k = 1, \dots, N_c)$ 。

在识别阶段, 首先把待识别图像 X 映射到特征脸空间, 并得到向量 Y :

$$Y = W^T (X - \mu) \quad (3.11)$$

Y 与每个人脸集的距离定义为:

$$\varepsilon_k^2 = \|Y - Y_k\|^2 \quad (k = 1, \dots, N_c) \quad (3.12)$$

为了区分人脸和非人脸, 还需要计算原始图像 X 与由特征脸空间重建的图像 X_f 之间的距离 $\varepsilon: \varepsilon = \|X - X_f\|^2$, 其中 $X_f = UY + \mu$ 。

人脸分类规则如下:

1. 若 $\varepsilon \geq \theta_c$, 则输入图像不是人脸图像;
2. 若 $\varepsilon < \theta_c$, 且 $\forall k, \varepsilon_k \geq \theta_c$, 则输入图像包含未知人脸;
3. 若 $\varepsilon < \theta_c$, 且 $\forall k, \varepsilon_k = \min\{\theta_k\} < \theta_c$, 则输入图像为人脸库中的第 k 个人的人脸。

从压缩能量的角度看, PCA 方法是最优的。它不仅使得从 d 维空间降到 L 维空间前后的和原样本的均方误差最小, 而且变换后的低维空间有很好的人脸表达能力, 其中最大的前面几个特征值对应的分量集中了大部分的信息, 较小的特征值对应的分量则代表了人脸的高频信息, 去掉后面的几个分量也就意味着可以忽略一些表情饰物等细小的差别。

但是, PCA 方法对外在因素带来的图像差异和人脸本身带来的差异是不加任何区分的, 即没有考虑到人脸的类别信息, 实验表明, PCA 方法还存在着以下不足:

1. PCA 具有很强的表征能力, 但辨别能力较差, 即由 PCA 方法表征人脸, 非常相似不仅会出现在同一人的不同脸部图像之间, 而且也会出现在不同人的图像之间;

2. 特征脸方法随着类别数目的增加,光照、姿势等因素的引入,识别率急剧下降^[48],因此特征脸方法还存在着理论上的缺陷。

PCA 方法需要大量的训练时间,训练时间由图像向量的维数决定,如对分辨率为 $d = m \times n$ 的图像,训练时间为 $O(d^2)$ 。即使采用奇偶定理,若训练集大小为 N , N 通常远远小于 d ,训练时间减少为 $O(N^2)$,但对于大型数据库,仍然是很大的负担。

§ 3.4 线性判别分析方法 (LDA)

在模式识别领域中,判别分析法是一种经典的监督特征提取算法,其基本思想就是在 Fisher 准则函数取得最大值的条件下,寻找一组变换使得变换后样本的类内散度最小,类间散度最大。与非监督的 PCA 相比,判别分析法是以样本的类别可分性为目标来提取最优判别特征 (Most Discriminant Feature, MDF),而 PCA 是以样本的最小重建误差为目标来提取最优描述特征 (Most Expressive Feature, MFF),因此,对于模式识别问题,通常判别分析法要优于 PCA。目前,判别分析法已经广泛的应用于文本分类、人脸识别、语音识别等领域。

线性判别分析 (Linear Discriminant Analysis, LDA) 是判别分析的线性形式,也称为 Fisher 脸方法^[49],是基于样本的类别进行降维的方法,利用了类别归属信息,它选择类内散度正交的矢量作为特征脸空间,从而压制了图像之间与识别信息无关的差异,强调了不同人脸之间的差别,同时弱化了同一人脸由于光照、视角和表情而引起的变化,获得了比特征脸更好的识别效果。Belhumeur^[24]对 16 个人得各 10 幅图像进行识别,PCA 方法的识别率为 81%,而 Fisher 脸方法的识别率为 99.4%。在模式识别中有着广泛的应用。

3.4.1 线性判别分析的原理

线性判别分析 (Linear Discriminant Analysis, LDA) 方法和特征脸法一样都认为图像的全局结构对于识别最为重要,这两种方法都将人脸图像对应的象素点连接起来看作一个矩阵,利用一定的方法获得图像的整体特征信息,这一整体的特征信息能够以较少的数据量代表原始图像数据的尽可能多的信息量。特征脸法的目的是选择一个缩小维数的变换方向使得在该方向上的投影的样本类间散度最大,即投影之后的方差为最大,这种方法的一个最主要的缺点就是,在增加样本类间散度的同时也加大了样本的类内散度,从而造成了投影空间中类间数据混乱,不能很好的进行聚类,因此特征脸法不能为识别阶段提供可靠的识别数据。而线性判别方法 (LDA) 则是在特征脸方法的基础上进行的改进,其目标是建立一个

更低维的，且只包含最基本元素的空间，可以更好的描述一组人脸的脸部特征空间数据，这些脸部空间是基于线性合并的方法生成的，它们之间互不相关，并且转化后的脸部空间是可分的，投影使得在扩大类间散度的同时，缩小了类内散度，弥补了特征脸方法的不足，为有效的人脸识别提供了可靠的数据。

LDA 方法的目的是寻找正交变换矩阵 $W = (u_1, u_2, \dots, u_L) \in \mathbb{R}^{d \times L}$, $WW^T = 1$, 使得 $J(W) = \text{tr}(W^T S_b W) / \text{tr}(W^T S_w W)$ 达到最大。也就是线性变换之后使得不同类样本（平均类间距离）尽可能远，同类样本（平均类内距离）尽可能近。这里求得 $W = (u_1, u_2, \dots, u_L) \in \mathbb{R}^{d \times L}$ 是对应于矩阵 $S_w^{-1} S_b$ 的前 L 个最大特征值的特征向量，即对特征方程 $S_b u_i = \lambda_i S_w u_i$ 求解。由于 S_w 的秩 $\text{rank}(S_w) \leq N - K < d$, 因此 S_w 是奇异的，Fisher 脸的方法就是针对该问题（Fisher 脸方法是从 LDA 方法导出的）。首先用 PCA 将原问题的 S_w 降维，然后再采用 LDA 的方法求解。

首先，用 PCA 降维，将 S_w 降至 p 维，其中令 $p = N - K$ 。

$$\widehat{S}_w = W_{pca}^T S_w W_{pca}, \quad \widehat{S}_b = W_{pca}^T S_b W_{pca}, \quad \text{其中 } W_{pca} = \arg \max_W \text{tr}(W^T S_i W),$$

$W_{pca} = (u_{pca_1}, u_{pca_2}, \dots, u_{pca_p})$ 为 S_i 最大的前 p 个特征值对应的特征向量。

$$\text{然后，用 LDA 方法求出 } W_{lda}, \quad W_{lda} = \arg \max_W \frac{\text{tr}(W^T \widehat{S}_b W)}{\text{tr}(W^T \widehat{S}_w W)}.$$

最后求出的投影矩阵为 $W_{opt}^T = W_{lda}^T W_{pca}^T$ 。

不难看出，对于图像高维数据（比如 256 行 256 列的图像堆起一个样本向量来就是 65536 维）， S_i, S_b, S_w 都是很大的矩阵，即使根据奇异值分解定理(SVD)计算其特征根的复杂度也为 $O(d^3)$ ，相应的存储空间也大得惊人，因此要达到实时处理的要求在普通计算机上几乎不可能实现。

3.4.2 基于 Fisher 脸的人脸识别

Fisher 线性判别在计算最优的投影矩阵 W_{opt} 时，利用了类别信息。它使投影后的样本的类间散度矩阵和类内散度矩阵的比值最大，也就是在投影空间中，属于不同类别的样本尽可能的分开，而属于同一类别的样本尽可能的聚合在一起。但是，单纯采用 LDA 方法获取子空间也存在着问题：

1. 矩阵计算非常复杂，因此计算复杂度高，计算起来非常耗时；
2. 由于图像像素个数通常比训练图像要多，因此样本类内散度矩阵 S_w 是奇异的，造成 $S_b u_i = \lambda_i S_w u_i$ 不可解，也就是所谓的小样本集合问题。

第四章 基于小波包矩特征提取的改进方法

基于统计特征的方法将人脸图像视为随机向量,从而用一些统计方法来分析人脸模式,这类方法有着完备的统计学理论支持,是人脸识别领域较为成功的方法。统计方法是通过计算人脸库中的人脸的统计特征如二阶矩、高阶矩,从而建立人脸空间中的一个新的坐标系。人脸图像在这个坐标系下的投影即为该图像的特征。对于一张待识别的人脸图像,求出该图像在这个坐标系中的坐标也就是求出了这个图像的特征,然后通过与人脸库中的图像在该坐标系下的坐标相比较,就可以完成对这一幅人脸图像的识别。这类方法主要有主特征分析(PCA)^{[50][51]}、特征脸(Eigenface)方法^[29]、Fisher脸方法^[49]、独立成分分析(ICA)^[52]、贝叶斯人脸识别方法^{[53][54][55]}等。相对于其他几类人脸识别的方法,如基于几何特征、基于模型、神经网络等,基于统计的方法具有计算相对简单,性能稳定、识别率较高等优点,具有较强的实用价值,已经有较为成熟的产品。但是,对于光照角度和光照强度较为敏感,同时对于尺度变化、旋转和遮挡等影响因素,识别率急剧下降。随着商业、政府、军事领域对人脸识别精度与速度要求的提高,特别多媒体数据高层语义自动提取与挖掘技术的需要,我们希望能够保证更高的识别率的同时实现实时识别的目标。

考虑到人脸图像是一种非平稳信号,用于分类的特征往往包含在局部的时频信息中,用一般的变换方法提取有效特征比较困难。小波变换是近年来发展起来的一种分析非平稳信号的有效方法,而且可以获得与人的视觉特性更为接近的多分辨率的特征提取效果,同时,二维小波变换快速算法的出现极大的降低了特征提取的计算复杂度,于是小波被人们应用于人脸识别中。而小波包变换具有任意的多分辨率分解特性,可以提供更丰富的基以供分类选择。近来,如何将小波变换和基于统计特征的人脸识别结合起来以提高人脸识别的正确率和识别速度成为人们研究的热点。G C Feng^[56]等研究了在小波包子带上应用PCA的方法,通过在特定的部分小波包子带重构的低精度图像上使用PCA方法,一定程度上降低了计算复杂度并提高了识别正确率。Christophe Garcia^[57]等假设在同类人脸图像条件下,特定人脸属于其中某类的概率分布为高斯分布,提出以小波包分解子带系数均值和方差为特征进行人脸识别,极大的降低了计算复杂度,同时在约束环境下(正面、无遮挡、光照均匀)保持了较高人脸识别的正确度。N. Saito 和 R. R. Coifman^{[58][59][60][61]}提出了 Local Discriminate Bases(LDB)的算法,这种方法在给定的训练样本集对信号进行小波包分解后,按类能量的差异来选择最优基,将该小波包变换对应的具有最大可分度的部分小波包系数作为特征矢量,提取出不同模

式类差别较大的有效的分类特征,提高了人脸识别的正确度,同时对姿态、光照、表情的变化具较好的自适应性。

§ 4.1 人脸识别的 LDB 方法

N. Saito 和 R. R. Coifman^{[58][59][60][61]}提出了 Local Discriminate Bases(LDB)的算法,这种方法在利用训练样本集对人脸图像小波包分解后的系数进行筛选,按类能量的差异来选择最优基,其筛选的原则是将具有最大可分度的部分小波包系数作为特征矢量,提取出不同模式类差别较大的有效的分类特征。

这里,我们依然把图像看成是 d 维的随机变量,并服从多维正态分布。设有 K 类 d 维样本空间 $X = \{w_1, w_2, \dots, w_K\}$, 其中 $w_k = \{x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,N_k)}\}$ 是第 k 类样本集,含有 N_k 个样本。 $x^{(k,i)} = (x_1^{(k,i)}, x_2^{(k,i)}, \dots, x_d^{(k,i)})^T$ 是第 k 类的第 i 个样本向量, d 为样本向量维数(图象像素数)。 $N = \sum_{k=1}^K N_k$ 为样本总数,则每类样本的先验概率 $\hat{P}(w_k)$

由各类样本个数来估计, $\hat{P}(w_k) = \frac{\#(w_k)}{\#(X)}$, 第 k 类样本均值 μ_k 的估计

$$\hat{\mu}_k = m_k = \frac{1}{N_k} \sum_{x \in w_k} x, \text{ 总体均值 } \mu \text{ 的估计 } \hat{\mu} = m = \frac{1}{N} \sum_{k=1}^K \sum_{x \in w_k} x.$$

4.1.1 最初的 LDB 分类器设计方法

最初的 LDB 以规范化能量差作为可分度的度量,利用 LDB 设计分类器主要有四个步骤。

第一步,将所有样本 $\{x^{(k,i)}, 1 \leq k \leq K, 1 \leq i \leq N_k\}$ 先进行小波包分解,分解级数为 $j=1, 2, \dots, L$ 。第 j 级第 n 个小波包子空间为 U_j^n ($j=1, 2, \dots, L$ $n=0, 1, 2, \dots, 2^{j-1}$), $x^{(k,i)}$ 在 U_j^n 中的小波包系数为 $\{d_p^{(k,i,j,n)} | p=1, 2, \dots, P_j\}$ 。计算所有样本 $\{x^{(k,i)}, 1 \leq k \leq K, 1 \leq i \leq N_k\}$ 第 j 级第 n 个小波包子空间 ($j=1, 2, \dots, L$ $n=0, 1, 2, \dots, 2^{j-1}$) 上的小波包系数 $\{d_p^{(k,i,j,n)} | p=1, 2, \dots, P_j\}$ 。

第二步,计算 K 类样本 $\{x^{(k,i)}, 1 \leq k \leq K, 1 \leq i \leq N_k\}$ 在每个小波包系数上的可分度 $J_{j,n,p}^E$ (如图 8), $J_{j,n,p}^E$ 定义为规范类能量之差。 K 类样本在子空间 U_j^n 的第 p 个

系数上的可分性为 K 类样本规范化能量的差异: $J_{j,n,p}^E = \sum_{u=1}^K \sum_{v=1}^K d(E_{j,n,p}^{(u)}, E_{j,n,p}^{(v)})$, 其中

$$d(x, y) \text{ 为距离范数, } d(x, y) = \sqrt{(x-y)^2}. E_{j,n,p}^{(u)} = \sum_{i=1}^{N_k} |d_p^{(u,i,j,n)}|^2 / \sum_{i=1}^{N_k} \|x^{(u,i)}\|^2 \text{ 为第 } u \text{ 类}$$

在子空间 U_j^n 的第 p 个系数上的规范化能量。选取最优基中大于某个阈值 t 的 $J_{j,n,p}^E$

加起来: $J_{j,n}^E = \sum_{p=1}^{P_j} \delta_{j,p,n} J_{j,n,p}^E$ 作为子空间 U_j^n 上的可分度, 从而降低维数, 去掉一些

冗余信息或噪音。

第三步, 选取最优基。

- a) 把最下层的小波包基 $U_j^n (j=L, 0 \leq n \leq 2^{j-1})$ 标上*号 (如图 9)。当前层 $j=L$ 为最下层。
- b) 求第 j 层两个小波包子空间 U_j^{2n} 和 U_j^{2n+1} 的可分度和与相应的上层父结点的 U_{j-1}^n 可分度, 如果 $J_{j-1,n}^E > J_{j,2n}^E + J_{j,2n+1}^E$, 则把 U_{j-1}^n 标上*号, 否则令 $J_{j-1,n}^E = J_{j,2n}^E + J_{j,2n+1}^E$ 。
- c) 令 $j=j-1$, 如果 $j>1$, 重复第 b) 步。
- d) 从最上层开始, 只选第一次遇到标有*号的 U_j^n 的下标集合 (j,n) 为 Ω , 组成一组正交基 $\{U_j^n, (j,n) \in \Omega\}$, 这组基就是最优基 (如图 10)。

50							
20				28			
11		12		13		14	
1	2	3	4	5	6	7	8

图8 计算出所有子空间的可行性

40							
20				21			
12		11		10		9	
1	2	3	4	5	6	7	8

图9 把最下层的小波包基标上*号

49(40)							
23(20)				26(21)			
3 (12*)		7 (11*)		11(10)		15 (9)	
1	2	3	4	5*	6*	7*	8*

图10 确定最优基

第四步, 将所有样本 $\{x^{(k,j)}, 1 \leq k \leq K, 1 \leq i \leq N_k\}$ 在最优基 $\{U_j^n, (j,n) \in \Omega\}$ 上的可分度大于阈值 t 的小波包系数 $\{d_p^{(k,j,n)} | J_{j,n,p}^E > t, 1 \leq k \leq K, 1 \leq i \leq N_k, (j,n) \in \Omega, p=1,2,\dots,P_j\}$ 作为新的特征。由于在最优基上往往很少一部分子空间占据了大部分可分度, 为减少小波包分解和提取系数计算量, 我们从中选择部分子空间构成最优基的子集 $\{U_j^*, (j,n) \in \Omega^*\}$, 然后将训练样本在这个子集上的可分度大于某一个阈值 t 的小波包系数作为新的特征。对于待识别样本 x , 也在最优基的子集上进行小波包分解, 得到相应位置上的系数作

为新的特征，然后根据这些特征设计分类器。

4.1.2 LDB 方法的优点

LDB 方法具有以下优点：

1. 小波包变换可以弥补小波变换在某些情况下高频部分时间分辨率低的缺点，为分类选择提供更丰富的基；
2. 基于多分辨率分解的方法在信号的不同频带上分别以不同算子进行特征提取，因而可以获得与人的视觉特性更为接近的特征提取的效果；
3. LDB 方法从时域和频域上分别提取出结构信息和频率变化信息进行压缩，而不是像 PCA 和 LDA 对原有特征简单的重组来选择压缩信息；
4. 多分辨率分解具有完善的重构能力，可以确保信号在分解过程中没有信息损失；
5. 小波包分解没有冗余，即变换域和原始信号的数据量是一样的。

4.1.3 LDB 用于人脸识别

我们很容易将 LDB 方法用于人脸识别。此时，我们处理的是图像二维信号，我们使用二维小波包分解，每个空间有 4 个子空间，在计算可分度判据的时候，把每个子空间的系数排成一列向量并求和，在选择最优基的第 b)步中，计算四个子空间的可分度的和。

LDB 方法通过在图像不同频带上分别以不同算子提取特征，因而可以获得与人的视觉特性更为接近的特征提取效果，同时小波包变换可以弥补小波变换在某些情况下高频部分时间分辨率低的缺点，可以提供更丰富的基以供分类选择。

小波包分解的层数和使用的的小波要通过实验加以确定^[53]，通常分解层数进行 2 到 3 级，小波基函数选择 Daubechies 小波较好。

§ 4.2 LDB 方法的改进

本文提出了一种新的人脸识别方法 LDB_DS_OM，综合上述三种方法并从三个方面作了改进：一是以差可分度（Difference Separability, DS）代替 LDB 算法中的平均类间、类内距离，以选择小波包分解系数和最优基；二是以选定子带内选定系数的一、二阶原点矩（Origin Moments, OM）作为人脸特征；三是以 Bhattacharyya 距离分量作为子带分类特征向量距离度量，以各子带差可分度为权重对子带分类特征向量距离线性加权，以加权子带分类特征向量距离分类法取代 LDB 常用的最近邻分类识别法。

在 LDB 中，我们选择了部分子空间构成最优基的子集 $\{U_j^*, (j, n) \in \Omega^*\}$ ，然后

将训练样本在这个子集上的可分度大于某一个阈值 t 的小波包系数保留下来, 并用这些系数重构图像, 重构的图像去掉了大量对于分类不利的冗余和噪音。对于待识别样本 x , 也在最优基的子集上进行小波包分解, 得到相应位置上的系数并重构图像。

我们不妨仍然把图像 K 类 d 维样本空间记为 $X = \{w_1, w_2, \dots, w_K\}$, 其中 $w_k = \{x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,N_k)}\}$ 是第 k 类样本集, 含有 N_k 个样本。 $x^{(k,i)} = (x_1^{(k,i)}, x_2^{(k,i)}, \dots, x_d^{(k,i)})^T$ 是重构后的第 k 类的第 i 个样本向量, x 为重构后的待识别样本。

定义 14 最近邻距离

定义 x 与 w_k 类的距离为重构后的待识别样本与该类所有训练样本的最近距离, 即:

$$D(x, w_k) = \min_{y \in w_k} \rho(x, y) \quad (4.1)$$

$$\rho(x, y) = \text{Cor}(x, y) = \frac{\|x - y\|_1}{\|x\|_1 + \|y\|_1} = \frac{\sum_i |x_i - y_i|}{\sum_i (|x_i| + |y_i|)} \quad (4.2)$$

令 $k = \arg \min_j (D(x, w_j))$, 则将 x 识别为类别 w_k 。

其它的距离定义还包括平均样本距离, 修正 Hausdroff 距离等。

4.2.1 最优基的选择

在原始 LDB 中, 用规范类能量差异 $J_{j,n}^E$ 为标准来选择最优基和特征分解系数。事实上, 能量就是数值的大小, 在图像中就是像素点灰度值的大小, 因此以 $J_{j,n}^E$ 为标准往往会选取到那些灰度比较大的像素点, 而忽略那些灰度小的像素点, 而恰恰某些灰度小的像素点仍然是人脸局部差异性特征的表示, 因此以 $J_{j,n}^E$ 为标准选取的特征并不是模式分类问题中的最佳特征。受 PCA 与 LDA 方法的启发, 最佳特征应该由某种聚类分离度指标获得。

由于 LDB 方法中需要分别计算样本空间各子带小波包分解系数的可分度和子带的可分度, 按照 LDA 的类内类间距离的思想, 我们定义 U_j^n 中第 p 个小波系数 $d_p^{(k,i,j,n)}$ 对应的可分度:

$$J_{j,n,p}^D = \frac{\sum_{k=1}^K \frac{N_k}{N} (m_p^{(k,j,n)} - m_p^{(j,n)})^2}{\sum_{k=1}^K \frac{N_k}{N} \sum_{i=1}^{N_k} \frac{1}{N_k} (d_p^{(k,i,j,n)} - m_p^{(j,n)})^2}$$

其中 $m_p^{(k,j,n)} = \frac{1}{N_k} \sum_{i=1}^{N_k} d_p^{(k,i,j,n)}$, $m_p^{(j,n)} = \frac{1}{N} \sum_{k=1}^K \sum_{i=1}^{N_k} d_p^{(k,i,j,n)}$ 。 $J_{j,n,p}^D$ 越大, 对应的小波

系数的可分性越好。我们可以取大于某阈值 t 的系数对应的可分度的和作为该子带的可分度 $J_{n,p}^D$, 但由于可能存在某个系数的可分度 $J_{j,n,p}^D$ 太大而淹没了其它系数的可分度, 因而利用简单加和的方式求得的子带可分度不够稳定。一种解决的办法是在子带上用 LDA 方法以类间散度和类内散度的迹之比作为子带的可分度:

$$J_{j,n}^S = \frac{\text{tr}(S_b^{j,n})}{\text{tr}(S_w^{j,n})} = \frac{\sum_{p=1}^{P_j} \sum_{k=1}^K \frac{N_k}{N} (m_p^{(k,j,n)} - m_p^{(j,n)})^2}{\sum_{p=1}^{P_j} \sum_{k=1}^K \frac{N_k}{N} \sum_{i=1}^{N_k} \frac{1}{N_k} (d_p^{(k,i,j,n)} - m_p^{(j,n)})^2} \quad (4.3)$$

4.2.2 可分度定义的改进——差可分度

因为小波包分解后是选择部分系数, 然而这一准则的计算却不满足线性可加性, 因此再用其作为判据进行最优基选择时不能简单地将子节点的可分度相加, 需要额外的算法解决最优基选择的组合优化问题。这无疑增加了大量的计算复杂度, 无法满足实时性的需求。同时, 在分类识别的时候数值大的特征系数将会占主要部分。有些比值虽然比较大, 但它们的绝对值都比较小, 比起那些数值大的特征应该处于次要地位, 因此我们要尽量保留数值大而有用特征系数。结合能量与散度的优缺点, 我们给出了差可分度的定义:

定义 15 差可分度

特征系数的可分度为该系数对应的平均类间距离与平均类内距离的差, 即:

$$J_{j,n,p}^{DS} = \sum_{k=1}^K \frac{N_k}{N} (m_p^{(k,j,n)} - m_p^{(j,n)})^2 - \sum_{k=1}^K \frac{N_k}{N} \sum_{i=1}^{N_k} \frac{1}{N_k} (d_p^{(k,i,j,n)} - m_p^{(j,n)})^2 \quad (4.4)$$

相应的子代的差可分度 $J_{j,n}^{DS} = \text{tr}(S_b^{j,n}) - \text{tr}(S_w^{j,n})$ 由选定的大于某个特定阈值 t 的特征系数可分度累加得到。

借鉴 Fisher 脸方法的思想, 我们可以对散度矩阵进行正交变换, 使得可分度 $J_{j,n}^{FDS} = \text{tr}(W^T S_b^{j,n} W) - \text{tr}(W^T S_w^{j,n} W)$ 最大化并作为最优基与特征系数选择的判据, 但是考虑到正交线性变换矩阵求取所带来的计算复杂度不利于实时性的实现, 因此本文还是采取差可分度作为最优基与最优系数选择的判据。

阈值 t 的选择: $t = \min \left\{ t \mid \frac{\sum_{p \in B} J_{j,n,p}^{DS}}{\sum_p J_{j,n,p}^{DS}} = 0.9, B = \{p \mid J_{j,n,p}^{DS} > t\} \right\}$ 或者更方便的取

$t = 0$, 本文采用后者, 即选择类间散度比类内散度大的特征。

§ 4.3 一种新的分类特征提取方法

我们假设分类所用的样本集和待识别样本都是一个矩形区域内的人脸。众所周知, 分类的复杂度是随着模式空间的维数数目快速增长的, 所以非常有必要进行降维, 重要的辨别信息一般都基于提取出的特征。

4.3.1 传统分类特征

在 PCA 分类方法中, 分类特征是通过计算协方差矩阵的主特征向量构成。PCA 方法的核心就是寻找正交变换矩阵 $W = (u_1, u_2, \dots, u_L) \in \mathbb{R}^{d \times L}$, $WW^T = 1$, 使得 $J(W) = \text{tr}(W^T S_i W)$ 达到最大。也就是在 $y = W^T x$ 的正交变换后使得总体矩阵 $W^T S_i W$ 的迹最大, 即各个样本分离的尽量远, 将样本的方差尽量的保留下来, 和原样本的均方误差尽量的小。实际上, $W = (u_1, u_2, \dots, u_L)$ 就是对应于矩阵 S_i 的前 L 个最大特征值对应的特征向量组成的矩阵, 即 u_i 是方程 $S_i u_i = \lambda_i u_i$ 的解。LDA 方法是利用协方差矩阵的特性对协方差矩阵进行了降维处理。传统 LDB 方法也是基于 PCA、LDA 的特征提取方法, 只是在特征提取前对图像利用小波变换进行了降维处理。

4.3.2 基于矩特性的分类特征提取

首先对整幅图片进行小波包分解。通常进行 2-3 层的小波包分解, 没有必要进行更深层次的分解, 因为在第二层之后, 图片变得非常小, 已经不可能得到有价值的信息了。假设我们进行 2 层小波包分解, 我们得到一张近似图像 (低频图像) 和 15 张细节图像 (如图 11)。因此, 人脸图像可以用 16 个小波系数矩阵来表示, 这其中包含了非常多的信息量 (相当于一整幅图像), 16 个系数矩阵包含了人脸纹理的信息。

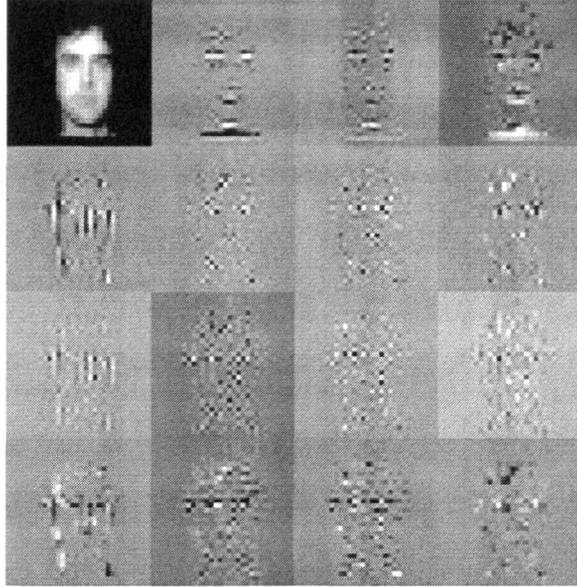


图11 人脸图像2层小波包分解示意图

由于人脸图像常常还包括人脸区域以外的信息，这些对于我们而言都是影响分类的噪音，需要在分类识别的时候尽量的去除掉。文献^[57]采用的方法是在人脸图像上经过量子化处理确定一个只包括人脸的矩形框（如图12），只考虑在人脸区域内的小波系数。对低频图像提取4个测度，人脸区域外的方差 σ_{out}^2 和均值 μ_{out} 以及人脸区域的方差 σ_{in}^2 和均值 μ_{in} 。另外15张细节图像，分别提取均值 μ_i 和方差 σ_i ($i=2, \dots, 16$)。实际上，它们的均值为0。这样，一共有19个特征值，定义如下：

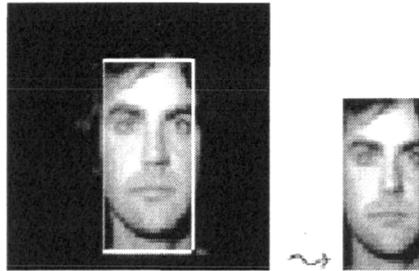


图12 选取人脸区域示意图

$$v = \bigcup_{i=0}^{16} \{\mu_i, \sigma_i^2\} \quad (4.5)$$

其中 $\forall i \geq 2, \mu_i = 0$, $\mu_0 = \mu_{out}, \sigma_0^2 = \sigma_{out}^2, \mu_1 = \mu_{in}, \sigma_1^2 = \sigma_{in}^2$ 。

由上一节的分析我们知道，并不是所有的小波子带和所有的小波分解系数都对分类识别有帮助。由于量子化处理选择矩形框准确度不高，从而会影响分类识别的准确度，同时量子化处理本身会带来额外的复杂计算量。利用上节中我们给出的差可分度的方法可以对上述方法进行改进。因此我们只考虑最优基子集 $U^* = \{U_j^n, (j, n) \in \Omega^*\}$ 上特征系数集 $D^* = \{d_p^{(k, i, j, n)} \mid d_p^{(k, i, j, n)} > t, 1 \leq k \leq K, 1 \leq i \leq N_k, (j, n) \in \Omega^*, p = 1, 2, \dots, P_j\}$ 的均值 $\mu^{(j, n)}$

和方差 $\sigma^{(j,n)}$ ，从而得到每个样本向量 $x^{(k,i)}$ 的分类特征一、二阶原点矩向量 $v^{(k,i)} = U\{u^{(j,n)}, \sigma^{(j,n)}\}$ 。

§ 4.4 分类识别方法

在模式识别问题中，最终目的是设计一个有最低误识率的、可以分类未知模式的识别系统，文章^{[59][62]}给出了一种误识率的定义。在人脸识别中，设有 K 类 d 维样本空间 $X = \{w_1, w_2, \dots, w_k\}$ ，其中 $w_k = \{x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,N_k)}\}$ 是第 k 类样本集，含有 N_k 个样本。人脸识别的目的就是对于任意一幅待识别的人脸图像 x （样本向量），需要确定 x 属于样本空间中的那一类样本集。

4.4.1 分类距离的定义

假设人连图像为一随机向量，则在概率意义上，当中有两类样本的情况下，识别错误的概率可以记为^[62]：

$$e = \frac{1}{2} \left[1 - \int |p(x|w_1)P(w_1) - p(x|w_2)P(w_2)| dx \right] \quad (4.6)$$

由方程(4.6)可以看到，误差在积分为 0 的时候最大。(4.6)式的积分可以看作两个密度函数间的概率距离。

定义 16 Bhattacharyya 距离

Bhattacharyya 距离定义为概率距离：

$$B(x) = -\ln \int [p(x|w_1)p(x|w_2)]^{\frac{1}{2}} dx \quad (4.7)$$

假设 x 属于某一类 w_i 的条件概率分布为 Gaussian 分布，则密度函数可以定义为：

$$p(x|w_i) = \left[(2\pi)^n |\Sigma_i| \right]^{-\frac{1}{2}} \times \exp \left\{ -\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right\} \quad (4.8)$$

其中， μ_i 和 Σ_i 为均值向量和方差矩阵。由(4.7)(4.8)得：

$$B = \frac{1}{4} (\mu_2 - \mu_1)^T [\Sigma_1 + \Sigma_2]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left[\frac{\frac{1}{2} (\Sigma_1 + \Sigma_2)}{\sqrt{|\Sigma_1| |\Sigma_2|}} \right] \quad (4.9)$$

由 Gaussian 分布特性，分类特征向量 v 的每一组分量 $\{\mu_j, \sigma_j^2\}$ 与其它分量都是独立的，因此我们给出如下子带分类特征向量距离定义：

定义 17 子带分类特征向量距离

任意两幅图像 x_p, x_q 对应的分类特征向量 v_p, v_q 在小波包子带 U_j^n 上的距离 D_j^n

定义为向量 v_p, v_q 在分量 $\{\mu_j^n, \sigma_j^{n2}\}$ 上的距离，即：

$$D_j^n(v_p, v_q) = \frac{1}{4} \frac{(\mu_{jp}^n - \mu_{jq}^n)^2}{(\sigma_{jp}^{n2} + \sigma_{jq}^{n2})} + \frac{1}{2} \ln \left[\frac{\frac{1}{2}(\sigma_{jp}^{n2} + \sigma_{jq}^{n2})}{\sqrt{\sigma_{jp}^{n2} \sigma_{jq}^{n2}}} \right] \quad (4.10)$$

4.4.2 分类距离的加权

由于我们在求分类特征距离时只考虑了部分可分度大的系数，因此在确定两幅图像分类特征向量的距离时，我们需要根据可分度 $J_{j,n}^{DS}$ 对子带分类特征向量距离进行加权，并以此作为两幅图像之间的距离。

令小波包子带 U_j^n 的权重为子带的相对可分度 c_j^n ：

$$c_j^n = \frac{J_{j,n}^{DS}}{\sum_{(j,n) \in \Omega} J_{j,n}^{DS}} \quad (4.11)$$

最终，我们得到两幅图像特征向量距离，也即两幅图像之间的距离：

$$D(v_p, v_q) = \sum_{(j,n) \in \Omega} c_j^n D_j^n(v_p, v_q) \quad (4.12)$$

基于图像距离的定义，定义待识别图像 x 与第 k 类样本 $w_k = \{x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,N_k)}\}$ 之间的距离为：

$$D(v_x, w_k) = \frac{1}{N_k} \sum_{i=1}^{N_k} D(v_x, v_{k,i}) \quad (4.13)$$

根据最短距离识别原则，如果存在 k^* 使得 x 与 w_{k^*} 之间的距离最小，即：

$$k^* = \arg \min_k D(v_x, w_k) \quad (4.14)$$

则将 x 识别为第 k^* 类。

4.4.3 识别阈值的确定

若存在 x 不属于训练图像所属的任何一类，那么我们就需要在识别中设计识别阈值 d_k ，用于确定待识别图像是否要归于当前训练集所属的类别。识别阈值 d_k 定义为类内最大距离或最大距离的加权值，权重根据经验或训练情况调整，即：

$$d_k = \alpha \max \{D(v_{x^{(k,l)}}), l=1, 2, \dots, N_k\} \quad \alpha \in (1, 2) \quad (4.15)$$

在识别中，若 $D(v_x, w_k) \leq d_k, k^* = \arg \min_k D(v_x, w_k)$ ，则将 x 识别为第 k^* 类；否则认为 x 不属于训练图像所属的任何一类。

第五章 人脸识别方法的实验分析与应用研究

如本文第一章所述，人脸识别算法的研究是目前的热点研究领域，并且在多个领域中得到了广泛的应用。本章针对提出的改进算法利用 Matlab 进行了实验分析，并与常用的 PCA 算法、LDB 算法进行了比较，针对实验结果进行了论证分析。在此基础上，结合本人参与的科研项目需要，对本文研究的内容进行了应用研究，设计并实现了相关系统与模块。

§ 5.1 实验方案设计

实验方案的设计取决于实验目的。本研究的目的在于全面或部分提高人脸识别的速度、正确率、稳定性和适应性，因此实验就是要针对算法改进的几个方面，通过比较检验算法在识别速度、正确率、稳定性和适应性方面有无明显的改进。因此实验方案包括的设计因子有：不同算法/改进、训练集与测试集的选取、图像类型的选取等，检验的指标有训练时间/识别时间、识别率/误识率等。

具体来讲，上述设计因子及其水平如下：

表1 人脸识别实验方案设计因子、水平表

因子	水平	
A 不同算法/改进	1. PCA (特征脸&欧式距离) (见 § 3.3)	
	2. LDB (小波包分解系数&能量可分度&特征脸&欧式距离) (见 § 4.1)	
	改进 LDB	3. 差可分度 (见定义 15)
		4. 差可分度&小波包矩特征 (见 4.3.2)
		5. 差可分度&小波包矩特征&加权距离 (见 4.4.2)
B 训练集与测试集的选取	1. 训练集 \cap 测试集=空集	
	2. 训练集 \cap 测试集=测试集	
C 图像类型的选取	1. 人脸灰度图像 (ORL 40 类*10 张)	
	2. 彩色人物正面图像(CPOEP)+人脸探测 (30 类)	
	3. 新闻视频截取图像(PON)+人脸探测 (3 类)	

在图像类型的选取上，我们选择了 ORL、CPOEP 和新闻视频中截取的人物图像 (PON) 三类图像库作为实验对象。其中 ORL 是人的头部正面灰度图像，黑色背景，同类图像具有小于 20 度的角度旋转和表情、光线、遮挡物等方面的差异，本文选取了 40 类图片，每类选取 10 张；CPOEP 图像是人的半身正面彩色图像，具有不同的复杂背景，同类图像具有表情、光线、遮挡物等差异，人物头部图像需要采用人脸探测的办法自动选取，由于人脸探测的误差，因此选取的头部图像质量较 ORL 图像有一定的差距，主要表现在框取位置偏差、具有复杂背景，本文

选取了 30 类，每类选取图片数由 5 到 25 张不等；新闻视频截选图像 PON 是从一段新闻中截取出的包含人的彩色图像，具有不同的复杂背景，同类图像具有大小、分辨率、姿态、表情、光线、角度、遮挡物等差异，较 CPOEP 效果更差，同样要采用人脸探测的办法自动选取人脸，我们重点检测两个已知人物是否在视频中出现的问題，通过实验检验算法的识别率和误识率。

来自 Olivette 研究实验室的 ORL 人脸数据库，这是一个人脸识别中常用的标准测试库。ORL 数据库共有 400 幅人脸图片，40 个人，每人 10 幅大小为 92×112 像素。这个数据库比较规范，大多数图像的光照方向和强度都差不多，表情、姿态有一些细微的变化，如图 13：



图13 ORL 图像库人物示意图

CPOEP，彩色的正面人物图像，但是光照差异较大，表情、姿态都有变化，但是差别都不是很大，如图 14：



图14 CPOEP 图像库

PON 是从新闻视频中截选的人物图像，对于特定的目标人物从不同的新闻片断进行截取，每个目标人物的图像，如图 15：



图15 视频中截取的人物图像

由表 1 可见, 按照因子水平组合数, 在不考虑其他因子变化的情况下, 我们至少需要进行 $5 \times 2 \times 3 = 30$ 次不同的实验才能够将上述情况全部覆盖。而事实上我们在实验中还要调整图片识别阈值 d_k (见公式(4.15))、最优基选择阈值 t (见 4.2.2) 等其它控制因子, 因此整个实验方案是一个复杂的、庞大的实验过程。

图 16 给出了整个试验过程的流程图, 同时也给出了识别率与误识率的计算方法。其中训练时间 T_r 为训练所用时间 t_r 与训练图片兆像素数的比值, 单位为毫秒/M, 测试时间为测试所用时间 t_t 与测试图片兆像素数的比值, 单位为毫秒/M。

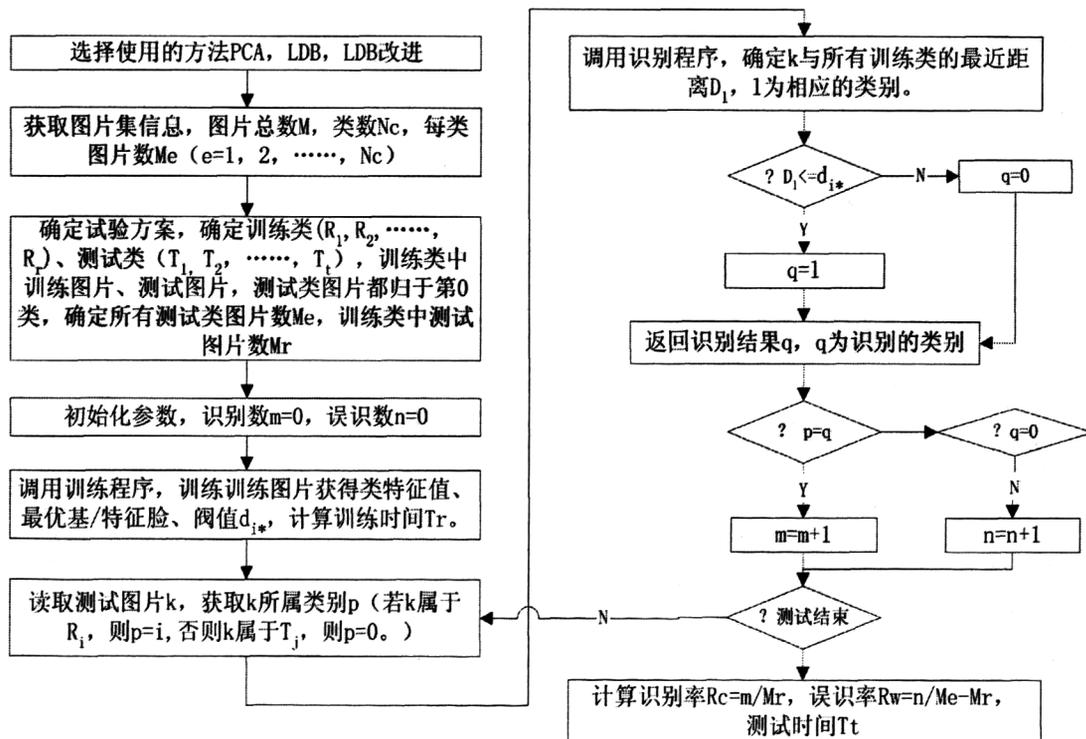


图16 人脸识别实验流程图

§ 5.2 实验结果分析

根据上一节的方案分析与设计,我们分别对三类图片应用 PCA、LDB、改进 LDB 进行了试验,记录了相应的识别率、误识率、训练时间、测试时间等数据。表 3 记录了表 1 对应的所有试验的结果。A1-A5 代表不同的人脸识别方法(A1、PCA; A2、LDB; A3、差可分度; A4、差可分度&小波包矩特征; A5、差可分度&小波包矩特征&加权距离),B1-B2 代表两种不同的训练、测试集划分方法(选取所有类中选 80%的类作为包含训练集的类,剩下 20%的类作为纯测试类。B1、随机选取每类中 50%的图片作训练集,其他图片作为测试集; B2、所有图片作为训练集,随机选取每类中 50%的图片作测试集),C1-C2 分别代表三类图片库(ORL、CPOEP 和 PON)。本文的方法 A5 是在最初的 LDB 方法的基础上,以差可分度(Difference Separability, DS)代替 LDB 算法中的平均类间类内距离,用来选择小波包分解系数和最优基;以选定子带内选定系数的一、二阶原点矩(Origin Moments, OM)作为人脸特征;以 Bhattacharrya 距离分量作为子带分类特征向量距离度量,以各子带差可分度为权重对子带分类特征向量距离线性加权,以加权子带分类特征向量距离分类法取代 LDB 常用的最近邻分类识别法。

表2 各种算法实验结果

		A1		A2		A3		A4		A5	
		B1	B2								
C1	Rc	73.6	88.32	86.112	91.264	89.056	94.208	72.128	86.848	82.432	92
	Rw	18.1	5.792	9.774	2.172	7.602	1.448	20.272	15.566	10.498	4.163
	Tr	4232.4	8168.6	126.97	253.95	126.97	253.95	0.8465	1.693	0.8465	1.693
	Tt	35.68	35.68	26.76	26.76	26.403	26.403	0.0571	0.1106	0.0571	0.1213
C2	Rc	67.712	85.376	82.432	86.848	80.224	87.584	67.712	80.224	77.28	84.64
	Rw	20.091	6.878	11.041	2.353	8.688	1.629	20.453	16.833	10.86	4.706
	Tr	4317.1	8337.9	126.97	253.95	126.97	253.95	0.8465	1.693	0.8465	1.693
	Tt	36.037	36.037	27.117	27.117	26.76	26.76	0.0571	0.1106	0.0571	0.1213
C3	Rc	55.2	72.128	61.824	68.448	66.976	72.864	58.144	70.656	65.504	69.92
	Rw	22.263	5.973	12.127	2.715	9.593	1.629	23.53	18.1	11.222	5.068
	Tr	4359.4	8422.6	126.97	253.95	126.97	253.95	0.8465	1.693	0.8465	1.693
	Tt	37.107	37.107	27.83	27.83	27.474	27.474	0.0607	0.1142	0.0607	0.10704

其中, A 表示各种不同算法及其改进,其中 A1 表示 PCA 算法, A2 表示 LDB 算法, A3 表示用差可分度改进 LDB 算法中, A4 表示用差可分度和小波包矩特征对 LDB 算法进行改进, A5 代表本文最终提出的算法---基于小波包矩特征的 LDB 人脸识别方法。B 表示两种训练集与测试集的不同选取(B1:训练集与测试集互不相同, B2:训练集中包含测试图片)。C 表示实验中图像的不同来源, C1 来自 Olivette 研究实验室的 ORL 人脸数据库。C2 是 CPOEP, 彩色的正面人物图像 PON, C3

是从新闻视频中截选的人物图像。Rc 是识别率，Rw 是误识率，Tr 为训练时间（单位为毫秒/M），Tt 为测试时间（单位为毫秒/M）。

为了能够更清晰的表现试验结果的差异，我们在表 3 中以 A1B1C1 所在的位置作为基准点，计算了各试验结果的相对值。即以 PCA 方法对 ORL 图像库进行测试，对每类中图片一半作为测试集，一半作为训练集。相应的识别率 Rc 单位为 73.6%、误识率 Rw 单位为 18.1%、训练时间 Tr 单位为 4232.45 毫秒/M，测试时间 Tt 单位为 35.68 毫秒/M，组成基向量 $B = (73.6\%, 18.1\%, 4232.45, 35.68)$ 。那么相应的基准点外的点对应的 Rc、Rw、Tr、Tt 又该位置的比率值乘以单位值。例如：要计算 A5B2C1（以方法 A5（差可分度&小波包矩特征&加权距离）对 C1（ORL 图像库）运用 B2 方法分类）进行试验得到的试验结果，首先由表得到相应的比率 $R_{A5B2C1} = \text{diag}(1.25, 0.23, 0.0004, 0.0034)$ ，则相应的试验结果：

$$ST_{A5B2C1} = BR_{A5B2C1} = (73.6\%, 18.1\%, 4232.45, 35.68) \text{diag}(1.25, 0.23, 0.0004, 0.0034) \\ = (92.0\%, 4.2\%, 1.69, 0.12) \quad (5.1)$$

表3 各种算法实验相对结果

A \ B \ C		A1		A2		A3		A4		A5	
		B1	B2	B1	B2	B1	B2	B1	B2	B1	B2
C1	Rc	1	1.20	1.17	1.24	1.21	1.28	0.98	1.18	1.12	1.25
	Rw	1	0.32	0.54	0.12	0.42	0.08	1.12	0.86	0.58	0.23
	Tr	1	1.93	0.03	0.06	0.03	0.06	0.0002	0.0004	0.0002	0.0004
	Tt	1	1	0.75	0.75	0.74	0.74	0.0016	0.0031	0.0016	0.0034
C2	Rc	0.92	1.16	1.12	1.18	1.09	1.19	0.92	1.09	1.05	1.15
	Rw	1.11	0.38	0.61	0.13	0.48	0.09	1.13	0.93	0.60	0.26
	Tr	1.02	1.97	0.03	0.06	0.03	0.06	0.0002	0.0004	0.0002	0.0004
	Tt	1.01	1.01	0.76	0.76	0.75	0.75	0.0016	0.0031	0.0016	0.0034
C3	Rc	0.75	0.98	0.84	0.93	0.91	0.99	0.79	0.96	0.89	0.95
	Rw	1.23	0.33	0.67	0.15	0.53	0.09	1.30	1.00	0.62	0.28
	Tr	1.03	1.99	0.03	0.06	0.03	0.06	0.0002	0.0004	0.0002	0.0004
	Tt	1.04	1.04	0.78	0.78	0.77	0.77	0.0017	0.0032	0.0017	0.003

行标、列标与表 2 相同。识别率 Rc 单位为 53.6%、误识率 Rw 单位为 28.1%、训练时间 Tr 单位为 4232.45 毫秒/M，测试时间 Tt 单位为 35.68 毫秒/M。

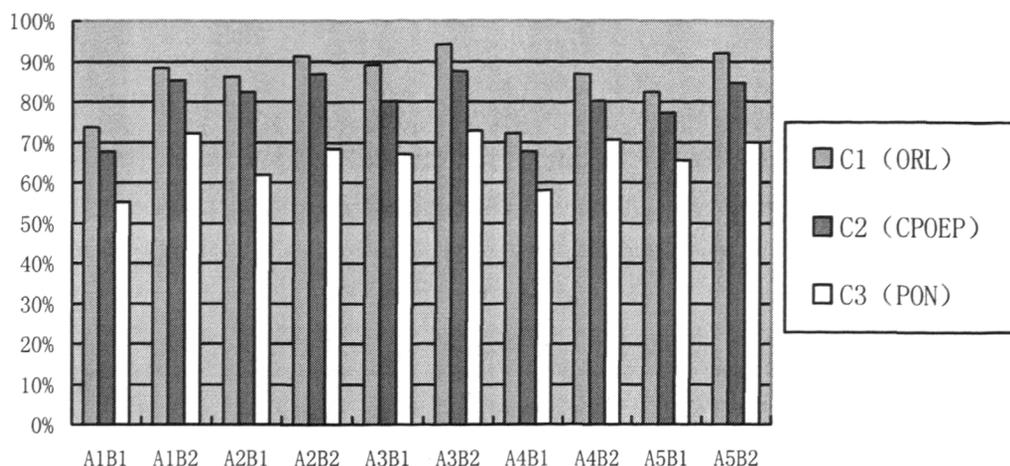


图17 识别率图

其中，A 表示各种不同算法及其改进，其中 A1 表示 PCA 算法，A2 表示 LDB 算法，A3 表示用差可分度改进 LDB 算法中，A4 表示用差可分度和小波包矩特征对 LDB 算法进行改进，A5 代表本文最终提出的算法---基于小波包矩特征的 LDB 人脸识别方法。B 表示两种训练集与测试集的不同选取（B1:训练集与测试集互不相同，B2:训练集中包含测试图片）。

由图可见本文提出的最终改进基本上能够获得与传统 LDB 方法相近的识别率，比 PCA 方法识别率高 8%左右。在改进过程中，改进可分度定义 A3 和改进特征权重 A5 都对识别率有所提高。但对提取特征的改进 A4 对识别率没有贡献，相反使得识别率有所下降。

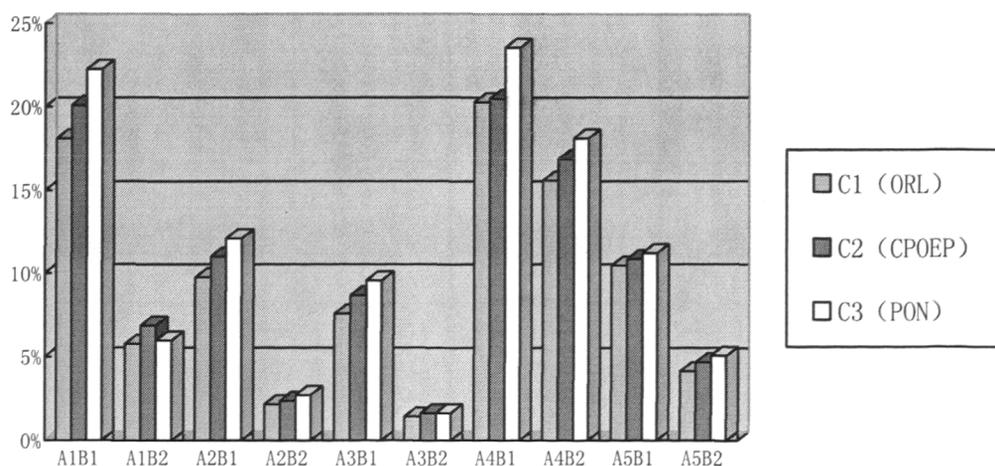


图18 误识率图

其中，A 表示各种不同算法及其改进，其中 A1 表示 PCA 算法，A2 表示 LDB 算法，A3 表示用差可分度改进 LDB 算法中，A4 表示用差可分度和小波包矩特征

对 LDB 算法进行改进, A5 代表本文最终提出的算法---基于小波包矩特征的 LDB 人脸识别方法。B 表示两种训练集与测试集的不同选取 (B1:训练集与测试集互不相同, B2:训练集中包含测试图片)。

误识率与识别率此消彼长, 在改进过程中, 改进可分度定义 A3 和改进特征权重 A5 都对误识率有所降低。但对提取特征的改进 A4 使得误识率有所提高。

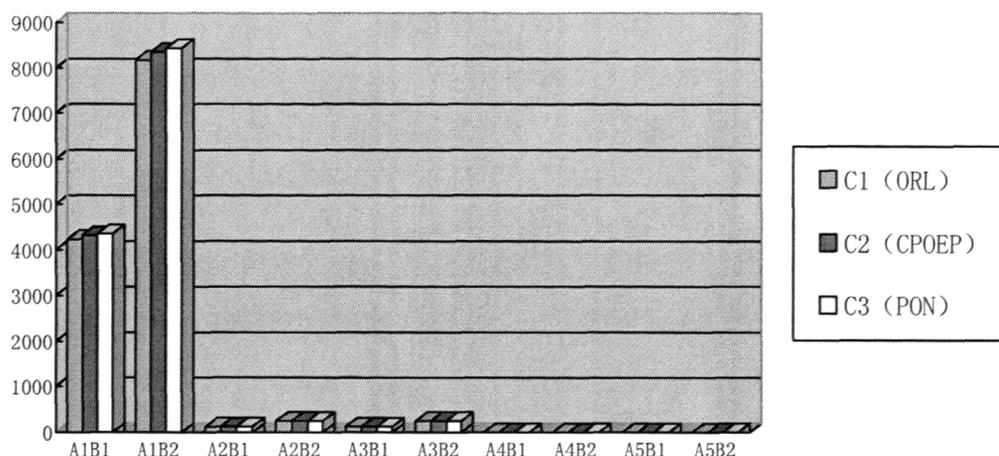


图19 训练时间图 (单位: 毫秒/M)

其中, A 表示各种不同算法及其改进, 其中 A1 表示 PCA 算法, A2 表示 LDB 算法, A3 表示用差可分度改进 LDB 算法中, A4 表示用差可分度和小波包矩特征对 LDB 算法进行改进, A5 代表本文最终提出的算法---基于小波包矩特征的 LDB 人脸识别方法。B 表示两种训练集与测试集的不同选取 (B1:训练集与测试集互不相同, B2:训练集中包含测试图片)。

由图可以看出, LDB 方法比 PCA 方法所用的训练时间明显减少。但是 LDB 与本文提出的改进方法的时间差别在图中体现的并不明显, 我们在下图中单独显示对 LDB 与改进 LDB 方法的用时。

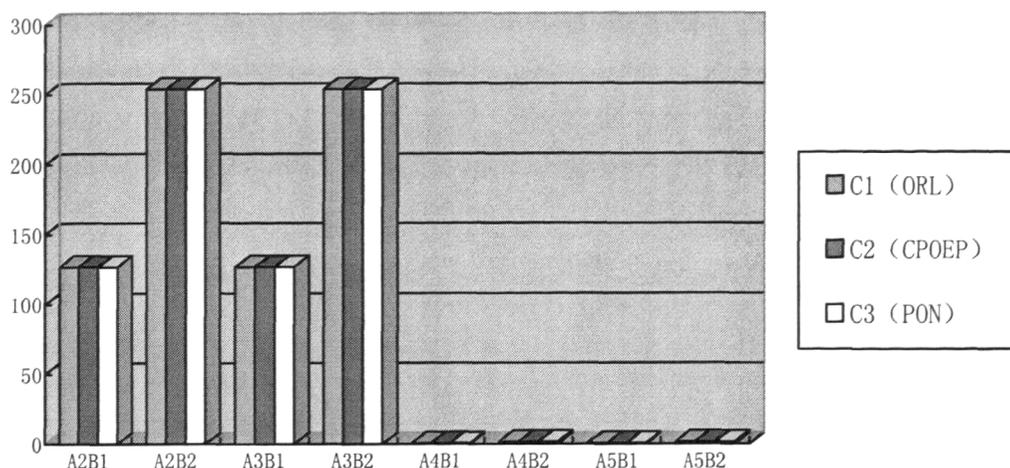


图20 部分训练时间图 (单位: 毫秒/M)

其中, A 表示各种不同算法及其改进, 其中 A1 表示 PCA 算法, A2 表示 LDB 算法, A3 表示用差可分度改进 LDB 算法中, A4 表示用差可分度和小波包矩特征对 LDB 算法进行改进, A5 代表本文最终提出的算法---基于小波包矩特征的 LDB 人脸识别方法。B 表示两种训练集与测试集的不同选取 (B1:训练集与测试集互不相同, B2:训练集中包含测试图片)。

可见, 在改进过程中, 改进可分度定义 A3 和和改进特征权重 A5 对训练时间没有多大影响, 但对提取特征的改进 A4 使得训练时间得到极大的降低, 即训练速度得到极大提高。

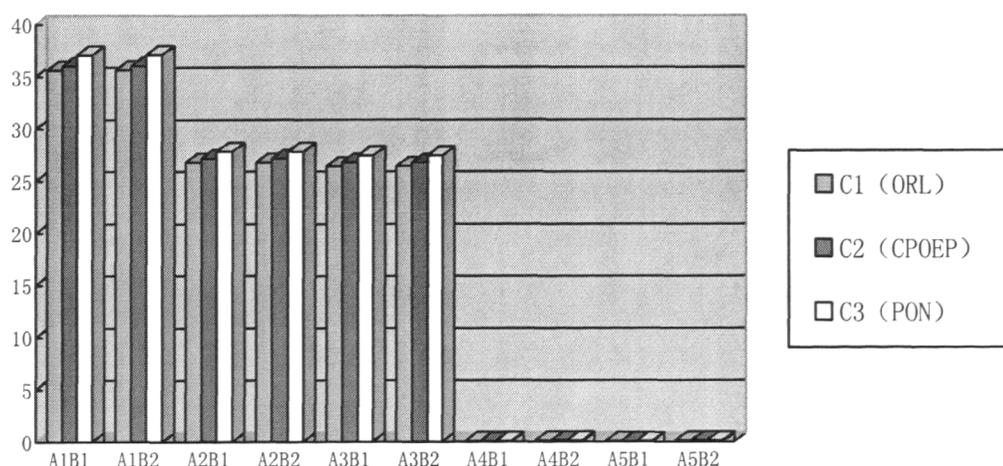


图21 测试时间图 (单位: 毫秒/M)

其中, A 表示各种不同算法及其改进, 其中 A1 表示 PCA 算法, A2 表示 LDB 算法, A3 表示用差可分度改进 LDB 算法中, A4 表示用差可分度和小波包矩特征对 LDB 算法进行改进, A5 代表本文最终提出的算法---基于小波包矩特征的 LDB

人脸识别方法。B 表示两种训练集与测试集的不同选取 (B1:训练集与测试集互不相同, B2:训练集中包含测试图片)。

由图可见测试时间的改进情况与训练时间类似。改进可分度定义 A3 和和改进特征权重 A5 对测试时间没有多大影响, 但对提取特征的改进 A4 使得测试时间得到极大的降低, 即测试速度得到极大提高。

综上所述, LDB 方法在 ORL 数据库中取得了比特征脸方法更好的效果, 表明小波包更能有效的提取多尺度下对分类有益的特征, 而本文提出的方法又进一步的提高了识别率, 并相应的大大缩短了识别时间。

§ 5.3 人脸识别在数字媒体信息处理系统中的应用

随着信息技术的发展和应用, 图像、音频、视频等数字媒体信息呈现出海量增长的趋势。如何根据用户的应用需求, 从已有海量信息中快速找到用户感兴趣的信息, 是目前信息领域迫切需要解决的重要问题之一。解决这类问题的基本方法是: 根据用户的应用需求从海量数字媒体信息中提取出相应的媒体特征, 并对这些特征进行合理有效的组织和管理, 进而根据特征实现对用户感兴趣信息的基于内容检索, 其中海量媒体特征的有效提取和组织是实现基于内容检索的关键问题之一。目前的媒体特征提取和组织, 研究较多且相对成熟的是对媒体低层特征 (如图像的颜色、形状等) 的提取和利用。而媒体特征提取和组织的新的研究前沿是, 关注媒体的高层特征 (如: 事件、人物、时间、地点等语义信息), 通过分析提取媒体语义特征, 同时通过建立不同类型媒体之间的语义关联, 实现相对快速地、高效地基于内容检索。人脸识别的技术研究及应用以数字媒体的基于内容检索为背景, 研究支持基于内容检索的媒体语义特征分析技术, 旨在通过有效地提取、组织和管理媒体的高层语义特征, 跨越图像、音频、视频等媒体数据的语义鸿沟, 实现对数字媒体的基于内容分析、检索服务和个性化的组织与表现服务, 提高军事、经济、文化等相关行业数字媒体的综合处理水平, 具有重要的应用价值和广阔的应用前景。

人脸识别系统主要围绕建立支持基于内容检索的数字媒体语义特征分析平台展开研究, 支持基于内容检索的数字媒体语义特征分析平台由两个子系统组成, 它们分别是数字媒体信息综合处理子系统、数字媒体信息服务子系统, 系统的体系结构如图 22 所示。解决的关键技术将集成在该平台的两个子系统中。



图22 支持基于内容检索的媒体特征分析平台体系结构图

数字媒体信息处理子系统是整个系统的基础，为最终的检索服务和组织与表现服务提供必须的信息。它在数字媒体信息收集工作的基础上完成媒体特征的选择、媒体特征的描述和提取、媒体数据的分类、媒体语义的提取和标注、关联分析等。数字媒体信息服务子系统是与上层应用结合最紧密的系统，也是需求与关键技术进行有机结合的关键。包括数据组织与可视化表现服务、基于内容的智能检索服务。

人脸识别主要应用于数字媒体信息处理子系统中，本系统主要研究数字媒体特征选择，数字媒体特征描述和提取，数字媒体语义提取和标注，数字媒体数据的分类等。

数字媒体的低层物理特征虽然能够一定程度上描述相应的媒体内容，但是与人们理解的媒体内容有着很大的差别，因此需要从数字媒体数据中自动提取和标注出方便人们理解的语义信息，以便于更好的利用这些数字媒体资源。这里主要研究数字媒体数据中一些基本的语义信息的提取与标注技术，主要包括：视频、音频、图像和文本中的人物的提取与标注技术研究；视频中的运动对象的提取与标注技术研究；图像场景的分类与标注技术研究。

由于在新闻视频中人们往往最关注的是一些重要人物，这里的重要人物主要指那些出现了特写镜头并且有一段讲话的人物，这些重要人物的出现往往是与一些重要的新闻事件相关联的，通过对重要人物的提取与标注能够很大的方便用户

最终查询相应的人物和事件。人脸识别则通过底层的物理特征对重要人物进行特征提取，通过与图像库中人物的特征匹配，建立关联关系，完成对重要人物的标注（如图 23）。



图23 人脸识别处理系统截图

首先，系统将从视频中截取的人物图像进行入库编辑，然后通过对待识别图像进行特征提取，来与训练样本集中的已知人脸图像进行特征匹配，最后用本文提出的加权子带分类特征向量距离计算出待识别图像所属类别，从而进行人物信息标注。

第六章 总结与展望

§ 6.1 总结

随着应用领域的迅速拓展，人脸识别研究日益成为模式识别领域研究的热门专题。出于对多媒体语义研究和应用的需要，本文对当前人脸识别的方法进行了研究，并在基于小波包进行特征提取的方法（Local Discriminant Basis）基础上，提出了一种改进的基于特定小波包分解系数矩特性的人脸识别方法（LDB_DS_OM），以差可分度（Difference Separability, DS）代替 LDB 算法中的平均类间类内距离，用来选择小波包分解系数和最优基；以选定子带内选定系数的一、二阶原点矩（Origin Moments, OM）作为人脸特征；以 Bhattacharyya 距离分量作为子带分类特征向量距离度量，以各子带差可分度为权重对子带分类特征向量距离线性加权，以加权子带分类特征向量距离分类法取代 LDB 常用的最近邻分类识别法。

实验结果显示，本文提出的最终改进方法基本上能够获得与传统 LDB 方法相近的识别率，比 PCA 方法识别率高 8% 左右。在识别效果方面，改进可分度定义和改进特征权重都对识别率有所提高。但对提取特征的改进对识别率没有贡献，相反使得识别率有所下降。在训练与识别时间方面，LDB 方法比 PCA 方法所用的训练与识别时间明显减少，而本文提出的改进可分度定义和和改进特征权重对训练与识别时间没有多大影响，但对提取特征的改进使得训练与识别时间得到极大的降低，即训练与识别速度得到极大提高。

通过实验结果分析，本文提出的人脸识别方法（LDB_DS_OM）保持较高识别效率的情况下，极大缩短了训练和识别的时间，能够满足项目提出的实时化处理需求。

在理论与算法研究的基础上，我们设计并实现了人脸探测与识别软件模块，实现了该模块与数字媒体信息处理系统的集成。

§ 6.2 改进与展望

人脸识别研究至今已经取得了一定的成果，但是还远远不能满足越来越高的社会需求。人脸识别问题对现实社会有着非常高的应用价值和广阔的发展前景，它仍将作为一项重要的研究课题内容有待我们进行研究。

对本文提出的方法而言，还有一些不足仍需要进一步的改进：

在人脸识别过程中，光照、姿势以及遮挡物这些外部的干扰是不可避免的，

对基于小波包的特征提取方法的方法影响还是很大的，因此，有必要研究鲁棒性好的人脸识别系统，能够有效的抑制外部因素的干扰。

人脸的特征提取对识别效果也是影响识别效果的一个重要条件，在今后的研究中需要对能够表达人脸的各种特征进行比较和选择，从而找到最有效的分类特征。

在数学研究领域中，对于各种距离的定义至今仍是一个热门话题，在以后的研究过程中，对分类距离需要进一步的研究，找出一种能够更好的分类距离的定义。

在海量数据分析与管理中，我们常常会面对大量混乱无序的数据束手无策。在图像文件管理中，一个棘手的问题就是面对大量的图像，如何按照图像中的人物对其进行分类。通常这些图像的标注是缺失的，不知道图像中有哪些人物，也不知道有多少人物、每个人物又在多少图片中出现，即不知道类别、类数和类内的样本数。我们提出的人脸识别技术可以有效地帮助这类问题的解决。在今后的研究中，希望通过有效的人脸识别技术帮助我们解决繁杂的分类问题。

人脸识别是一个非常复杂的问题，任何单一的算法都不能彻底解决人脸识别中的各种问题，这就需要对各种不同的算法进行研究，进行有效的组合，才有可能达到比较理想的效果。

致 谢

本文是在吴玲达教授的悉心指导下完成的。两年多来，导师言传身教，在百忙之中对我的教诲和培养使我终身难忘；导师严格的要求和不断的鞭策促使我的研究工作不断深入；同时，导师一丝不苟、严谨端正的学术态度为我树立了榜样，良好的学术研究习惯的养成不仅对我硕士阶段的学习和研究具有不可或缺的重要意义，而且必将在我今后的学习工作中大有帮助。在此，我向我的导师致敬并致以诚挚的谢意。

实验室良好的学术氛围和有广度、有深度的讨论对我的学习和研究具有很大的启发性，谢毓湘老师、文军师兄、曾璞师兄、栾悉道师兄、陈丹雯师姐也给予了我无私的帮助，同时，还要非常感谢万君博士和程光权博士，他们对我的工作提出了很多有价值的建议，为我指点迷津，帮助我排忧解难，在此表示感谢。

最后，还要感谢信息系统与管理学院研究生队的领导和同学们，感谢他们多年来对我在学习和生活上的热忱关心和热心帮助。

再一次对所有关心我、帮助我的人致以诚挚的谢意和崇高的敬礼！

参考文献

- [1] 王伟, 张佑生, 方芳. 人脸检测与识别技术综述[J]. 合肥工业大学学报(自然科学版), 2006, 29(2):158-163
- [2] 王俊艳, 苏光大, 林行刚. 人脸图像年龄估计[J]. 清华大学学报(自然科学版), 2007, 47(4)
- [3] W. Zhao, R. Chellappa, P. J. Phillips, et al. Face Recognition: A literature survey[J]. ACM Computing Surveys, 2003, 35(4): 399-459
- [4] 刘向东, 陈兆乾. 人脸识别技术的研究[J]. 计算机研究与发展, 2004, 141(17)
- [5] Demetri Terzopoulos and Keith Waters. Analysis and synthesis of facial image sequences using physical and anatomical models[J]. IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 15, No.6, June 1993, 596-579
- [6] 陈雅茜, 雷开彬. 人脸识别技术综述. 西南民族大学学报(自然科学版), 2007, 33(4):878-880
- [7] 何东风, 凌捷. 人脸识别技术综述. 危机发展. 2003, 13(12): 75-78
- [8] Chellappa R, et al. Human and Machine Recognition of Faces: A Survey[J]. Proceedings of the IEEE, 1995, 83(5): 705-740
- [9] 周杰, 卢春雨, 等. 人脸自动识别方法综述[J]. 电子学报. 2000, 28(4): 102-106
- [10] T. Kanade. Picture processing system by computer and recognition of human face [PhD Dissertation]. Kyoto: Kyoto University, 1973
- [11] R. J. Baron. Mechanisms of human facial recognition. Int J Man Machine Studies, 1981, 15:137-178
- [12] 丁嵘, 苏光大, 林行刚. 特征脸和弹性匹配人脸识别算法的比较[J]. 计算机工程与应用. 2002, 38(7):1-2
- [13] Lanitis A, Taylor C J, Cootes T F. Automatic Interpretation and Coding of Face Images Using Flexible Models[J]. IEEE Trans on PAMI, 1997, 19(7):743-756
- [14] H.Chan and W.W.Bledsoe. A man-machine facial recognition system: some preliminary results. Technical report. Panoramic Research Inc., Cal, 1965
- [15] 何家忠, 杜明辉. 结合小波低频子带的主成分分析方法[J]. 华南理工大学学报(自然科学版). 2007, 35(1): 44-48
- [16] 祝磊, 朱善安. 人脸识别的一种新的特征提取方法[J]. 光电工程. 2007, 34(6):122-125

- [17] 张志伟, 杨帆, 夏克文, 杨瑞霞. 基于小波变换和 NMF 的人脸识别方法的研究. 计算机工程, 2007, 33(6):176-178
- [18] 闫荣华, 彭进业, 李岩, 谢明华, 温文龙. 基于小波域奇异值分解的人脸识别方法[J]. 计算机工程. 2007,33(4):
- [19] 张丽丹, 黄凤岗, 李先伟. 基于监督式 ICA 的人脸识别[J]. 哈尔滨工程大学学报. 2006, 27(5)
- [20] 杨光正, 黄熙涛. 镶嵌图在人面定位中的应用[J]. 模式识别与人工智能. 1996, 9(3):213-220
- [21] 彭辉, 张长水. 基于 K-L 变换的人脸自动识别方法[J]. 清华大学学报(自然科学版). 1997, 37(3): 67-70
- [22] 荆晓远, 杨静宇. 基于小波变换和群体决策方法识别人脸图像[J]. 计算机研究与发展. 1999, 36 (1) :72-76
- [23] 杨键, 杨静宇, 叶辉. 线性鉴别分析的理论研究及其应用[J]. 自动化学报. 2003, 29(4):482-493
- [24] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720
- [25] A. Yuille, D. Cohn, and P. Hallinen. Feature extraction from faces using deformable temples[R]. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition, 1989, 104-109
- [26] M. Lades, J.C. Vorbruggen, et al. Distortion invariant object recognition in dynamic link architecture[J]. IEEE Trans. Computer, Vol. 42, 1993, 300-311
- [27] Bai-Bo Z, Chang-Shui Z. Lower bounds estimation to KL transform in face representation and recognition[A]. Proc of 2002 International Conference on Machine Learning and Cybernetics[C]. 2002, 314-1318
- [28] Kirby M, Sirovich L. Application of the Karhunen Loeve Procedure for the Characterization of Human Faces. IEEE Trans on Pattern Analysis and Machine Intelligence, 1990, 12(1):103-108
- [29] Turk M, Pentland A. Eigenfaces for Recognition. Journal of Cognitive Neuroscience, 1991,3(1):71-86
- [30] 边肇祺, 张学工, 等. 模式识别(第2版)[M]. 北京: 清华大学出版社, 2000
- [31] 高丽萍, 郭义民, 倪重匡. 一种改进的特征脸方法[J]. 计算机应用与软件, 2002, 44-47
- [32] Jun Zhang, Yong Yan, Martin Lades. Face Recognition: Eigenface, Elastic

- Mathcing and Neural Nets[J]. Proceeding of IEEE, 1997, 85(9):1422-1435
- [33] 李刚, 高政. 人脸自动识别方法综述[J]. 计算机应用研究, 2003, 20(8):4-9
- [34] Ping W, Wei-Qing X, Xiao-Quan W. A design and complement for face recognition[A]. Proc of 2004 International Conference on Machine Learning and Cybernetics[C]. 2004, 3666-3669
- [35] 蒋宗礼. 人工神经网络导论[M]. 北京: 高等教育出版社, 2002
- [36] Fleming M K, Cottrell G W. Categorization of faces using unsupervised feature extraction. Proceedings of International Joint Conference on Neural Networks, San Diego CA, 1990,II:65-70
- [37] Shang-Hung Lin, Sun-Yuan Kung, Long-Ji Lin. Face recognition/detection by probabilistic decision-based neural network. Neural Networks. IEEE Trans. On Issue:1, Vol. 8, Jan 1997, 114-132
- [38] Ricanek, K., Jr., Leiby, G.L., Haywood, K.. Hopfield like networks for pattern recognition with application to face recognition. Neural Networks, 1999. IJCNN'99. International Joint Conference on, 10-16 July 1999, vol. 5:3265-3269
- [39] Rowley R , Baluja S. Neural network-based face detection. IEEE Trans on Pattern Analysis and Machine Intelligence, 1998, 20(1): 23-38.
- [40] 崔锦泰 著, 程正兴 译. 小波分析导论. 西安: 西安交通大学出版社, 1995
- [41] Ste'phane Mallat. A Wavelet Tour of Signal Processing. Academic Press, 1998
- [42] Daubechies I.. Ten lectures on wavelet. In: CBMS. Philadelphia, PA: SIAM,1994
- [43] A. Boggess, F. J. Narcowich. A First Course in Wavelets with Fourier Analysis. Publishing House of Electronics Industry, 2002
- [44] Klivansky, S. K., Mukherjee, A. and Song, C. "On Long Range Dependence in NSFNET Traffic" Technical Report GIT-CC-94/61, Georgia Institute of Technology, Atlanta, GA 30332, USA, December 1994
- [45] 成礼智, 王红霞, 罗永 编著. 小波的理论与应用. 北京: 科学出版社, 2004
- [46] S. Eickeler, S. Muller, G Rigoll. Recognition of JPEG compressed face images based on statistical methods[J]. Image and Vision Computing, 2001, 18(4):279-287
- [47] Hong ZQ. Algebraic Feature Extraction of Image for Recognition. Pattern Recognition[J]. 1991, 24(3): 211-219
- [48] 苏宏涛, 张艳宁, 王晶, 等. 光照变化条件下的人脸识别研究[J]. 西北工业大学学报(自然科学版), 2004, 22(4):426-430
- [49] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection[J]. IEEE Trans on Pattern

Analysis and Machine Intelligence, 1997, 19(7): 711-720

- [50] Jolliffe I T. Principal Component Analysis. New York, USA, Springer-Verlag, 1986
- [51] Draper B A, Baek K, Bartlett M S, Beveridge J R. Recognizing Faces with PCA and ICA. Computer Vision and Image Understanding, 2003, 91(1-2): 115-137
- [52] Comon P. Independent Component Analysis --A New Concept of Signal Processing, 1994, 36(3):287-314
- [53] Moghaddam B, Jebara T, Pentland A. Bayesian Face Recognition. Pattern Recognition, 2000, 33(11): 1771-1782
- [54] Liwei W, Xiao W, Jufu F. Intrapersonal subspace analysis with application to adaptive Bayesian face recognition[J]. Pattern Recognition, 2005, 38(4):617-621
- [55] Shuicheng Y, Xiaofei H, Yuxiao H, et al. Bayesian shapelocalization for face recognition using global and local textures[J]. IEEE Trans on Circuits and Systems for Video Technology, 2004, 14(1):102-113
- [56] G C Feng, P C Yuen, D Q Dai. Human Face Recognition Using PCA on Wavelet Subband[J]. J of Electronic Imaging, 2000, 9(2)
- [57] Christophe Garcia, Giorgos Zikos, Giorgos Tziritas, A Wavelet-based Framework for Face Recognition, ECCV1998
- [58] N. Saito, R. R. Coifman, F. B. Geshwind, F. Warner. Discriminant feature extraction using empirical probability density estimation and a local basis library. Pattern Recognition, 35(2002): 2841-2852
- [59] N. Saito. Local feature extraction and its application using a library of bases, Ph.D. Thesis, Department of Mathematics, Yale University, December 1994
- [60] N. Saito, R. R. Coifman. Local discriminant bases, in: A. F. Laine, M. A. Unser(Eds.), Wavelet Applications in Signal and Image Processing II, Proceedings of the SPIE 2303, San Diego, CA, July 1994: 2-14
- [61] N. Saito, R. R. Coifman. Local discriminant bases and their applications, J. Math. Imaging Vision 4 (1995): 337-358
- [62] Y. Fu. Handbook of Pattern Recognition and Image Processing. Academic Press, 1986

作者在硕士学习期间取得的学术成果

1. 周蕾, 吴玲达, 程光权, 万君, 一种基于矩特性的改进 LDB 人脸识别方法, 第四届智能 CAD 与数字娱乐学术会议, 2007 年 9 月, 已发表
2. 周蕾, 吴玲达, 程光权, 万君, 一种改进的 LDB 人脸识别方法, 《计算机工程与应用》, 已录用, 2008 年 3 月正式刊出