

论文题目：基于块内相关性的 2DPCA 人脸识别方法

专    业：通信与信息系统

硕  士  生：张成言

指导教师：马争鸣 教授

## 摘    要

主元分析（PCA）方法是一种常用的人脸识别方法。由于 PCA 方法需要将图像从二维矩阵转化为一维向量，构造出巨大的协方差矩阵，并求解其特征值与特征向量，长期以来一直受到高计算复杂性的困扰。近年来出现的二维主元分析（2DPCA）人脸识别方法，可以利用二维图像矩阵直接构造出协方差矩阵，不但避免了庞大的运算量，运算时间大约是 PCA 方法的  $1/4 \sim 1/20$ ，而且其识别率明显高于 PCA 方法，一般有大于 5% 的显著提升。

通过对 2DPCA 的数学分析可以发现，2DPCA 方法是利用图像矩阵同一行像素之间的相关性来构造协方差矩阵的，其实质是一种基于行分块 PCA 方法。对于人脸图像来说，图像矩阵同一行像素之间的相关性没有同一个局部块像素之间的相关性强，而且人脸图像的一个局部块通常表示一个完整的语义，如鼻子、眼睛、嘴巴等等。然而，2DPCA 方法完全抛弃了图像矩阵行与行之间像素的相关性信息，割裂了人脸局部区域内的各个像素之间的相关性，因而难以表征人脸的局部特征。针对 2DPCA 方法的缺陷，本文提出了两种改进方法：基于结构分块的 2DPCA 方法和基于块内相关性的 2DPCA 方法。

基于结构分块的 2DPCA 方法，是针对不完整人脸图像的识别问题而提出的，其基本思想是：将人脸图像分割为互不重叠的图像块，用 2DPCA 方法分别计算各个图像块的主元，这些主元反映了人脸的结构特征，本文中称之为结构主元。这种方法结合了统计模式识别与结构模式识别的优点，通过调整各图像块的加权系数，可应用于各种不完整人脸图像的识别。实验结果表明，在选取较少主元数目的情况下，基于结构分块的 2DPCA 方法的人脸识别率与 2DPCA 方法相比大约有 3%~7% 的提升。

基于块内相关性的 2DPCA 方法的基本思想是：对人脸图像进行分块，然后把图像块内的像素转化为行向量，再把这些行向量按图像块的顺序排列成新的图像

矩阵，最后对这些新的图像矩阵采用 2DPCA 方法进行人脸识别。这种方法不但保留了 2DPCA 方法低计算复杂度的优点，而且能够充分地利用图像行与列像素之间的相关性信息，克服了 2DPCA 方法的不足。实验结果表明，在选取较少主元数目的情况下，基于块内相关性的 2DPCA 方法的人脸识别率与 2DPCA 方法相比大约有 2%~4% 的提升，且达到相同识别率所需的主元数目要少于 2DPCA 方法。

**关键字：**模式识别 人脸识别 特征提取 二维主元分析 块内相关性

Title: 2DPCA Face Recognition Based on Block-Correlation  
Major: Communication and Information Systems  
Name: Chengyan Zhang  
Supervisor: Professor Zhengming Ma

## ABSTRACT

Principal Component Analysis (PCA) is a commonly used method for face recognition. PCA is based on vector-space model; 2D image matrix is transformed into 1D long vector, and then represented as a point in a high-dimensional vector space. This leads to the Curse of Dimensionality dilemma of PCA. A new technique called two-dimensional principal component analysis (2DPCA) was proposed for human face representation and recognition. Different from classical PCA, 2DPCA is based on 2D matrix rather than 1D vector. 2DPCA is computationally more efficient, its time consuming is about 1/4 ~ 1/20 of PCA. 2DPCA can achieve better performance in face recognition, usually has a significant improvement of 5% as compared to PCA.

Through the mathematic analysis of 2DPCA, we can find that 2DPCA constructs the covariance matrix using only the correlation information among the rows of images, it is essentially the PCA performed on the row vectors of all the images. For facial images, the correlation information among the rows of images is less significant than the correlation information among the blocks of images. Furthermore, a block from a facial image commonly indicates a specific semantics, such as nose, eye, mouth and so forth. However, 2DPCA discards all correlation information among the columns of images, thus it is difficult to characterize the local features of human face. To overcome the shortcomings of 2DPCA, two new methods are developed in this essay, namely Frame 2DPCA and Block- Correlation 2DPCA. The experimental results indicated that the recognition rate of Frame 2DPCA has a significant improvement of 3% ~ 7% as compared to that of 2DPCA when the numbers of principal components is small.

Frame 2DPCA method is proposed for solving the problem of incomplete facial image recognition. This method first divide facial image into several nonoverlapping blocks, then calculate principal components of each block using 2DPCA. These principal components, which called frame principal components, reflect the structure characteristic of human face. Frame 2DPCA combines the advantages of statistical

pattern recognition and frame pattern recognition, can be applied to various incomplete facial images by adjusting the weight coefficients.

The process of Block-Correlation 2DPCA method is as follow: Firstly, a facial image is divided into nonoverlapping blocks, with the restriction that each block contains the same number of pixels of a row of the original image. Secondly, each block is rearranged to a row vector by concatenate all the rows in the block. Thirdly, a new image matrix is form by all row vectors that generated in previous step. According to the steps mentioned above, all facial images will be transformed to new image matrices, and then the standard 2DPCA is performed. Block-Correlation 2DPCA retains the low computational complexity of 2DPCA, and takes full advantage of the correlation information among the rows and columns of images. The experimental results indicated that the recognition rate of Block-Correlation 2DPCA has an improvement of 2% ~ 4% as compared to that of 2DPCA when the numbers of principal components is small. Moreover, Block-Correlation 2DPCA can achieve the same recognition rate as 2DPCA while using fewer principal components.

**Key Words:** Pattern Recognition, Face Recognition, Characteristic Extraction, 2D Principal Component Analysis, Block-Correlation

# 第1章 绪论

## 1.1 人脸识别概述

人脸识别是计算机视觉与模式识别领域的一个重要研究课题，近几十年来已经成为一个非常活跃的研究方向。所谓人脸识别技术，就是指利用计算机技术分析包含人脸的静态图像或视频序列，从中提取人脸的个性化特征，并以此自动识别别人的身份的一门技术。尽管人类早在婴幼儿时期就已经具有了辨识人脸的能力，但构建一个自动的、高识别率的计算机人脸识别系统却并非易事，有许多困难仍未得到解决，这些困难主要表现在光线、姿态、年龄、遮挡物等方面。人脸识别研究跨越了图像处理、模式识别、计算机视觉、人工智能以及生理学、心理学、认知科学等研究领域，并与其他基于生物特征的身份鉴别方法以及计算机人机感知交互领域都有着密切的联系。

人脸识别研究从二十世纪六七十年代起就引起了诸多学科领域研究者的浓厚兴趣，进入二十世纪九十年代后，随着各行业对人脸识别系统的迫切需求，人脸识别研究再次成为热门课题。当前世界各国有许多研究机构在从事这方面的研究，这些研究受到军方、警方以及大公司的高度重视和资助，美国军方还专门组织了人脸识别竞赛以促进人脸识别研究的发展<sup>[1]</sup>。经过几十年的研究，人脸识别已经成为图像分析与图像理解领域最成功的应用之一，研究人员提出了许多识别方法，建成了一些实验系统，也有不少成功的人脸识别商业软件已经投入市场。

人脸识别技术因其无需用户过多参与、非接触式的数据采集方式、对用户无任何损害、便于隐藏等优点而普遍为人们所看好。人脸识别技术在国家重要机关、社会安防等各种领域都具有广泛的应用，典型应用如：

### （1）犯罪刑侦领域中的应用

在刑侦工作中，对罪犯的抓捕是至关重要的一环。如果掌握了犯罪嫌疑人的脸图像，则可应用人脸识别技术，在公安机关海量的人脸数据库里找出与其最相像的人，从而确定其身份。如果此项搜索工作由人工完成，则效率低且容易出错，由计算机来完成则不会遇到此问题。同时，应用人脸识别技术还可依据犯罪嫌疑人的人脸图像，对机场、车站、港口等重要交通场所进行监控，从而大大促

进了罪犯抓捕工作的开展。

### (2) 安全防范领域中的应用

社会上有许多重要的部门，如军事、金融、保密等部门都需要对进出人员进行身份识别，以防止信息泄漏和不法现象的发生。与指纹、掌形、虹膜和语音等生物识别技术相比，人脸识别具有直接、方便和友好等特点。当前计算机系统的安全管理也倍受重视，通常使用的由字符和数字组成的口令，可能会被遗忘或者被破解，如果把人脸作为身份标识，其效率与安全性都会得到很大的提高。

### (3) 公共事业领域中的应用

在现代社会，许多领域都需要对人的身份进行验证，如银行、保险、交通等公共事业部门。采用传统的密码、IC 卡等手段和技术对人进行身份验证具有安全性差、易丢失、易伪造等缺点，而采用人脸识别技术进行身份验证则能够很好地克服传统身份验证手段和技术的缺陷。

除了以上这几类应用外，人脸识别技术在视频会议、机器人的智能化研究以及医学等各个方面都有着广阔的应用前景。

## 1.2 人脸识别的研究与发展

人脸识别过程一般可分为三个步骤：从场景中检测分割出人脸、提取人脸识别特征、匹配和识别人脸，因此人脸识别的研究主要集中在以下三个方面：

### (1) 人脸检测与定位

即从各种复杂的背景中检测出人脸的存在，并确定其在图像中的位置，最后将其从背景中分割出来。这一任务主要受光照、噪声、人脸在图像中的位置、旋转角度、尺度缩放以及各种遮挡的影响。

### (2) 人脸特征提取

即采用某种表示方法来表征检测出的人脸和数据库中的已知人脸，其关键是确定表征人脸的描述方式。具体的特征形式随识别方法的不同而不同，如：基于几何特征的识别方法主要是提取特征点，然后构造特征向量；基于统计的特征脸方法是利用图像协方差矩阵的特征向量来构造特征；隐马尔可夫方法是对多个样本图像的空间序列训练出一个隐马尔科夫模型，它的参数就是特征值；而大部分神经网络方法则是直接用归一化后的灰度图像作为输入，网络的输出就是输出结

果，没有专门的特征提取过程。

在特征提取之前，通常需要对人脸图像进行预处理，一般包括几何归一化和光线归一化，前者是指根据人脸定位结果将图像中的人脸变化到同一位置和大小，后者则是指对图像进行光照补偿等处理，以克服光照变化的影响。

### (3) 人脸识别

数据库里预先存放了已知的人脸图像或其对应的特征形式，人脸识别的目的就是将待识别图像或特征与数据库里的进行匹配，其核心是选择适当的人脸表征方式和匹配策略。人脸识别的任务主要有两类：一是身份鉴定，即判断输入图像是否为数据库中的人，是一对多的匹配过程；二是身份确认，即验证某个人的身份是否属实，是一对一的匹配过程。

人脸识别技术的应用包括了从静态的、受控背景的照片到动态的、非受控背景的视频等各个方面，本论文研究的是人脸识别在静态图像中的应用，即是给定一组静态的人脸图像，利用已有的人脸数据库来对人脸图像进行身份鉴定。

#### 1.2.1 人脸的检测与定位

人脸检测与人脸定位是两个密切相关而又有所区别的概念。人脸检测的目的是检测输入的未知图像中是否包含人脸，而人脸定位则是确定输入图像中人脸的位置和大小，并将其从图像背景中分割出来。本文重点在于研究特征提取和识别，因此将人脸检测与人脸定位结合起来讨论。

人脸检测与定位是一项非常复杂的工作，这是由于光线、角度、姿态、表情、遮挡物等各种因素，以及可能出现的噪声和干扰，导致即便是同一个人的人脸图像也会有相当大的差别，这使得人脸检测与定位工作变得相当困难。作为人脸识别的基础，人脸的检测与定位已经成为人脸识别研究的一个重要方面。

人脸检测的基本思想是：用统计或知识的方法对人脸进行建模，比较所有待检测区域与人脸模型的匹配度，从而得到可能存在人脸的区域。人脸检测的方法大致可分为基于统计和基于知识两类，前者将人脸图像视为一个高维向量，将人脸检测问题转化为高维空间中分布信息的检测问题，后者则利用关于人脸的知识建立若干规则，从而将人脸检测问题转化为假设 / 验证问题。

下面对常见的各类人脸检测方法作一下介绍：

##### 一、基于统计的人脸检测方法

### (1) 事例学习方法

将人脸检测看作区分非人脸样本与人脸样本的两类模式分类问题，用人脸样本集和非人脸样本集进行学习训练以产生分类器。目前国际上普遍采用人工神经网络方法<sup>[2-7]</sup>。Juell 等人<sup>[2]</sup>利用了四个子网构成的多层网络结构进行检测，三个子层神经网络分别判别口、眼、鼻，父层神经网络对子层网络判别的结果进行综合。Roudey<sup>[5]</sup>建立的三层误差反传网络以多种方式对人脸区域进行了划分和组合，每次在大小为  $20 \times 0$  的窗口区域内检测双眼距离为  $12 \times 2$  的人脸，通过将图像尺度按  $1:1.2$  的比例逐级缩放，并在各个尺度上用  $20 \times 0$  的窗口扫描整个图像，保证了图像中不同尺度的人脸区域总在某个缩放比例下基本完整地落入检测窗口，从而解决了图像中人脸尺度不一的问题。

### (2) 子空间方法

Turk 与 Pentland 将 K-L 变换引入了人脸检测中<sup>[8][9]</sup>，在人脸识别中利用的是主元子空间（特征脸空间），而人脸检测利用的则是次主元子空间（即特征脸空间的补空间）。用待检测区域在次主元子空间上的投影能量，即待检测区域到特征脸子空间的距离作为检测统计量，距离越小表明越像人脸。

## 二、基于知识建模的人脸检测方法

### (1) 器官分布规则

虽然人脸在外观上变化很大，但都遵循一些普遍适用的规则，检测图像中是否有人脸，即测试该图像中是否存在满足这些规则的图像块。这种方法一般有两种思路：一是“从上而下”，其中最简单有效的是 Yang 等人提出的 Mosaic 方法<sup>[11]</sup>，它给出了基于人脸区域灰度分布的规则，依据这些规则对图像从粗分辨率到高分辨率进行筛选，以样本满足这些规则的程度作为检测的判据。二是“从下至上”，先直接检测几个器官可能分布的位置，然后将这些位置点分别组合，用器官分布的几何关系准则进行筛选，找到可能存在的人脸<sup>[12]</sup>。

### (2) 轮廓规则

人脸的轮廓可以简单地看成一个近似椭圆，而人脸检测可以通过椭圆检测来完成。Baddeley 等人<sup>[13]</sup>提出认知模型方法，将人脸建模为两条直线（左右两侧面颊）和上下两个弧（头部和下巴），通过修正 Hough 变换来检测直线和弧。Tankus 等人<sup>[14]</sup>利用凸检测的方法进行人脸检测。

### (3) 颜色与纹理规则

由于相同民族的人的面部肤色在颜色空间中的分布相对比较集中，根据颜色信息在一定程度上可以将人脸同大部分背景区分开来。Lee 等人<sup>[15]</sup>设计了肤色模型以表征人脸颜色，利用感知模型进行复杂背景下人脸及器官的检测与分割。Dai 等人<sup>[16]</sup>把人脸图像看作灰度分布具有特殊结构特征的区域，利用 SGLD（空间灰度共生矩阵）的纹理信息作为特征进行低分辨率的人脸检测。

### 1.2.2 人脸的特征提取与识别

Brunelli 和 Poggio<sup>[18]</sup>认为，人脸识别方法可以分为两大类：基于几何特征（Geometric Feature-Based）的方法和基于模板匹配（Template Matching-Based）的方法。基于几何特征方法的思想是首先检测出嘴巴、鼻子、眉毛、眼睛等脸部主要部件的位置和大小，然后利用这些部件的总体几何分布关系以及相互之间的参数比例来识别人脸。基于模板匹配方法则是利用模板和整个人脸图像的像素值之间的自相关性进行识别，这类方法也称作基于表象（Appearance-Based）的方法。

在 Brunelli 和 Poggio 以后，又出现了许多新的人脸识别方法，下面介绍一下常见的各类人脸识别方法：

#### （1）基于几何特征的方法

几何特征也称为直观特征，它是以人脸器官的形状和几何关系为基础的向量，其分量通常包括人脸指定两点间的欧氏距离、曲率、角度等。将人脸用一组几何特征向量表示，利用模式识别中层次聚类的思想，设计分类器以达到识别的目的。

采用几何特征进行人脸识别，一般是通过提取眼睛、嘴巴、鼻子等脸部主要器官的位置与几何形状作为分类特征的，如 Brunelli 和 Poggio<sup>[18]</sup>通过计算鼻子的宽度和长度、嘴巴位置和下巴形状等进行人脸识别。

侧影（Profile）识别<sup>[19]</sup>也是早期基于几何特征人脸识别中的一种重要方法，其基本原理是从人脸的侧影轮廓线上提取特征点，将侧影转化为轮廓曲线，从中提取基准点，根据这些点之间的几何特征来进行识别。

基于几何特征的方法的特点是非常直观，识别速度快，内存要求较少，提取的特征在一定程度上对光照变化不太敏感，但当人脸具有一定的表情或者姿态变化时，特征提取不精确，而且由于忽略了整个图像的很多细节信息，人脸识别率较低，所以近年来已经很少有新的发展。

#### （2）基于数学模型的方法

隐马尔可夫模型（Hidden Markov Model, HMM）是一种常用的数学模型，基于 HMM 的方法首先被用于声音识别等身份识别上，之后被 Nefian 和 Hayes 引入到人脸识别领域<sup>[20]</sup>。HMM 是用于描述信号统计特征的一组统计模型，用马尔科夫链来模拟信号统计特性的变化，而这种变化又是间接通过观察序列来描述的，因此隐马尔可夫过程是一个双重的随机过程。在 HMM 中结点表示状态，有向边表示状态之间的转移，一个状态可以具有特征空间中的任意特征，对同一特征，不同形态表现出这一特征的概率不同。在人脸识别过程，Nefian 首先采用二维离散余弦变换（Discrete Cosine Transform, DCT）抽取人脸特征，得到观察向量，构建 HMM 人脸模型，然后用 EM（Expectation Maximization）算法<sup>[21]</sup>训练。利用该模型就可以算出每个待识别人脸观察向量的概率，从而完成识别。HMM 方法的鲁棒性较好，对表情、姿态变化不太敏感，识别率高。

主动形状模型（Active Shape Model, ASM）方法由 Cootes 等人<sup>[22]</sup>提出，Cootes 对形状和局部灰度表象建模，用建立的 ASM 在新的图像中定位易变的物体。后来，Lanitis 将其应用于解释人脸图像，在使用 ASM 找出人脸的形状后，将人脸切割并归一到统一的框架，对这个与形状无关的人脸采用亮度模型来进行解释和识别。

主动表象模型（Active Appearance Model, AAM）方法可以看成是对 ASM 方法的进一步扩展<sup>[23][24]</sup>，是一种通用的非线性图像编码模式，通过变形处理将通用人脸模型与输入图像进行匹配，并将控制参数作为分类的特征向量。

### （3）基于子空间的方法

Kirby 等人<sup>[25]</sup>和 Truk 等人<sup>[8]</sup>首次把主元分析的子空间思想引入到了人脸识别中，并获得了较大的成功<sup>[26]</sup>。随后子空间分析方法就引起了人们的广泛注意，从而成为了当前人脸识别的主流方法之一。子空间分析的思想就是根据一定的性能目标来寻找一线性或非线性的空间变换，把原始信号数据压缩到一个低维子空间，使数据在子空间中的分布更加紧凑，为数据的更好描述提供了手段，另外计算的复杂度也得到了大大降低。

#### ① 主元分析（PCA）方法

主元分析的思想来源于 K-L 变换，目的是通过线性变换寻找一组最优的标准正交向量基，通过它们的线性组合来重建原始样本，并使得重建后的新样本与原始样本之间的均方误差最小。PCA 的两大优点是：一、消除了模式样本之间的相关性；二、实现了模式样本的维数压缩。

PCA 方法通常采用训练样本协方差矩阵的特征向量系作为展开基（即 K-L 坐标轴），若干个最大特征值所对应的特征向量被称为主元（或主成分），模式样本在这些主元上线性投影后，所得的投影系数就是主元特征。原模式样本可以表示为这些主元与投影系数乘积的代数和，这种方法已被证明在最小均方差意义上是最优的。识别时，只需将测试图像投影到主元子空间上，得到一组投影系数。再将测试图像的投影系数与训练图像的投影系数作比较，以确定哪一幅训练图像与测试图像最接近。

特征脸 (Eigenface) 方法由 Turk 和 Pentland 提出<sup>[8]</sup>，是 PCA 方法在人脸识别上早期的成功应用之一。Moghaddam 等人<sup>[27]</sup>根据主元分解后的两个正交的子空间的特征，提出基于主元分析的贝叶斯框架。Liu 等人<sup>[28][29]</sup>提出结合主元分析的概率推理模型，假设在主元子空间中每类的类条件概率服从正态分布。文献[30]中提出了 Eigenhill 的方法实际上是先对原始图像采用边缘滤波和局部平滑的预处理，再利用主元分析提取特征，实验表明可以更有效地处理人脸图像中的丰富表情变化。

为了减少 PCA 方法的运算量，Yang 和 Zhang 对 PCA 方法进行了分析，提出了一种称作二维主元分析 (2DPCA) 的人脸识别方法<sup>[31][32]</sup>。2DPCA 方法可以直接在二维图像矩阵上进行处理，而不需要事先将图像矩阵转换为一维向量，从而避免了庞大的计算量，解决了 PCA 方法在计算复杂性上的困扰。

## ②线性鉴别分析 (LDA) 方法

尽管 PCA 方法在最小均方差意义上给出了模式样本的最优表示，但由于它是以所有样本的最优重建为目的，因此对描述不同类样本之间的差异而言，不一定是最优的描述，即用 PCA 来描述人脸识别的特征是不充分的。线性鉴别分析不同于主元分析，它是以样本的可分性为目标，寻找一组线性变换使每类的类内离散度最小、类间的离散度最大，因此从理论上说，比较适合于模式识别问题。

经典的线性鉴别分析中使用的是 Fish 准则函数，所以线性鉴别分析又被称为 Fish 线性鉴别分析 (Fisher LDA, FDA)。FDA 方法的基本思想是：选择使得 Fisher 准则函数达到极大值的向量作为最优投影方向，使得模式样本在该方向上投影后，达到最大的类间离散度和最小的类内离散度。

Swets 等人<sup>[34]</sup>最先提出结合主元分析的线性判决方法，后来 Belhumeur 等人<sup>[35]</sup>把它发展为 Fisher 脸 (Fisherface) 方法，并在光照变化、表情变化以及有无戴眼镜的情况下与特征方法作了比较，实验结果表示 Fisher 脸要比特征脸鲁棒性好。

Liu 等人<sup>[36]</sup>用同时对角化类间散度矩阵  $S_b$  和类内散度矩阵  $S_w$  的方法, 来避免求逆的过程, 在本质上就是直接去除  $S_w$  的零空间。Huang 等人<sup>[37]</sup>则是先用 PCA 去除不含有用鉴别信息的  $S_b$  零空间, 再在 PCA 降维后的低维空间中运用零空间 FDA 方法。Yu 等人<sup>[38]</sup>提出了 Direct LDA 方法, 即先去掉  $S_b$  的零空间, 然后再使类内离散度最小化。Cevikalp 等人<sup>[39]</sup>发现样本投影到  $S_w$  的零空间后都汇聚到一点上, 从而提出了鉴别公共矢量这一概念。

Li 等人<sup>[40]</sup>和 Xiong 等人<sup>[41]</sup>提出了基于二维图像矩阵的 2DFDA 方法, 其基本思想是: 跳过将图像矩阵转换为向量的过程, 直接利用二维图像矩阵构造类内散度矩阵和类间散度矩阵, 并在此基础上利用 Fisher 鉴别准则寻找最佳映射方向。2DFDA 不仅解决了一维 FDA 遭遇的小样本问题, 同时还避免了庞大的计算量。

### ③独立元分析(ICA)方法

主元分析和线性鉴别分析都是基于训练样本集的二阶统计信息的, 而忽视了高阶统计信息, 实际上高阶统计信息有时对识别来说也是非常有用的。独立元分析的思想就是通过线性变换, 从训练样本中找到一组互相独立的基(独立元), 并以此来描述样本数据。主元分析实际上是在二阶统计意义下的去相关, 而独立元分析则是在所有阶统计意义下的去相关, 从而使信号的二阶统计和高阶统计信息都得到有效利用<sup>[42]</sup>, 因此独立元分析可以看作是对主元分析的推广。

独立元分析要求基于所有阶的统计信息, 所以在求解上就没有主元分析和线性鉴别分析那么容易, 目前比较流行的求解方法有三种<sup>[43][44]</sup>: 基于信息论的方法, 固定点算法和联合近似对角化特征矩阵的方法, 其中固定点算法<sup>[45]</sup>最为简单快捷。

独立元分析主要是应用在盲信号分离上<sup>[44]</sup>, 它在人脸识别上的应用是由 Bartlett 等人<sup>[46]</sup>首次提出的, 把人脸图像看作是多个互相独立的基图像的线性叠加。文献[47]是先用独立元分析提取特征, 然后结合支持向量机来提高分类效果。

### (4) 基于神经网络的方法

神经网络在人脸识别领域有很长的应用历史, 1994 年就出现了神经网络用于人脸处理的综述性文章<sup>[48]</sup>。神经网络因其固有的并行运算机制, 以及对模式的分布式全局存储, 后来被用于人脸识别中。

Kohonen<sup>[49]</sup>最早将自组织映射(Self-Organizing Map, SOM)神经网络应用于人脸识别, 他利用 SOM 的联想能力“回忆”人脸, 即使当输入人脸图像具有较大噪

音或者有部分图像丢失时，也能恢复出完整的人脸。

动态链接结构（Dynamic Link Architecture, DLA）是用于人脸识别的最有影响的神经网络方法<sup>[50]</sup>。DLA 试图解决传统神经网络中的一些概念性问题，其中最突出的是网络中语法关系的表达。DLA 利用突触的可塑性将神经元集合划分成若干图结构，同时保留了神经网络的优点。DLA 使用 Gabor 小波来表示图像的特征。

Lawrence 和 Giles 等人<sup>[51]</sup>采用将自组织映射神经网络（SOM）与卷积神经网络相结合方法进行人脸识别。SOM 实现对图像的采样向量降维，且对图像样本的小幅度变形不敏感；卷积网络用来实现相邻像素间的相关性知识，在一定程度上对图像的平衡、旋转、尺度和局部变形不敏感。

神经网络方法较其他人脸识别方法有着特有的优势，通过对神经网络的训练可以获得其他方法难以实现的关于人脸图像的规则和特征的隐性表示，避免了复杂的特征抽取工作，并有利于硬件的实现。神经网络方法的缺点主要在于其方法的可解释性较弱，且要求多张人脸图像作为训练集，一般只适合于小型人脸库。

### 1.3 人脸识别系统

#### 1.3.1 人脸识别系统的组成

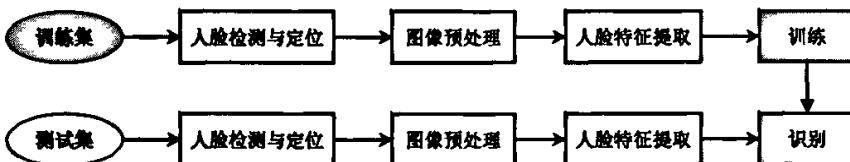


图 1-1 人脸识别系统的基本框图

如图 1-1 所示，人脸识别系统主要由五个功能模块组成：人脸检测与定位模块、图像预处理模块、人脸特征提取模块、训练模块和识别模块。

- (1) 人脸检测与定位模块：从原始图像中找到并确定人脸的位置，然后从中提取出人脸图像；
- (2) 图像预处理模块：对人脸图像进行几何归一化、光线归一化、去除噪声等操作，以便在同一条件下完成训练和识别；
- (3) 人脸特征提取模块：完成对人脸特征的提取，如何提取有效、稳定的特征

是人脸识别系统成败的关键；

(4)训练模块：对训练集进行训练，得到参数以供识别模块使用，它是人脸识别系统的核心；

(5)识别模块：根据训练所得的参数，完成最后的模式分类与判别工作。

### 1.3.2 人脸识别系统的测评指标

人脸识别系统的测评指标主要包括识别率、计算时间、数据存储量和可扩展性等，分别叙述如下：

#### (1) 识别率

在计算机人脸识别中，与识别率相关的有两个概念：正识率和拒识率。若待识别的样本图像属于人脸库中的人脸模式，且系统正确识别出该人脸模式，则这些样本图像的总数占测试样本总数的百分比为正识率，即属于该系统的某个人被正确判定是谁的概率。若待识别的样本图像不属于人脸库中的人脸模式，且系统识别出该样本图像不属于人脸库中的人脸模式，则这些样本图像的总数占测试样本总数的百分比为拒识率。这两种情况均为正确识别，识别率由这两部分构成。

在实际应用中需要给出拒识人脸，即通过设置合适的拒识阈值，识别出不属于人脸库的人脸图像。在通常的人脸识别方法研究中，大多数系统都没给出拒识人脸，而是将系统的拒识阈值设为无穷，即给定一待识别样本，在已知人脸库中找到和该样本最相近的人脸，而不考虑该样本是否应该是人脸库中的已知人脸。

#### (2) 计算时间

人脸识别的实际应用对实时性的要求比较强，因此计算时间是人脸识别系统中的一个重要指标。计算时间主要有两个方面：一是训练阶段，即人脸识别系统训练需要的时间；二是识别阶段，即人脸识别系统识别需要的时间。通常人脸识别系统的训练为离线训练，因此训练阶段所需的时间可以不考虑，但识别时间却相当重要，它直接影响人脸识别系统的实时性，对系统是否可应用于实践起着决定性作用。

#### (3) 数据存储量

在计算机人脸识别系统中，人脸库的存储量也是一个不得不考虑的问题。存储大量的人脸数据将会给识别系统造成一定的负担。因此，在开展计算机人脸识别算法研究时，也要考虑数据存储量的大小。

#### (4) 可扩展性

在实际应用中，经常需要对已有人脸库进行改动，因此，如何更加方便地对人脸库的进行动态维护，也是在研究人脸识别技术时需要考虑的问题。

### 1.4 本文的研究工作

本文围绕基于统计特征的二维主元分析(2DPCA)人脸识别方法，对影响人脸识别的因素进行了研究和探讨，提出了两种新的2DPCA改进方法。主要工作和成果如下：

#### (1) 研究了PCA方法、2DPCA方法以及二者的相互关系

本文从PCA人脸识别方法出发，在透彻地理解PCA方法的数学理论与基本原理的基础上，研究了基于二维图像矩阵的2DPCA人脸识别方法，阐述了2DPCA方法如何避免PCA方法庞大的计算量，并获得优于PCA方法的人脸识别率的原因。

通过对2DPCA方法的数学分析，给出了2DPCA方法的协方差矩阵与PCA方法的协方差矩阵之间的联系，证明了2DPCA方法的实质是基于行分块的PCA方法，同时指出了2DPCA方法抛弃了行与行之间像素的相关性信息的缺陷，为后续的研究工作奠定了理论基础。

#### (2) 提出了基于结构分块的2DPCA人脸识别新方法

针对人脸识别中的一类特殊应用——不完整人脸图像的识别问题，考虑到人脸图像具有局部特征，即人脸图像一个局部块往往表示一个完整的语义，如鼻子、眼睛、嘴巴等，本文提出了基于结构分块的2DPCA人脸识别方法。其基本思想是：首先把人脸图像分割为若干个互不重叠的图像块，然后对每一个图像块应用2DPCA方法，求出其特征矩阵，则该特征矩阵代表了确定的人脸特征（本文中称之为结构主元）。在计算图像相似度的时候，先计算出各图像块对应的相似度，然后用加权和的方法计算出最终的相似度。

基于结构分块的2DPCA人脸识别方法结合了统计模式识别与结构模式识别的优点，通过调整各图像块的加权系数，可以应用于各种不完整人脸图像的识别。本文在在FERET人脸库上进行了相关实验，证实了本文提出的基于结构分块的2DPCA方法在人脸识别率上优于2DPCA方法，而且可以通过调整各图像块的加权系数，获取更好的识别性能。

### (3) 提出了基于块内相关性的 2DPCA 人脸识别新方法

通过对 2DPCA 的数学分析可以发现，2DPCA 方法是利用图像矩阵同一行像素之间的相关性来构造协方差矩阵的，其实质是基于行分块 PCA 方法。2DPCA 方法只利用了图像矩阵同一行像素之间的相关性信息，完全抛弃了行与行之间像素的相关性信息，然而对于人脸图像来说，图像矩阵同一行像素之间的相关性没有一个局部块像素之间的相关性强，而且人脸图像一个局部块通常具有特定的语义。基于上述考虑，本文提出了基于块内相关性的 2DPCA 人脸识别方法，其基本思想是：首先对人脸图像进行分块，然后把图像块的像素转化为行向量，再把这些行向量按图像块的顺序重排为新的图像矩阵，最后对这些新的图像矩阵采用 2DPCA 方法进行人脸识别。

基于块内相关性的 2DPCA 人脸识别方法不但保留了 2DPCA 方法低计算复杂度的优点，而且能够充分地利用图像行与列像素之间的相关性信息，克服了 2DPCA 方法的不足。本文在 FERET 人脸库上进行了相关实验，证实了本文提出的基于块内相关性的 2DPCA 方法在识别率、特征向量数目、数据存储量等多个方面都优于 2DPCA 方法。

## 1.5 本文的内容安排

本文总共分为六章，各章的内容安排如下：

### 第 1 章：综述

介绍了人脸识别的基本概念和应用场合；人脸识别研究的主要内容，人脸识别研究的各种主流方法及其发展现状，并讨论了人脸识别系统的组成以及系统的测评指标，最后给出了本论文的主要研究工作与成果以及各章的内容安排。

### 第 2 章：研究主元分析（PCA）人脸识别方法

介绍了主元分析（PCA）人脸识别方法的基本原理与实现过程，对主元分析所涉及的数学理论作了简明而清晰的介绍，并对经典 PCA 方法中的特征脸（Eigenface）方法作了详细的说明。

### 第 3 章：研究基于二维图像矩阵的 2DPCA 人脸识别方法

介绍了二维主元分析（2DPCA）人脸识别方法的基本原理与实现过程，并给出了 2DPCA 方法与 PCA 方法的性能比较，验证了 2DPCA 方法的优越性并确定其

为本论文的改进对象。此外，还介绍了 2DPCA 人脸识别的一些改进方法。

**第4章：提出基于结构分块的 2DPCA 人脸识别新方法**

这是本文的重点，针对不完整人脸图像的识别问题，在考虑了人脸图像所具有局域特性后，提出了基于结构分块的 2DPCA 人脸识别方法。该方法结合了统计模式识别与结构模式识别的优点，通过调整各图像块的加权系数，可应用于各种不完整人脸图像的识别。本文在 FERET 人脸库上对基于结构分块的 2DPCA 人脸识别方法进行了实验验证。

**第4章：提出基于块内相关性的 2DPCAA 人脸识别新方法**

这是本文的另一重点，提出了基于块内相关性的 2DPCA 人脸识别方法。通过对 2DPCA 方法的数学分析，得知它是利用图像矩阵同一行像素之间的相关性来构造协方差矩阵的，并未利用行与行之间像素的相关性信息。基于块内相关性的 2DPCA 方法通过对人脸图像进行分块、重排，利用图像矩阵局部块像素之间的相关性来构造协方差矩阵，充分地利用图像行与列像素之间的相关性信息。本文在 FERET 人脸库上对基于块内相关性的 2DPCA 人脸识别方法进行了实验验证。

**第6章：总结全文，并指出进一步要做的工作。**

## 第2章 PCA 人脸识别方法

### 2.1 引言

主元分析（Principal Component Analysis，简称 PCA）是统计学中分析数据的一种有效的方法<sup>[62]</sup>，其基本思想是提取出高维数据空间中的主要特征（主元），使数据在一个低维的特征空间被处理，同时保持原始数据的绝大部分的信息，从而解决数据空间维数过高的瓶颈问题。主元分析在信号处理、模式识别、数字图像处理等领域已得到广泛的应用。

K-L 变换是图像压缩中的一种最优正交变换<sup>[25]</sup>，人们将它用于统计特征提取，从而形成了子空间法模式识别的基础。PCA 方法根据 K-L 变换从最大信息压缩方向获得图像在低维空间的信息表达，从信息论的观点来看，在所有的正交变换中，K-L 变换所对应的信息熵最小，所以用 PCA 方法所获得的识别空间是原图像空间的一个最优低维逼近<sup>[8][10]</sup>。

PCA 技术首先由 Kirby 与 Stroovich 引入到人脸识别领域<sup>[25]</sup>，并且证明了 PCA 是使原始图像与重构图像之间的均方误差极小化的最佳压缩方式，后来 Turk 和 Pentland 将它成功地发展为特征脸（Eigenface）方法<sup>[9][10]</sup>。

PCA 人脸识别方法是从人脸图像整体代数特征出发，对图像的总体信息进行分类识别的，是基于人脸全局表达的一种相当有效识别方法。

本章详细叙述了 PCA 人脸识别方法的基本原理，包括涉及到的各种数学知识，然后介绍特征脸方法，并给出了经典 PCA 人脸识别的实现过程。

### 2.2 主元分析的基本原理及计算

主元分析方法是满足在特征提取过程中，测量空间到特征空间的映射准则的一种数据压缩方法，其思想来源于 K-L 变换（Karhunen-Loeve Transform, KLT），目的是通过线性变换寻找一组最优的标准正交向量基，通过它们的线性组合来重建原始样本，并使得重建后的新样本与原始样本之间的均方误差最小。在数学上，主元分析就是通过解特征值问题来对角化协方差矩阵。

PCA 方法的基本思想如图 2-1 所示，这是一个二维情况下的例子<sup>[65]</sup>。图(a)中实线轴 ( $x_1$  与  $x_2$ ) 表示原始数据所在空间的正交基，虚线轴 ( $\phi_1$  与  $\phi_2$ ) 表示经过 K-L 变换后得到的正交基。第一主轴是  $\phi_1$ ，对应于方差最大的投影方向；第二主轴是  $\phi_2$ ，在二维的情况下  $\phi_2$  可以根据正交约束唯一确定，但在高维空间中，坐标轴的选取还要根据在投影方向上的方差的大小顺序来确定。图(b)是将 2D 数据降维为 1D 的情况，即只保留第一主轴  $\phi_1$ ，并将数据点往  $\phi_1$  方向上投影的情况。在数据降维的过程中，由于方差则越大投影后的数据点越分散，所以保留方差较大的坐标轴有利于保持数据的分布特征，即有利于模式分类。

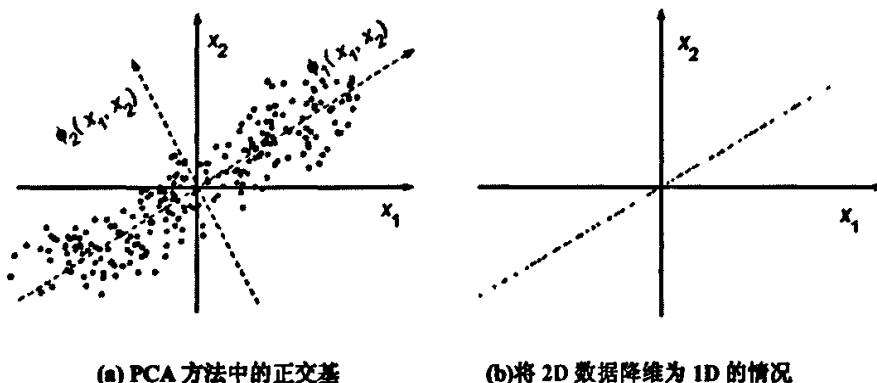


图 2-1 主元分析（PCA）基本思想示意图

### 2.2.1 特征提取的基本概念

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），这样产生的特征叫做原始特征。原始特征的数量可能很大，或者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）。

所谓特征提取在广义上就是指一种变换，若  $R^n$  是测量空间， $R^m$  是特征空间，则变换  $T: R^n \rightarrow R^m$  就叫做特征提取器。在数学上，特征提取就是从测量空间  $R^n$  到特征空间  $R^m$  的映射，映射通常要遵循以下两个准则：

一、特征空间必须保留测量空间中的主要分类信息；

二、特征空间的维数必须远低于测量空间的维数（即有  $m \ll n$ ）。

### 2.2.2 离散 K-L 变换

离散 K-L 变换 (Discrete Karhunen-Loeve Transform, DKLT) [63][64] 是一种基于目标统计特征的最优正交变换，它具有非常优良的特性：使变换后产生的新的分量正交或不相关；以部分新的分量表示原矢量的均方误差最小；使变换矢量更趋确定、能量更趋集中等等。这些特性使得离散 K-L 变换在特征选取、数据压缩等方面都有着极为重要的应用。

设  $n$  维随机向量  $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n]^T$ ，其均值向量为  $E[\mathbf{X}]$ ，协方差矩阵为  $\Sigma_{\mathbf{X}} = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T]$ 。设矩阵  $\Phi = [\phi_1 \ \phi_2 \ \cdots \ \phi_n]$  是由  $n$  个线性无关的列向量组成的标准正交矩阵，即有  $|\Phi| \neq 0$ ，且

$$\phi_i^T \phi_j = \begin{cases} 1 & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \quad (2-1)$$

则矩阵  $\Phi$  的列向量  $\{\phi_1, \phi_2, \dots, \phi_n\}$  可以张成一个  $n$  维的空间，向量组  $\{\phi_1, \phi_2, \dots, \phi_n\}$  是这个  $n$  维空间的一组标准正交基。

随机向量  $\mathbf{X}$  经矩阵  $\Phi^T$  正交变换后，成为随机向量  $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$ ，即有

$$\mathbf{Y} = \Phi^T \mathbf{X} = [\phi_1 \ \phi_2 \ \cdots \ \phi_n]^T \mathbf{X} = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \vdots \\ \phi_n^T \end{bmatrix} \mathbf{X} \quad (2-2)$$

从而  $\mathbf{Y}$  的各分量

$$y_i = \phi_i^T \mathbf{X} \quad (i=1,2,\dots,n) \quad (2-3)$$

而

$$\mathbf{X} = (\Phi^T)^{-1} \mathbf{Y} = \Phi \mathbf{Y} = [\phi_1 \ \phi_2 \ \cdots \ \phi_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n y_i \phi_i \quad (2-4)$$

假设我们想只保留  $m$  个 ( $m < n$ ) 特征投影向量  $\phi_i$ ，并在最小均方误差准则之下线性估计随机向量  $\mathbf{X}$ ，我们可以保留向量  $\mathbf{Y}$  的前  $m$  个分量  $y_1, y_2, \dots, y_m$ ，而余下的

( $n-m$ )个分量  $y_{m+1}, y_{m+2}, \dots, y_n$  则分别由预选的( $n-m$ )个常数  $b_i$  ( $i = m+1, \dots, n$ )代替, 此时的估计式为:

$$\hat{X}(m) = \sum_{i=1}^m y_i \phi_i + \sum_{i=m+1}^n b_i \phi_i \quad (2-5)$$

估计的均方误差为

$$\begin{aligned} s^2(m) &= E[(X - \hat{X}(m))^T (X - \hat{X}(m))] \\ &= E\left[\left(\sum_{i=m+1}^n (y_i - b_i) \phi_i\right)^T \left(\sum_{i=m+1}^n (y_i - b_i) \phi_i\right)\right] \\ &= E\left[\sum_{i=m+1}^n \sum_{j=m+1}^n (y_i - b_i)(y_j - b_j) \phi_i^T \phi_j\right] \\ &= \sum_{i=m+1}^n E[(y_i - b_i)^2] \end{aligned} \quad (2-6)$$

(1)求最佳的  $b_i$

由  $\frac{\partial s^2(m)}{\partial b_i} = 0$  得

$$\frac{\partial}{\partial b_i} \{E[(y_i - b_i)^2]\} = -2\{E[y_i] - b_i\} = 0 \quad (2-7)$$

解得

$$b_i = E[y_i] = \phi_i^T E[X] \quad (2-8)$$

于是

$$\begin{aligned} s^2(m) &= \sum_{i=m+1}^n E[(y_i - E[y_i])^2] \\ &= \sum_{i=m+1}^n \phi_i^T E[(X - E[X])(X - E[X])^T] \phi_i \\ &= \sum_{i=m+1}^n \phi_i^T \Sigma_X \phi_i \end{aligned} \quad (2-9)$$

(2)求最佳的  $\phi_i$

我们希望在  $\Phi$  为标准正交矩阵的约束下, 使  $s^2(m)$  的取值最小, 作准则函数

$$J = \sum_{i=m+1}^n [\phi_i^T \Sigma_X \phi_i - \lambda_i (\phi_i^T \phi_i - 1)] \quad (2-10)$$

由  $\frac{\partial J}{\partial \phi_i} = 0$  得

$$\Sigma_x \phi_i = \lambda_i \phi_i \quad (i = m+1, \dots, n) \quad (2-11)$$

上式表明，这时的  $\phi_i$  为  $X$  的协方差矩阵  $\Sigma_x$  的对应于特征值  $\lambda_i$  的特征向量。

将上式代入  $\varepsilon^2(m)$  式中，可得

$$\varepsilon^2(m) = \sum_{i=m+1}^n \lambda_i \quad (2-12)$$

由上述讨论可知，为了使随机向量  $X$  与其估计值  $\hat{X}(m)$  之间的均方误差  $\varepsilon^2(m)$  最小化，应选取对应于协方差矩阵  $\Sigma_x$  前  $m$  个最大特征值的特征向量，来构造正交变换矩阵  $\tilde{\Phi} \in R^{n \times m}$ 。这  $m$  个特征向量所组成正交坐标系，称作  $X$  所在的  $n$  维空间的  $m$  维 K-L 变换坐标系， $X$  在 K-L 变换坐标系上的展开系数向量（即  $Y$ ）称作  $X$  的 K-L 变换系数。

图 2-2 是二维情况下 K-L 变换的一个例子，数据分布于椭圆轮廓的范围之内，协方差矩阵的特征向量  $\phi_1$  与  $\phi_2$  是该分布的两个主轴，特征值  $\lambda_1$  与  $\lambda_2$  分别是沿着  $\phi_1$  轴与  $\phi_2$  轴的数据分布的方差。图中  $X$  为某一样本点，由  $y_i = \phi_i^T X$  可知， $y_1$  与  $y_2$  分别是  $X$  在  $\phi_1$  轴与  $\phi_2$  轴上的投影，若  $\lambda_2 = E[(y_2 - b_2)^2]$  比  $\lambda_1$  要小，则  $y_2$  与  $b_2$  之间距离小于  $y_1$  与  $b_1$  之间的距离，因此可以用  $\hat{X} = y_1 \phi_1 + b_2 \phi_2$  对  $X$  进行估计，即图中的点  $\hat{X}$ 。可以看出，保留较大特征值对应的特征向量，估计所得到的误差较小。

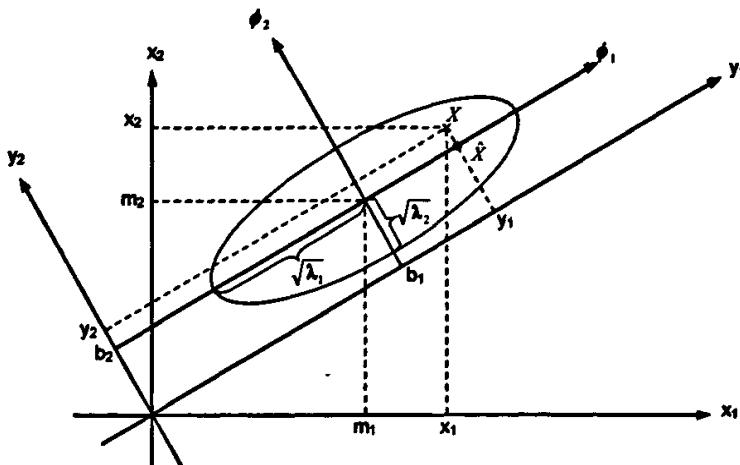


图 2-2 二维情况下的 K-L 变换示意图

### 2.2.3 SVD 定理

由于 PCA 方法需要求解随机向量  $X$  的协方差矩阵  $\Sigma_x$  的特征值与特征向量，但通常随机向量  $X$  的维数是相当高的，因而协方差矩阵  $\Sigma_x$  相当巨大（以 PCA 人脸识别为例，FERET 人脸库的图像尺寸为 92? 12 像素，转换为一维列向量则维数为 10304，产生的协方差矩阵的维数为 10304? 10304），对如此高维数的矩阵求解特征值和特征向量是相当困难的，超出了一般计算机所能处理的范围。因此，PCA 方法中引入了奇异值分解（Singular Value Decomposition，简称 SVD）定理<sup>[67][68]</sup>来解决数据维数过高的问题。

SVD 定理的描述如下：

假设  $A$  是一个  $n \times m$  维的实矩阵，即  $A \in R^{n \times m}$ ，矩阵  $A$  的秩为  $r$ ，则存在一个  $n \times n$  维的正交矩阵和一个  $m \times m$  维的正交矩阵  $V$

$$U = [u_1 \ u_2 \ \dots \ u_n], \quad U \in R^{n \times n}, \quad U^T U = I \quad (2-13)$$

$$V = [v_1 \ v_2 \ \dots \ v_m], \quad V \in R^{m \times m}, \quad V^T V = I \quad (2-14)$$

以及一个  $n \times m$  维的广义对角矩阵  $\Sigma$

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma \in R^{n \times m}, \quad \Sigma_r \in R^{r \times r} \quad (2-15)$$

$$\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 \quad (2-16)$$

则有

$$A = U \Sigma V^T \quad (2-17)$$

其中， $\sigma_i (i=1,2,\dots,r)$  是矩阵  $A$  的奇异值，它是矩阵  $AA^T$  或矩阵  $A^T A$  的特征值  $\lambda_i (i=1,2,\dots,r)$  的非负平方根。向量  $u_i (i=1,2,\dots,r)$  与  $v_i (i=1,2,\dots,r)$  分别为矩阵  $AA^T$  与  $A^T A$  对应于特征值  $\lambda_i (i=1,2,\dots,r)$  的正交归一特征向量。

PCA 方法中，若样本为  $X_i \in R^n (i=1,2,\dots,m)$ ，矩阵  $A = [\hat{X}_1, \hat{X}_2, \dots, \hat{X}_m] \in R^{n \times m}$ ，其中  $\hat{X}_i = X_i - \frac{1}{m} \sum_{i=1}^m X_i (i=1,2,\dots,m)$ ，则协方差矩阵  $C = AA^T$ 。由于维数过高难以直接计算协方差矩阵  $C$  的特征值与特征向量，由 SVD 定理知矩阵  $AA^T$  与矩阵  $A^T A$  具有相同的特征值，则可以构造矩阵

$$R = A^T A \quad (2-18)$$

通常有  $m \ll n$ , 因此容易求得矩阵  $R$  的特征值  $\lambda_i (i=1,2,\dots,m)$  以及相应的正交归一特征向量  $v_i (i=1,2,\dots,m)$ , 由如下推导

$$A^T A v_i = \lambda_i v_i \Rightarrow A A^T A v_i = \lambda_i A v_i \Rightarrow C A v_i = \lambda_i A v_i \quad (2-19)$$

可知, 协方差矩阵  $C$  的正交归一特征向量  $u_i (i=1,2,\dots,m)$  为

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad (i=1,2,\dots,m) \quad (2-20)$$

#### 2.2.4 Gram-Schmidt 正交化

在 K-L 变换中, 我们需要选取  $n$  维随机向量  $X$  的协方差矩阵  $\Sigma_x$  的对应于前  $m$  个最大特征值的特征向量, 来构造正交变换矩阵  $\tilde{\Phi}$ 。这一组  $m$  个特征向量能够张成  $n$  维空间的一个  $m$  维子空间, 然而它们未必是一组标准正交基(即不一定满足两两正交与及单位化的条件), 因而我们需要有一种方法能够求出这个  $m$  维的子空间的一组标准正交基, Gram-Schmidt (格兰姆-施密特) 正交化就提供了这种方法。

Gram-Schmidt 正交化<sup>[69][70]</sup>是线性代数理论中一个很著名的定理, 它是将欧氏空间  $V^n$  中一组线性无关的向量  $\{v_1, v_2, \dots, v_m\} (m \leq n)$  化为一组标准正交基  $\{\eta_1, \eta_2, \dots, \eta_m\} (m \leq n)$  的一种典型方法。Gram-Schmidt 正交化的基本想法, 是利用投影原理, 在已有正交基的基础上构造一个新的正交基, 其计算过程如下:

设  $m$  维向量组  $\{v_1, v_2, \dots, v_m\}$  是线性无关的, 则

$$\beta_1 = v_1 \quad \eta_1 = \frac{\beta_1}{\|\beta_1\|} \quad (2-21)$$

$$\beta_2 = v_2 - \langle v_2, \eta_1 \rangle \eta_1 \quad \eta_2 = \frac{\beta_2}{\|\beta_2\|} \quad (2-22)$$

$$\beta_3 = v_3 - \langle v_3, \eta_1 \rangle \eta_1 - \langle v_3, \eta_2 \rangle \eta_2 \quad \eta_3 = \frac{\beta_3}{\|\beta_3\|} \quad (2-23)$$

⋮

$$\beta_m = v_m - \sum_{i=1}^{m-1} \langle v_m, \eta_i \rangle \eta_i \quad \eta_m = \frac{\beta_m}{\|\beta_m\|} \quad (2-24)$$

这样就得到了  $span\{v_1, v_2, \dots, v_m\}$  上的一组正交基  $\{\beta_1, \beta_2, \dots, \beta_m\}$ , 以及相应的标准正交基  $\{\eta_1, \eta_2, \dots, \eta_m\}$ 。

### 2.2.5 主元分析的计算步骤

前面几节叙述了主元分析的基本原理和涉及到的数学知识，现在我们可以把主元分析的计算步骤归纳如下：

假设总共有  $M$  个样本  $X_1, X_2, \dots, X_M$ ，每个样本都是  $n$  维的列向量 ( $M \ll n$ )，则样本均值向量为

$$\bar{X} = \frac{1}{M} \sum_{i=1}^M X_i \quad (i=1,2,\dots,M) \quad (2-25)$$

作如下定义

$$\hat{X}_i = X_i - \bar{X} \quad (i=1,2,\dots,M) \quad (2-26)$$

$$A = [\hat{X}_1 \quad \hat{X}_2 \quad \dots \quad \hat{X}_M] \quad (2-27)$$

则样本协方差矩阵

$$C = \frac{1}{M} \sum_{i=1}^M (X_i - \bar{X})(X_i - \bar{X})^T = \frac{1}{M} \sum_{i=1}^M \hat{X}_i \hat{X}_i^T = \frac{1}{M} AA^T \quad (2-28)$$

矩阵  $C$  是一个  $n \times n$  维的巨型矩阵，直接求解如此高维数矩阵的特征值与特征向量是不现实的，我们可以引入 SVD 定理来解决这个问题，构造如下矩阵

$$R = A^T A \quad (2-29)$$

矩阵  $R$  是一个  $M \times M$  维的矩阵，远低于矩阵  $C$  的维数，容易求出其特征值  $\lambda_i$  ( $i=1,2,\dots,M$ ) (注意应将特征值按从大到小的顺序排列，即  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M > 0$ ) 以及相应的正交归一特征向量  $v_i$  ( $i=1,2,\dots,M$ )，由 SVD 定理可得，矩阵  $C$  的正交归一特征向量  $\phi_i$  ( $i=1,2,\dots,M$ ) 为

$$\phi_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad (i=1,2,\dots,M) \quad (2-30)$$

这样，我们就通过计算低维矩阵  $R$  和特征向量而间接求出了样本协方差矩阵的特征向量。

然后，可以根据各个特征向量  $\phi_i$  ( $i=1,2,\dots,M$ ) 贡献率的大小，根据如下公式来选取一定的主元数目  $k$ ，来构造正交变换矩阵  $\tilde{\Phi}$ 。

$$\gamma = \frac{\sum_{i=1}^k |\lambda_i|}{\sum_{i=1}^M |\lambda_i|} \quad (i=1,2,\dots,M) \quad (2-31)$$

其中，阈值 $\gamma$ 表示前 $k$ 个特征向量的累积贡献率，通常取值范围为85%~95%。经由上述步骤得到正交变换矩阵 $\tilde{\Phi} = [\phi_1 \ \phi_2 \ \cdots \ \phi_k]$ 之后，就可以对任意 $X$ 进行如下正交变换，得到相应的降维后的向量 $Y$

$$Y = \tilde{\Phi}^T X = [\phi_1^T \ \phi_2^T \ \cdots \ \phi_k^T]^T X = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \vdots \\ \phi_k^T \end{bmatrix} X \quad (2-32)$$

从而 $Y$ 的各分量

$$y_i = \phi_i^T X \quad (i=1,2,\cdots,k) \quad (2-33)$$

这样，就完成了从测量空间 $R^n$ 到特征空间 $R^k$  ( $k \ll n$ ) 的特征提取过程。

## 2.3 经典 PCA 人脸识别

一幅人脸图像按照某种方式（通常是按行）排列成一个 $n$ 维向量，则可以看作是 $n$ 维空间中的一个点，由于人脸结构的相似性，人脸图像在这在一高维空间中的分布并不是散乱的，而是呈现出一定规律，即在 $n$ 维空间中聚集为一个相对较低维的“人脸子空间”。PCA 人脸识别方法的基本思想就是试图确定一组最能反映这一分布特点的正交基向量，使得所有人脸图像经过投影之后均落入“人脸子空间”中，从而获得人脸图像在低维空间的信息表达。

基于 K-L 变换的 PCA 人脸识别方法，在 20 世纪 90 年代初受到很大的重视，其人脸识别效果也相当不错，识别率从 70% 到 100% 不等，主要取决于人脸库图像的质量。从能量压缩的角度来看，K-L 变换是最优的，它不仅使得从高维空间降到低维空间产生的均方误差最小，而且变换后的低维空间有很好的人脸表达能力。

### 2.3.1 特征脸 (Eigenface) 算法

在各种 PCA 人脸识别方法中，特征脸 (Eigenface) 方法是一个相当经典的人脸识别算法，于 1991 年由 Turk 和 Pentland<sup>[8]</sup> 提出，是一种基于主元分析的线性子空间人脸识别方法。特征脸方法由于计算简单，概念清晰易于实现，而且效果良好，目前已经成为人脸识别的基准程序 (Benchmark) 和事实上的工业标准。

特征脸方法的基本思想如下：

人脸图像被看作是一个随机向量，任意给定的一幅人脸图像则是这个随机向量的一个实现。对训练图像集进行主元分析，可以得到一组正交归一的特征向量，即低维主元子空间的标准正交基。如果将这些特征向量按照图像阵列排列，则可以看出这些特征向量呈现人脸的形状（如图 2-3 所示），因此这些特征向量被称作特征脸（Eigenface）。任意给的一幅人脸图像都可以近似为这组特征脸图像的线性组合，用组合的系数作为人脸的特征向量，这组系数表明了该人脸图像在降维子空间中的位置。这样，原来的人脸图像识别的问题就转化为根据子空间中的样本点进行模式分类的问题，经典的特征脸方法采用基于 Euclidean 距离的最近中心分类器，后来比较常用的是基于 Euclidean 距离的最近邻分类器。



图 2-3 特征脸（Eigenface）示意图（最左边是平均脸，其余是前 7 个特征脸）

在实际计算中，对于每一幅人脸图像，都按照从左到右、从上到下的扫描顺序，将所有像素点的灰度值排列成一个高维列向量，然后通过主元分析（PCA）进行数据降维，即对高维数据进行特征提取以获取相应的低维表示，最后再使用基于欧氏距离的最近邻分类器进行模式分类。

在特征脸方法中，使用 PCA 进行降维主要基于以下三点：

- (1) 压缩功能，在低维空间内比较图像可以大大减少计算量，提高计算效率；
- (2) 人脸样本的分布近似于正态分布，方差大的维可能与有用信号相关，而方差小的维可能对应于噪声，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，且被缩放成单位向量，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。

尽管特征脸方法的人脸识别效果良好，但相关研究表明，随着光线、姿态及人脸的尺寸等外界因素的引入，特征脸方法的识别率会有较大幅度的下降，这是因为对 K-L 变换而言，外界因素带来的图像差异和人脸自身所造成的图像差异是不加任何区分的。对特征脸方法进行的改善通常有两个思路：

一、针对外界因素带来的干扰，对输入图像作规范化处理，如人脸尺寸归一化、光线补偿、灰度归一化等等，以消除各种不利因素带来的图像差异。

**二、考虑到局部人脸图像受到的外界干扰相对较小，在进行人脸识别时，除计算特征脸之外，还可以用同样的原理计算出“特征眼睛”、“特征嘴巴”等等，将局部特征向量与全局特征向量加权之后再进行匹配，就能得到更好的效果。**

### 2.3.2 经典 PCA 人脸识别的实现过程

完整的 PCA 人脸识别算法包括以下两个过程：

**一、建库过程，其步骤如下：**

- (1) 获取一系列人脸图像并进行相关预处理，建立训练集；
- (2) 对训练集进行主元分析，得到一组特征脸，形成人脸子空间；
- (3) 把训练集中的图像投影到人脸子空间中，保存相应的人脸特征向量。

**二、识别过程，其步骤如下：**

- (1) 对测试图像进行预处理；
- (2) 将测试图像投影到人脸子空间中，得到相应的人脸特征向量；
- (3) 选择一定的距离函数，对人脸特征向量进行分类判别。

下面详细描述 PCA 人脸识别的实现过程。

### 2.3.3 建库过程

#### (1) 建立训练集

首先，应对输入的人脸图像进行预处理，包括几何归一化、光线归一化等，尽量去除尺度、光线等因素对 PCA 人脸识别过程带来的不利影响。然后，需要将经过预处理的二维人脸图像数据转化为一维的向量，通常采用各行像素首尾相接的方式产生，所有这些一维向量就构成了人脸图像的训练集。

假设每一幅经过预处理的人脸图像，都可以用一个  $m \times n$  的二维数组  $I(x, y)$  来表示，数组元素表示像素点的灰度值，将二维数组转化为一维向量则有  $X \in R^m$ 。若训练集中一共有  $P$  个人，每人有  $L$  张不同的人脸图像，且训练集中人脸图像的总数为  $M = P \times L$ ，则整个训练集可记作  $\{X_1, X_2, \dots, X_M\}$ ，其中，每一个向量  $X_i \in R^m$  ( $i = 1, 2, \dots, M$ ) 对应于一幅人脸图像。

#### (2) 通过训练形成特征子空间

训练集中  $M$  幅人脸图像的均值向量：

$$\bar{X} = \frac{1}{M} \sum_{i=1}^M X_i \quad (2-34)$$

每幅人脸图像  $X_i$  与平均人脸  $\bar{X}$  的差值向量为：

$$\hat{X}_i = X_i - \bar{X} \quad (i=1,2,\dots,M) \quad (2-35)$$

则训练集的协方差矩阵可表示为：

$$C = \frac{1}{M} \sum_{i=1}^M (\hat{X}_i - \bar{\hat{X}})(\hat{X}_i - \bar{\hat{X}})^T = \frac{1}{M} \sum_{i=1}^M \hat{X}_i \hat{X}_i^T = \frac{1}{M} A A^T \quad (2-36)$$

其中，矩阵  $A = [\hat{X}_1 \ \hat{X}_2 \ \dots \ \hat{X}_M]$  是所有差值向量构成的矩阵。

为了计算矩阵  $C$  的特征值与特征向量，我们可以构造低维矩阵  $R = A^T A$ ，容易求出其特征值  $\lambda_i$  ( $i=1,2,\dots,M$ ，且有  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M > 0$ ) 及相应的正交归一特征向量  $v_i$  ( $i=1,2,\dots,M$ )，然后由 SVD 定理，即可得到协方差矩阵  $C$  的正交归一特征向量  $u_i$  ( $i=1,2,\dots,M$ )

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad (i=1,2,\dots,M) \quad (2-37)$$

然后，可根据下式选定前  $k$  个特征向量  $u_1, u_2, \dots, u_k$  构成特征子空间：

$$\gamma = \frac{\sum_{i=1}^k |\lambda_i|}{\sum_{i=1}^M |\lambda_i|} \quad (2-38)$$

其中，阈值  $\gamma$  表示前  $k$  个特征向量的累积贡献率，通常取值范围为 85%~95%

### (3) 计算训练集的人脸特征向量

人脸特征向量即将人脸图像  $X_i$  向特征子空间投影之后得到的系数向量，即：

$$Y_i = U^T X_i = [u_1 \ u_2 \ \dots \ u_k]^T X_i = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_k^T \end{bmatrix} X_i \quad (i=1,2,\dots,M) \quad (2-39)$$

将训练集中所有已知人脸图像的特征向量  $Y_i$ ，连同其图像向量  $X_i$  以及对应的个人身份信息保存起来，就构成了已知人脸图像的数据库。

### 2.3.4 识别过程

#### (1) 图像预处理

与建立训练集类似地，对测试图像进行相应的预处理，然后再将二维人脸图像转化为一维向量。

#### (2) 计算测试图像的人脸特征向量

假设向量  $X^*$  表示一幅待识别的测试图像，将  $X^*$  向特征子空间投影，即可得到相应的人脸特征向量：

$$Y^* = U^T X^* = [u_1 \ u_2 \ \cdots \ u_k]^T X^* = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_k^T \end{bmatrix} X^* \quad (i=1,2,\cdots,M) \quad (2-40)$$

#### (3) 选择距离函数进行分类判别

将人脸图像投影到特征子空间，得到相应的人脸特征向量之后，余下的任务就是如何判别测试图像所属的类别，即归类问题或判别决策问题。

首先，需要对图像之间的相似性进行度量，常见的度量方式有：

##### ① 欧氏 (Euclidean) 距离

也称欧几里德距离， $d$  维向量  $X$  与  $Y$  之间的欧氏距离定义为：

$$D(X, Y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2} \quad (2-41)$$

##### ② 马氏 (Mahalanobis) 距离

两个服从同一分布且协方差矩阵为  $\Sigma$  的向量  $X$  与  $Y$  之间的马氏距离定义为：

$$D(X, Y) = \sqrt{(X - Y)^T \Sigma^{-1} (X - Y)} \quad (2-42)$$

马氏距离排除了不同特征之间相关性的影响。只有当  $\Sigma$  是使对角阵时，各特征之间才是完全独立的。当  $\Sigma$  是单位矩阵时，马氏距离等于欧氏距离。

##### ③ 民氏 (Minkowsky) 距离

民氏距离是若干距离的同式，其定义如下：

$$D(X, Y)_\lambda = \sqrt{\sum_{i=1}^d |x_i - y_i|^\lambda} \quad (2-43)$$

当  $\lambda=2$  时，它等于欧氏距离；当  $\lambda=1$  时，它又称作街区 (City Block) 距离。

#### ④角度相似性度量

向量  $X$  与  $Y$  之间的角度相似性定义为它们之间夹角的余弦，即：

$$S(X, Y) = \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|} \quad (2-44)$$

角度相似性度量对于坐标系的旋转和放大缩小是不变的量，但是对于位移和一般的线性变换则不是不变的。

计算出图像之间的相似度以后，需要进行判别决策，常用方法有最近邻分类器、K-近邻分类器、贝叶斯分类器等。经典 PCA 人脸识别方法中，使用的是基于欧氏距离的最近邻分类器，其分类过程如下：

对所有训练样本  $X_1, X_2, \dots, X_M$  ( $M$  为训练集中图像总数)，每一个样本都属于某个确定的类别  $\omega_k$  ( $k = 1, 2, \dots, P$ ) ( $P$  为训练集中人的个数)，相应的人脸特征向量分别为  $Y_1, Y_2, \dots, Y_M$ 。对于某个给定的测试图像  $X^*$ ，其人脸特征向量为  $Y^*$ ，若有  $dist(Y^*, Y_i) = \min_j dist(Y^*, Y_j)$  ( $dist(\bullet)$  表示相似性度量)，且有  $Y_i \in \omega_k$ ，则分类的结果是  $X^* \in \omega_k$ 。

## 第3章 2DPCA 人脸识别方法

### 3.1 引言

长期以来，PCA 人脸识别方法一直受到计算复杂性的困扰，这是由于 PCA 是基于向量的特征提取方法，即在进行 PCA 处理之前，需要将图像从二维矩阵转化为一维向量，然后才能计算协方差矩阵和求解主元特征向量。对尺寸为  $N \times N$  的图像，其协方差矩阵的维数为  $N^2 \times N^2$ ，随着  $N$  的增大，计算一个  $N^2 \times N^2$  矩阵的特征向量所需的计算量是相当惊人的，相当于一个 NP 完全的问题。

因此，在 PCA 人脸识别中，一般会引入奇异值分解（SVD）的方法来解决维数过高的问题。但由于协方差矩阵的维数与样本数目通常有很大差距，用样本估计产生的协方差矩阵的并不准确，导致计算所得的特征向量也不够准确，从而会影响到人脸识别率。然而，当训练样本的数目增多时（理论上样本越多越好，因为样本越多，样本平均就越接近统计平均），这种引入 SVD 技术的做法仍然难以摆脱计算量上的难题。

为了减少 PCA 方法的运算量，Yang 和 Zhang 对 PCA 方法进行了分析，提出了一种称作二维主元分析（Two-Dimensional Principal Component Analysis, 2DPCA）的人脸识别方法<sup>[31][32]</sup>。2DPCA 方法是对传统 PCA 方法的一种改进，可以直接在二维图像矩阵上进行处理，而不需要事先将图像矩阵转换为一维向量，从而避免了庞大的计算量，解决了 PCA 方法在计算复杂性上的困扰。

2DPCA 方法直接利用二维图像矩阵去构造图像协方差矩阵，进而求出协方差矩阵的主元特征向量（即最优投影方向），然后将二维图像直接投影到最优投影方向上，从而获得图像的特征表示。现有实验表明，2DPCA 方法的运算时间大约是 PCA 方法的  $1/4 \sim 1/20$ ，且 2DPCA 方法的人脸识别率高于 PCA 方法，特别是在样本数目较少的情况下，人脸识别率一般有大于 5% 的显著提升<sup>[31]</sup>。

本章详细叙述了 2DPCA 人脸识别方法的原理与实现过程，并给出了文献中 2DPCA 方法与 PCA 方法的性能比较，最后还介绍了 2DPCA 人脸识别的一些改进方法，如 2DPCA+PCA、K2DPCA、B2DPCA 等。

### 3.2 2DPCA 人脸识别的基本原理

#### 3.2.1 基本思想

设  $X$  表示  $n$  维归一化的列向量, 图像  $A$  是一个  $m \times n$  的随机矩阵, 2DPCA 方法的基本思想是通过如下线性变换<sup>[32][33]</sup>:

$$Y = AX \quad (3-1)$$

将矩阵  $A$  投影到向量  $X$  上, 得到一个  $m$  维的列向量  $Y$ , 称之为图像  $A$  的投影特征向量, 向量  $X$  则称作投影向量。

为了确定最优的投影向量  $X$ , 我们使用投影后训练样本的总体散度来衡量投影向量  $X$  的鉴别能力。由于投影后训练样本的总体散度特性可以用投影特征向量的协方差矩阵的迹来描述, 故定义如下准则函数:

$$J(X) = \text{tr}(S_x) \quad (3-2)$$

其中,  $S_x$  表示投影特征向量的协方差矩阵,  $\text{tr}(S_x)$  表示  $S_x$  的迹。

协方差矩阵  $S_x$  可由下式计算:

$$\begin{aligned} S_x &= E(Y - EY)(Y - EY)^T \\ &= E[AX - E(AX)][AX - E(AX)]^T \\ &= E[(A - EA)X][(A - EA)X]^T \end{aligned} \quad (3-3)$$

则有

$$\text{tr}(S_x) = X^T [E(A - EA)^T E(A - EA)] X \quad (3-4)$$

定义如下矩阵:

$$G_i = E[(A - EA)^T (A - EA)] \quad (3-5)$$

矩阵  $G_i$  称作图像协方差矩阵(散度矩阵), 易知  $G_i$  是一个  $n \times n$  的非负定矩阵。我们可以用训练样本图像来直接估计矩阵  $G_i$ , 设总共有  $M$  幅训练样本图像, 第  $i$  幅图像用  $m \times n$  的矩阵  $A_i$  ( $i = 1, 2, \dots, M$ ) 来表示, 训练样本图像的均值图像为

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i \quad (3-6)$$

则可用下式估计图像协方差矩阵  $G_i$

$$G_i = \frac{1}{M} \sum_{i=1}^M (A_i - \bar{A})^T (A_i - \bar{A}) \quad (3-7)$$

因此，准则函数又可写作：

$$J(\mathbf{X}) = \mathbf{X}^T G_i \mathbf{X} \quad (3-8)$$

其中， $\mathbf{X}$ 是归一化的列向量，称为最优投影方向。

我们的目的在于通过最大化准则函数  $J(\mathbf{X})$ ，来寻找一个最优的投影方向  $\mathbf{X}$ ，使得所有的训练样本在  $\mathbf{X}$  方向上投影之后，各投影特征向量尽可能地相互分离，以达到总体散度最大化的效果。

实际上，最优投影方向  $\mathbf{X}_{opt}$  即是图像协方差矩阵  $G_i$  的最大特征值所对应的特征向量，可简要证明如下：

若向量  $\mathbf{X}_{opt}$  是矩阵  $G_i$  对应于特征值  $\lambda_{opt}$  的特征向量，则有

$$G_i \mathbf{X}_{opt} = \lambda_{opt} \mathbf{X}_{opt} \quad (3-9)$$

故

$$\mathbf{X}_{opt}^T G_i \mathbf{X}_{opt} = \lambda_{opt} \mathbf{X}_{opt}^T \mathbf{X}_{opt} \quad (3-10)$$

由于  $\mathbf{X}_{opt}$  是归一化的向量，则有

$$\mathbf{X}_{opt}^T G_i \mathbf{X}_{opt} = \lambda_{opt} \quad (3-11)$$

若  $\lambda_{opt}$  是  $G_i$  的最大特征值，则准则函数  $J(\mathbf{X})$  有最大值  $\lambda_{opt}$ 。

在样本类别数较多的情况下，如果只取一个最优投影方向，分类鉴别能力是远远不够的，通常情况下会选择相互正交且极大化准则函数的一组投影向量  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_d$ ，即：

$$\begin{cases} \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_d\} = \arg \max J(\mathbf{X}) \\ \mathbf{X}_i^T \mathbf{X}_j = 0, i \neq j, i, j = 1, 2, \dots, d \end{cases} \quad (3-12)$$

实际上，投影向量  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_d$  就是图像协方差矩阵  $G_i$  的前  $d$  个最大特征值  $\lambda_1, \lambda_2, \dots, \lambda_d$  所对应特征向量 ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ )。

### 3.2.2 特征提取

给定一组 2DPCA 的最优投影向量  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_d$ ，对一幅样本图像  $\mathbf{A}$ ，令

$$\mathbf{Y}_k = \mathbf{A} \mathbf{X}_k \quad (k = 1, 2, \dots, d) \quad (3-13)$$

可获得一组投影特征向量  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_d$ ，称为样本图像  $\mathbf{A}$  的主元。在 PCA 中，每一个主元是一个系数，但在 2DPCA 中每一个主元则是一个向量。样本图像  $\mathbf{A}$  的

所有主元构成一个  $m \times d$  ( $d \ll m$ ) 的矩阵  $B = [Y_1 \ Y_2 \ \dots \ Y_d]$ , 通常称之为样本图像  $A$  的特征矩阵或特征图像。

### 3.2.3 模式分类

经过 2DPCA 的特征提取过程之后, 每一幅人脸图像都对应一个特征矩阵, 可以使用基于欧氏距离的最近邻分类器进行模式分类, 分类过程的描述如下:

对任意两幅人脸图像  $A_i$  和  $A_j$ , 分别对应于特征矩阵  $B_i = [Y_1^{(i)} \ Y_2^{(i)} \ \dots \ Y_d^{(i)}]$  和  $B_j = [Y_1^{(j)} \ Y_2^{(j)} \ \dots \ Y_d^{(j)}]$ , 则特征矩阵  $B_i$  与  $B_j$  之间的距离定义为:

$$d(B_i, B_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|_2 \quad (3-14)$$

其中,  $\|Y_k^{(i)} - Y_k^{(j)}\|_2$  表示两个主元向量  $Y_k^{(i)}$  与  $Y_k^{(j)}$  之间的欧氏距离。

对所有的训练样本  $A_1, A_2, \dots, A_M$  ( $M$  为训练样本总数), 每一个样本都属于某个确定的类别  $\omega_k$  ( $k = 1, 2, \dots, C$ ), 它们对应的特征矩阵分为  $B_1, B_2, \dots, B_M$ 。对于某个给定的测试图像  $A^*$ , 其对应的特征矩阵为  $B^*$ , 若有  $d(B^*, B_i) = \min_j d(B^*, B_j)$ , 且有  $B_i \in \omega_k$ , 则分类的结果是  $A^* \in \omega_k$ 。

### 3.2.4 图像重构

在特征脸 (Eigenface) 方法中, 只要给定的一组特征脸图像以及某一幅人脸图像的主元, 就能够重构出一幅人脸图像, 类似地, 2DPCA 可以按照下面的方法进行图像重构:

假设  $X_1, X_2, \dots, X_d$  是对应于图像协方差矩阵  $G_i$  的前  $d$  个最大特征值所对应的正交归一的特征向量, 将图像  $A$  向这些投影轴投影, 则可得到相应的主元向量  $Y_k = AX_k$  ( $k = 1, 2, \dots, d$ ), 令  $V = [Y_1 \ Y_2 \ \dots \ Y_d]$ ,  $U = [X_1 \ X_2 \ \dots \ X_d]$ , 得

$$V = AU \quad (3-15)$$

由于  $X_1, X_2, \dots, X_d$  两两正交, 即有  $UU^T = I$ , 易知图像  $A$  的重构图像为

$$\tilde{A} = VU^T = \sum_{k=1}^d Y_k X_k^T \quad (3-16)$$

显然, 图像  $\tilde{A}_k = Y_k X_k^T$  ( $k = 1, 2, \dots, d$ ) 与原始图像  $A$  的尺寸一致, 我们将其称作

原始图像  $A$  的“重构子图像”。因此，原始图像  $A$  的重构图像可以由  $d$  幅“重构子图像”叠加产生，特别地，当主元向量的数目  $d = n$ （ $n$  是矩阵  $G_i$  特征向量的总数）时，有  $\tilde{A} = A$ ，即表明可以根据图像的主元向量无损地重构原始图像，若主元向量的数目  $d < n$ ，则重构图像  $\tilde{A}$  是原始图像  $A$  的近似。

### 3.3 2DPCA 人脸识别的实现过程

基于 PCA 方法的人脸识别系统，一般都包括以下两个过程：

**一、建库过程：**通过训练产生低维人脸空间，抽取一系列人脸图像的低维数据表示建立人脸数据库。

**二、识别过程：**将测试图像的低维数据表示与人脸数据库中的数据进行比对，判定测试图像所属的类别。

2DPCA 人脸识别算法也包括以上两个过程，下面详细叙述其实现。

#### 3.3.1 建库过程

##### (1) 建立训练集

对输入的人脸图像进行相应的预处理，如几何规整、光线补偿、直方图均衡等等，以消除各种外在因素对图像造成的影响，这有利于提高人脸识别率。经过预处理的人脸图像称作标准化人脸图像。

假设任意一幅标准化人脸图像都可以用一个  $m \times n$  的二维矩阵来表示，矩阵的元素表示像素的灰度值。若训练集中人脸图像的总数为  $M$ ，则整个训练集可记作  $\{A_1, A_2, \dots, A_M\}$ ，其中，每一个二维矩阵  $A_i \in R^{m \times n}$  对应于一幅人脸图像。

##### (2) 计算最优投影方向

根据训练集中的二维图像矩阵可以计算出图像协方差矩阵，公式如下：

$$G_i = \frac{1}{M} \sum_{i=1}^M (A_i - \bar{A})^T (A_i - \bar{A}) \quad (3-17)$$

其中， $\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i$  是所有二维图像矩阵的均值矩阵。

然后根据如下准则函数求出最优投影方向：

$$J(X) = X^T G_i X \quad (3-18)$$

其中,  $X$  是归一化的列向量, 称为最优投影方向。在样本类别数较多的情况下, 通常会选择一组相互正交且最大化准则函数的最优投影向量组  $X_1, X_2, \dots, X_d$ , 即满足如下要求:

$$\begin{cases} \{X_1, X_2, \dots, X_d\} = \arg \max J(X) \\ X_i^T X_j = 0, \quad i \neq j, \quad i, j = 1, 2, \dots, d \end{cases} \quad (3-19)$$

在实际计算中, 最优投影向量组  $X_1, X_2, \dots, X_d$  恰好就是图像二维协方差矩阵  $G_i$  的最大的前  $d$  个特征值所对应的特征向量。

### (3) 计算训练集图像的特征矩阵

得到最优投影向量组  $X_1, X_2, \dots, X_d$  之后, 根据如下线性变换公式:

$$Y_k = AX_k \quad (k = 1, 2, \dots, d) \quad (3-20)$$

将训练集中的每一幅标准化图像  $A_i$  ( $i = 1, 2, \dots, M$ ) 往最优投影方向上投影, 得到相应的一组投影特征向量  $Y_1^{(i)}, Y_2^{(i)}, \dots, Y_d^{(i)}$ , 称作图像  $A_i$  的主元向量组, 写成矩阵的形式即是  $B_i = [Y_1^{(i)} \quad Y_2^{(i)} \quad \dots \quad Y_d^{(i)}]$  ( $i = 1, 2, \dots, M$ ),  $B_i$  称作图像  $A_i$  的特征矩阵。

将本步骤得到的特征矩阵  $B_1, B_2, \dots, B_M$  ( $M$  为训练集中图像的总数) 以及对应的个人身份信息保存起来, 就构成了已知人脸图像的特征数据库。

## 3.3.2 识别过程

### (1) 图像预处理

对所有的测试图像都进行与训练过程中相同的预处理操作, 得到相应的标准化人脸图像。

### (2) 计算测试图像的特征矩阵

假设矩阵  $A'$  表示一幅  $m \times n$  维的待识别测试图像, 将  $A'$  往最优投影方向  $X_1, X_2, \dots, X_d$  上投影, 则可得到图像  $A'$  的特征矩阵  $B' = [Y'_1 \quad Y'_2 \quad \dots \quad Y'_d]$ 。

### (3) 判定测试图像所属的类别

对于任意两幅人脸图像  $A_i$  和  $A_j$ , 对应的特征矩阵分别  $B_i = [Y_1^{(i)} \quad Y_2^{(i)} \quad \dots \quad Y_d^{(i)}]$  和  $B_j = [Y_1^{(j)} \quad Y_2^{(j)} \quad \dots \quad Y_d^{(j)}]$ , 则特征矩阵  $B_i$  与  $B_j$  之间的距离定义为:

$$d(B_i, B_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|_2 \quad (3-21)$$

其中,  $\|Y_k^{(i)} - Y_k^{(j)}\|_2$  表示两个主元向量  $Y_k^{(i)}$  与  $Y_k^{(j)}$  之间的欧氏距离。

对所有的训练样本  $A_1, A_2, \dots, A_M$  ( $M$  为训练集中的图像总数), 每一个样本都属于某个确定的类别  $\omega_k$  ( $k = 1, 2, \dots, C$ ) ( $C$  为训练集图像中的总人数), 它们对应的特征矩阵分为  $B_1, B_2, \dots, B_M$ 。对于某个给定的测试图像  $A^*$ , 其对应的特征矩阵为  $B^*$ , 若有  $d(B^*, B_i) = \min_j d(B^*, B_j)$ , 且有  $B_i \in \omega_k$ , 则分类的结果是  $A^* \in \omega_k$ 。

### 3.3.3 2DPCA 与 PCA 的比较

由于 2DPCA 方法可以直接通过二维图像构造去协方差矩阵, 对于尺寸为  $m \times n$  的图像, 2DPCA 方法的协方差矩阵  $G$  维数仅为  $n \times n$ , 远低于 PCA 方法的协方差矩阵维数  $mn \times mn$ , 因而所耗费的运算时间也远低于 PCA 方法。表 3-1 中给出了 2DPCA 方法与 PCA 方法的特征抽取所用时间的比较<sup>[31]</sup> (机器配置: CPU: Pentium III 800MHz, RAM: 256MB), 可以看出 2DPCA 所用时间约为 PCA 的 1/4~1/20。

表 3-1 特征抽取所用时间比较 (ORL 数据库)      单位: 秒

训练样本数 / 类	1	2	3	4	5
PCA(Eigenface)	44.45	89.00	139.36	13.40	304.61
2DPCA	10.76	11.23	12.59	198.95	14.03

在图像的重构上, 2DPCA 方法也是优于 PCA 方法的, 图 3-1 展示了不同主元数目下 2DPCA 的重构图像与 PCA 的重构图像的比较<sup>[31]</sup>, 其中, 5 幅 2DPCA 的重构图像分别选取了  $d = 2, 4, 6, 8, 10$  个主元, 5 幅 PCA 的重构图像分别选取了  $d = 5, 10, 20, 30, 40$  个主元。

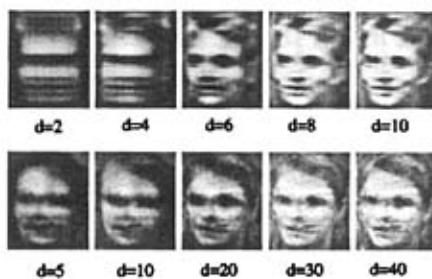


图 3-1 2DPCA 的重构图像 (上) 与 PCA 的重构图像 (下) 的比较

可见，2DPCA 方法的重构图像的质量要优于 PCA 方法的重构图像，这说明 2DPCA 方法的特征抽取比 PCA 方法更加有效，2DPCA 方法的数据降维带来的误差比 PCA 方法要小，最终体现在 2DPCA 方法的人脸识别率高于 PCA 方法。

表 3-2 给出了 2DPCA 方法与 PCA 方法最佳识别率比较<sup>[31]</sup>，括号中的数字表示特征向量（或矩阵）的维数，带 \* 的栏表示 2DPCA 方法的人脸识别率相对于 PCA 方法的人脸识别率有大于 5% 的显著提升。

表 3-2 2DPCA 与 PCA 的最佳识别率比较（ORL 数据库）

训练样本数 / 类	1*	2*	3	4*	5
PCA(Eigenface)	66.9%(39)	84.7%(79)	88.2%(95)	90.8%(60)	93.5%(37)
2DPCA	76.7%(112×2)	89.1%(112×2)	91.8%(112×6)	95.0%(112×5)	96.0%(112×3)

### 3.4 2DPCA 人脸识别的改进方法

#### 3.4.1 2DPCA+PCA

在 PCA 中，每一个主元是一个系数，但在 2DPCA 中，每一个主元则是一个向量，该向量的维数与图像的行数相同。因此，与 PCA 相比，2DPCA 需要更多的系数才能表征一幅图像，这是 2DPCA 的一个较为明显的不足。例如，在表 3-2 中，在达到最高识别率的情况下，2DPCA 的图像特征矩阵的维数 ( $112 \times d$ ) 总是远高于 PCA 的图像特征矩阵的维数。

那么，如何降低 2DPCA 抽取出的图像特征数据的维数呢？文献[31]中提出了一种简单而直接的思路：首先使用 2DPCA 方法对图像进行特征提取，然后再对得到的图像特征矩阵进行 PCA 降维。这样既降低了图像特征数据的维数，又能够保持 2DPCA 的高人脸识别率的优点，文献[31]中将这种方法称作“2DPCA+PCA”。

为了验证“2DPCA+PCA”方法的有效性，文献[31]在 ORL 数据库上进行了如下实验：选择每类 5 幅图像进行训练，然后保留每幅图像前 8 个 2DPCA 的主元向量，即图像特征矩阵的维数为  $112 \times 8$ （这里的 112 是 ORL 数据库中人脸图像的行数），然后再使用 PCA 方法进行数据降维，并用基于欧氏距离的最近邻分类器进行模式分类。该实验结果如图 3-2 所示，可见，在数据维数相同的情况下，“2DPCA+PCA”方法的人脸识别率优于 PCA 方法。

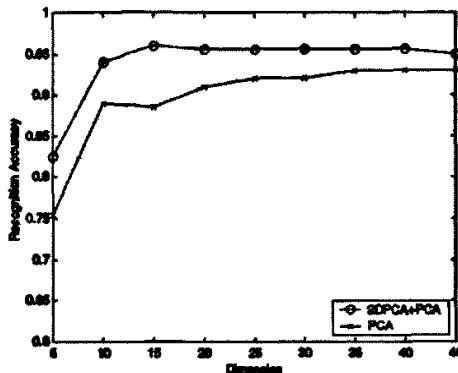


图 3-2 2DPCA+PCA 与 2DPCA 的人脸识别率比较 (ORL 数据库)

### 3.4.2 K2DPCA

核方法是一系列非线性数据处理技术的总称，其共同特征是这些数据处理方法都应用了核映射。核方法的特点是可以把在测量空间无法线性分类的数据变换到特征空间来实现线性分类。

核主元分析，其缩写为 KPCA (Kernel Principal Component Analysis)，是一种广义的 PCA 方法，其基本思想是通过一个非线性变换  $\Phi$ ，将输入数据由测量空间映射到更高维的甚至无穷维的特征空间中，然后在特征空间中进行 PCA 分析。KPCA 方法能够捕捉到输入数据的高阶统计特性，使获得的降维数据能够更有效地表征原始数据，由于通过核技巧 (kernel trick) 隐式地实现了数据映射，因而处理高维数据空间带来的计算开销相当有限。

KPCA 方法已经成功地应用到了人脸识别领域，并获得了比 PCA 人脸识别更好的性能，类似地，将 2DPCA 方法核化也会有助于刻画输入数据的非线性结构，因此文献[71]中提出了 K2DPCA (Kernel 2DPCA) 方法。

在 K2DPCA 方法中，同样是通过隐式的方法实现了非线性的数据映射，但与 KPCA 方法不同的是，K2DPCA 中的映射是对二维图像矩阵每一行实施的，即对图像的每一行实施映射  $\Phi: R^f \rightarrow R^f (f > t)$ ，其中， $t$  是图像行的维数， $f$  可以任意大。特征空间  $R^f$  中的内积运算（点乘）可以很方便地由预定义的核函数（kernel function）来计算，常用的核函数包括：线性核函数、多项式核函数、高斯径向基函数 (RBF) 核函数等。

假设训练集中总共有  $M$  幅  $m \times n$  维的人脸图像  $A_1, A_2, \dots, A_M$ ，所有人脸图像的

均值图像为  $\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i$ ，则  $\hat{A}_i = A_i - \bar{A}$  ( $i=1, 2, \dots, M$ ) 为减去图像均值后的差值图像。令  $\Phi(\hat{A}_i)$  表示第  $i$  幅图像  $\hat{A}_i$  映射到特征空间后的图像， $\Phi(\hat{A}_i^j)$  表示图像  $\Phi(\hat{A}_i)$  的第  $j$  行的向量，则特征空间  $R'$  中的图像协方差矩阵  $C^*$  可由下式计算：

$$C^* = \frac{1}{M} \sum_{i=1}^M \Phi(\hat{A}_i)^T \Phi(\hat{A}_i) \quad (3-22)$$

其中， $\Phi(\hat{A}_i) = [\Phi(\hat{A}_i^1)^T \quad \Phi(\hat{A}_i^2)^T \quad \dots \quad \Phi(\hat{A}_i^m)^T]^T$ ， $m$  为图像的行数。

若  $R'$  是无限维的空间，则矩阵  $C^*$  的维数是  $\text{inf} \times \text{inf}$ ，显然难以直接计算特征值  $\lambda_i$  以及相应的特征向量  $v_i$ ，这里的  $\lambda_i$  与  $v_i$  满足下式：

$$C^* v_i = \lambda_i v_i \quad (3-23)$$

然而，文献[71]中证明了如下定理，使得 K2DPCA 可以通过 KPCA 来实现：

**【定理】** 对二维图像实施 K2DPCA，实质上相当于对所有训练图像的每一行实施 KPCA，即将图像的每一行看作是 KPCA 的一个计算单元。

任意一幅  $m \times n$  维的人脸图像经过 K2DPCA 的特征提取之后，可得到一个  $m \times d$  维的特征矩阵 ( $d \ll n$ )，然后就可以用与 2DPCA 中相似的步骤进行模式分类。

表 3-3 给出了来自文献[71]的 K2DPCA 与 2DPCA 的识别率比较。

表 3-3 K2DPCA 与 2DPCA 的识别率比较 (UMIST 数据库)

训练图像编号	#5, #14	#1, #7, #13	#2, #8, #14	#3, #9, #15	#4, #10, #16	#5, #11, #17	#6, #12, #18
2DPCA	90.3%	91.0%	93.0%	95.0%	95.0%	93.7%	92.3%
K2DPCA	92.7%	94.0%	94.3%	95.7%	97.0%	95.7%	94.0%

### 3.4.3 B2DPCA

文献[71]中证明，2DPCA 方法实质上相当于基于“行分块”的 PCA 方法，即把二维图像的每一行看作是 PCA 的一个输入样本，因此，2DPCA 只利用了图像各行内的相关性信息，而完全抛弃了图像各列之间的相关性信息。为了充分利用图像行与列的相关性信息，文献[71]中提出了 B2DPCA (Bilateral 2DPCA) 方法，与 2DPCA 方法相比，B2DPCA 方法能够有效地去除图像中行与列的冗余信息，因而用于表示图像的特征矩阵的维数要远小于 2DPCA，这解决了 2DPCA 所需存储空间较大的问题，同时，由于 B2DPCA 方法充分地利用了图像行与列的相关性信息，

其人脸识别率要优于 2DPCA 方法。

B2DPCA 的算法如下：

令  $\mathbf{U} \in R^m \times R'$  为左乘投影矩阵， $\mathbf{V} \in R^n \times R'$  为右乘投影矩阵，对一幅  $m \times n$  维的图像  $\mathbf{A}_i$ ，可通过如下双向投影公式获得相应的  $l \times r$  维的图像特征矩阵  $\mathbf{B}_i$ ：

$$\mathbf{B}_i = \mathbf{U}^T \mathbf{A}_i \mathbf{V} \quad (3-24)$$

为了求出最优投影矩阵  $\mathbf{U}_{opt}$  与  $\mathbf{V}_{opt}$ ，我们可以将问题转化为求解图像  $\mathbf{A}_i$  ( $i=1,2,\dots,M$ ) 的最佳逼近  $\mathbf{U}_{opt} \mathbf{B}_i \mathbf{V}_{opt}^T$  的问题，即

$$[\mathbf{U}_{opt}, \mathbf{V}_{opt}] = \arg \min \sum_{i=1}^M \|\mathbf{A}_i - \mathbf{U} \mathbf{B}_i \mathbf{V}^T\|_F^2 \quad (3-25)$$

其中， $M$  是训练集中人脸图像的数目， $\|\cdot\|_F$  是矩阵的 Frobenius 范数，即有

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}.$$

上述最小化问题可以通过迭代的方式来求解，在此不作详细说明，若有需要可参看文献[71]。

任意一幅  $m \times n$  维的人脸图像经过 B2DPCA 的特征提取之后，可得到一个  $l \times r$  维的特征矩阵 ( $l \ll m, r \ll n$ )，然后就可以用与 2DPCA 中相似的步骤进行模式分类。表 3-4 给出了来自文献[71]的 B2DPCA 与 2DPCA 的识别率比较。

表 3-4 B2DPCA 与 2DPCA 的识别率比较 (UMIST 数据库)

训练图像编号	#5, #14	#1, #7, #13	#2, #8, #14	#3, #9, #15	#4, #10, #16	#5, #11, #17	#6, #12, #18
2DPCA	90.3%	91.0%	93.0%	95.0%	95.0%	93.7%	92.3%
B2DPCA	90.7%	91.7%	93.4%	95.3%	95.8%	94.0%	92.8%

## 第4章 基于结构分块的2DPCA 人脸识别方法

### 4.1 引言

2DPCA 人脸识别方法不需要将二维图像转换为一维向量，可以直接利用二维图像矩阵构造协方差矩阵，如果图像是 $m \times n$ 的矩阵，则 2DPCA 的协方差矩阵仅为 $n \times n$ 维，远远小于 PCA 方法中由图像向量构造的 $mn \times mn$ 维的协方差矩阵。2DPCA 方法不但避免了庞大的运算量，而且能获得优于 PCA 方法的人脸识别率。

在人脸识别的实际应用中，有一类问题是不完整人脸图像的识别。对于 PCA、2DPCA 这类基于统计的模式识别方法而言，在统计平均之下，不完整人脸图像的真实部分不但不能突显，反而会被不真实的部分消磨甚至淹没，因而对不完整人脸的识别率是相当低的。

由于人脸图像具有局域特性，如果对图像进行预处理，使得相同的人脸局部特征（如眼睛、嘴巴等）在所有图像中的位置都一致，那么对图像进行分块以后，各个图像块就具有了确定的语义。基于上述考虑，本文提出一种结合了统计模式识别和结构模式识别优点的人脸识别方法，称作基于结构分块的 2DPCA 人脸识别方法。这种方法针对每一个图像块应用 2DPCA 方法，求出其特征矩阵，由于该特征矩阵代表了确定的人脸特征，本文中称之为结构主元。在计算图像相似度的时候，先计算出各图像块对应的相似度，然后用加权和的方法计算出最终的相似度。通过调整各图像块的加权系数，这种方法可应用于各种不完整人脸图像的识别。

本章详细叙述了基于结构分块的 2DPCA 人脸识别方法的基本思想与实现过程，并在 FERET 人脸库上做相关实验，证明本文提出的基于结构分块的 2DPCA 人脸识别方法优于 2DPCA 人脸识别方法。

### 4.2 基于结构分块的 2DPCA 新方法

#### 4.3.1 基本思想

在人脸识别的一些特定应用中，待识别的人脸图像不一定是完整且真实的。例如在公安刑侦业务中，由于罪犯为了躲避追捕往往会遮掩脸部，因而公安机关

通常较难获得完整的人脸照片。再比如说公安机关在做人脸拼图时，目击者通常只对人脸的某些特征印象深刻，对其它特征的印象则较为模糊，因此获得的人脸图像只有部分区域比较真实。对于这一类特殊的人脸图像，通常的人脸识别算法的识别率是相当低的，甚至会判别为非人脸图像，但显然这一类图像为公安机关提供了丰富的破案线索。

考虑到人脸图像具有局域特征，即人脸的局部特征如鼻子、眼睛、嘴巴等等，是最能刻画与表征人与人之间的差异的，如果将输入图像经过预处理，使得所有人脸图像的各个器官都处在相同的位置上，再采用图像分块的方法将每一幅人脸图像分成数目固定的图像块，则根据图像块的排列顺序，可以推断出每个图像块的语义，也即图像块代表了确定的人脸局部特征。

根据上述关于人脸图像局域特性的考虑，参考 2DPCA 人脸识别的思想，本文提出了一种基于结构分块的 2DPCA 人脸识别方法。这种方法先将人脸图像分割为若干个互不重叠的图像块，然后用 2DPCA 方法计算每个图像块的协方差矩阵，进而求出各个图像块的特征矩阵（即主元）。这些主元称为人脸图像的结构主元，因为这些主元尽管仍是统计平均的结果，但却是人脸局部区域的统计平均，反映了人脸的结构特征。人脸图像的全部结构主元，配上表示其重要性的加权系数，再按照一定的顺序排列，就构成了整幅人脸图像的主元。在计算人脸图像相似度的时候，先计算出相同位置图像块的结构主元之间的相似度，然后使用求加权和的方法计算出最终的相似度，最后再使用最近邻分类器进行模式分类。

本文提出的基于结构分块的 2DPCA 人脸识别方法，结合了统计模式识别和结构模式识别的优点，通过调整各图像块的加权系数，可应用于前述各种不完整或部分区域真实的人脸图像的识别。

### 4.3.2 算法描述

基于结构分块的 2DPCA 人脸识别方法的算法描述如下：

#### (1) 图像预处理

首先要对人脸图像进行预处理，包括几何归一化、光线补偿、灰度归一化等操作，经过预处理的人脸图像称作标准化人脸图像。标准化人脸图像的尺寸固定，而且眼睛、嘴巴均处于预先指定的标准位置，这使得经过图像分块之后，相同的人脸局部特征能落入编号相同的图像块内。此外，标准化人脸图像的光线分布被

统一到了标准水平，消除了光线差异对人脸识别的影响。

### (2) 人脸图像分块

对于一幅 $m \times n$ 维的标准化人脸图像，若有 $m = p \times m'$ 与 $n = q \times n'$ ，则可将原图像分割为 $p \times q$ 个互不重叠的图像块，每个图像块的尺寸为 $m' \times n'$ 。显然，为了保证 $p, q, m', n'$ 均为整数，需要图像预处理中几何归一化步骤的配合。以图 4-1(a)为例，将人脸图像经图像预处理转化成尺寸为 $96 \times 80$ 的标准化图像，取图像块的尺寸为 $16 \times 16$ ，则可将原图像分割为 $6 \times 5 = 30$ 个图像块。

### (3) 计算最优投影矩阵

假设训练集中共有 $M$ 幅人脸图像，则经过图像分块步骤以后，可得到 $p \times q$ 组图像块，每组有 $M$ 个尺寸为 $m' \times n'$ 的图像块。对每一组图像块，使用 2DPCA 方法计算其协方差矩阵，进而求出其最优投影矩阵，即得到该组图像块对应的特征子空间。以图 4-1 为例，经过图(a)的分块操作，可得到图(b)所示 30 组图像块，每组有 $M$ 个尺寸为 $16 \times 16$ 的图像块，然后用 2DPCA 方法计算各组图像块的最优投影矩阵，即如图(c)所示的 $16 \times d$ 维的矩阵，一共有 30 个。

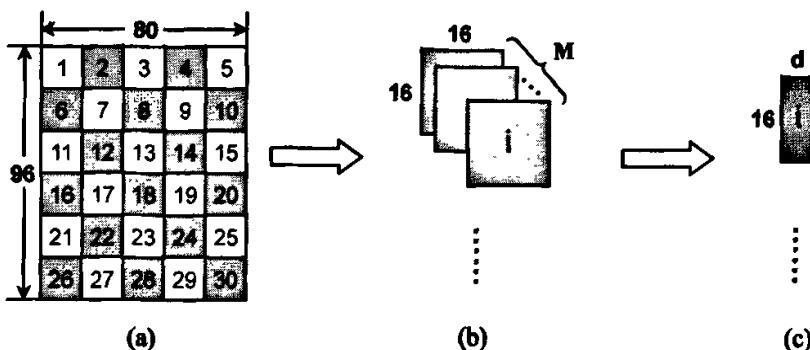


图 4-1 基于结构分块的 2DPCA 方法示意图 ( $i=1,2,\dots,30$ )

### (4) 特征提取

计算出各组图像块的特征子空间后，对任意一幅人脸图像均可通过步骤(1)、(2)，将人脸图像分割成 $p \times q$ 个图像块，然后可以使用 2DPCA 方法中的特征投影方法，提取人脸图像的各图像块对应的特征矩阵。本文将这些特征矩阵称作人脸图像的结构主元，即每一幅人脸图像可由 $p \times q$ 个结构主元来表征。

### (5) 模式分类

对待识别的人脸图像，通过步骤(1)(2)(4)计算出它的全部 $p \times q$ 个结构主元，再分别计算出各结构主元与数据库中某已知人脸图像的相同编号的结构主元之间的相似度，然后根据实际需要确定各图像块的加权系数，并计算出 $p \times q$ 个相似度的加权和作为最终的相似度。将数据库中与待识别人脸图像相似度最大的人脸图像，判定为与待识别人脸图像最相似的图像，该最相似图像对应的身份则是待识别人脸图像的人物身份。

下面详细地介绍基于结构分块的2DPCA 人脸识别方法的实现过程，特别地，设定标准化人脸图像的尺寸为 $96 \times 80$ ，将其分割为 30 个互不重叠的图像块，每个图像块的尺寸为 $16 \times 16$ 。

### 4.3.3 建库过程

#### (1) 建立训练集

建立训练集即是将输入图像经过图像预处理，转化成为标准化人脸图像，并保存到数据库中的过程，图像预处理主要包括如下几个步骤：

##### ①几何归一化

几何归一化对于基于结构分块的2DPCA 人脸识别方法来说是非常重要的，经过几何归一化后图像的尺寸固定，而且眼睛、嘴巴均处于预先指定的标准位置，这样，经过图像分块之后，可以保证编号相同的图像块对应于相同的人脸局部特征。几何归一化的具体实施方法如下：

我们设定标准化图像的尺寸为 $96 \times 80$ ，右眼中心点在 $(16, 24)$ ，左眼中心点在 $(64, 24)$ ，嘴巴中心点在 $(40, 72)$ 。对任意一幅给定的人脸图像，假设其右眼中心点在 $(x_1, y_1)$ ，左眼中心点在 $(x_2, y_2)$ ，嘴巴中心点在 $(x_3, y_3)$ ，令

$$X = 64 - 16 = 48 \quad (4-1)$$

$$Y = 72 - 24 = 48 \quad (4-2)$$

几何扩张系数分别如下：

$$C_x = X / (x_2 - x_1) \quad (4-3)$$

$$C_y = Y / (y_3 - y_1) \quad (4-4)$$

将图像的横坐标按照 $C_x$ 的比例，纵坐标按照 $C_y$ 的比例进行放大( $C > 1$ )或者缩小( $C < 1$ )。然后，将图像的原点定在 $(x_1 - 16, y_1 - 24)$ ，切割出 $96 \times 80$ 的图像，

并抛弃多余的部分。

图 4-2 是几何归一化的示意图，其中，图(a)为检测出眼睛和嘴巴位置的原始图像，图(b)为缩放后图像的裁切示意图，图(c)为几何归一化后的人脸图像。

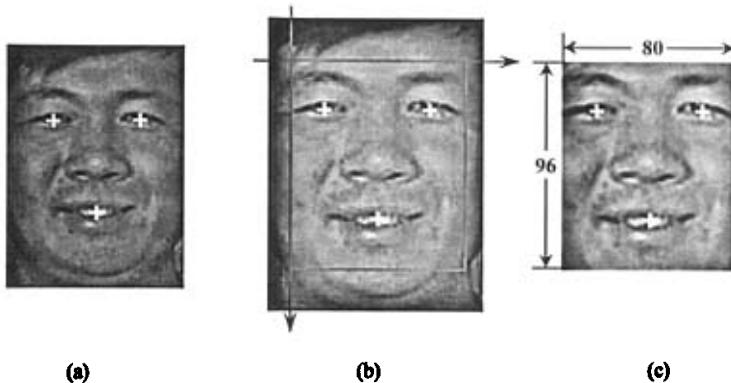


图 4-2 几何归一化示意图

### ② 直方图均衡

直方图均衡可以看作是一种简单的光线补偿处理，可以提高图像的对比度，扩大灰度动态范围，使图像的灰度层次更加分明、细节更加清晰，而突显图像的局部细节则是有利于提高 2DPCA 方法的人脸识别率的。

### ③ 灰度归一化

灰度归一化使得所有人脸图像的灰度平均值处于同一水平，这对基于统计分析的 2DPCA 人脸识别方法也是必要的，可以避免由于图像灰度差异过大而造成的误差。灰度归一化具体实施方法如下：

选取一张灰度适中的人脸图像作为灰度归一化的标准图，计算其灰度平均值：

$$W = \bar{\sigma} = \frac{1}{N} \sum_{i=1}^N \sigma_i \quad (4-5)$$

其中， $\sigma_i$  代表人脸图像中第  $i$  个像素的灰度值， $N$  为人脸图像的像素总数。

然后计算出需要归一化的图像的灰度平均值  $\omega$ ，则灰度调整比例为：

$$C_w = \frac{W}{\omega} \quad (4-6)$$

将该图像中所有像素的灰度值按照  $C_w$  的比例进行调整，即完成了灰度归一化。

## (2) 人脸图像分块

如图 4-3 所示，将经过预处理的尺寸为  $96 \times 80$  的标准化人脸图像分割为 30 个非重叠的图像块，每个图像块的尺寸为  $16 \times 16$ 。由于在标准化图像中，眼睛、嘴巴均处于预先指定的标准位置，经过图像分块之后，可以根据图像块的排列顺序来推断每个图像块的语义，也即图像块代表了确定的人脸局部特征。如图 4-3 中的示例，编号为 6 和 7 的图像块表示右眼，编号为 9 和 10 的图像块表示左眼，编号为 22、23、24、27、28、29 的图像块则表示嘴巴（包括下巴），这样在识别阶段计算相似度的时候，为各图像块赋予不同的加权系数就具有了结构语义上的依据。

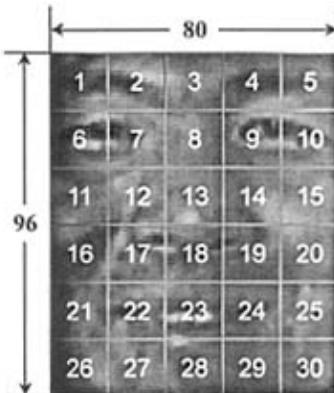


图 4-3 人脸图像分块示意图

假设训练集中  $M$  幅人脸图像，定义矩阵  $A_i^k$  ( $i=1,2,\dots,M$ ;  $k=1,2,\dots,30$ ) 为第  $i$  幅人脸图像的第  $k$  个图像块所对应的二维矩阵，则每一幅图像可用 30 个二维矩阵来表示： $A_i^1, A_i^2, \dots, A_i^{30}$  ( $i=1,2,\dots,M$ )，这些二维矩阵的维数都是  $16 \times 16$ 。

### (3) 计算最优投影矩阵

经过上述步骤(1)、(2)以后，可得到 30 组图像块，每组有  $M$  个尺寸为  $16 \times 16$  的图像块，对每一组图像块，使用 2DPCA 方法计算其协方差矩阵，进而求出图像块对应的最优投影矩阵，即得到该图像块对应的特征子空间，具体实施方法如下：

对第  $k$  组 ( $k=1,2,\dots,30$ ) 图像块，其图像协方差矩阵  $G_i^k$  的计算方法如下：

$$G_i^k = \frac{1}{M} \sum_{m=1}^M (A_i^k - \bar{A}^k)^T (A_i^k - \bar{A}^k) \quad (4-7)$$

其中， $M$  为训练集中人脸图像的总数，矩阵  $\bar{A}^k$  为第  $k$  组图像块的均值，即

$$\bar{A}^k = \frac{1}{M} \sum_{i=1}^M A_i^k \quad (4-8)$$

计算出协方差矩阵  $G_i^k$  后，可根据以下准则函数求出最优投影方向：

$$J(X^k) = (X^k)^T G_i^k (X^k) \quad (4-9)$$

其中， $X^k$  是归一化的列向量，称为最优投影方向。最大化准则函数  $J(X^k)$  的意义是：第  $k$  个图像块在  $X^k$  方向上投影后得到的特征向量的类间分散程度最大。实际上，最优投影方向  $X^k$  即是协方差矩阵  $G_i^k$  的最大特征值所对应的特征向量。

在样本类别数较多的情况下，如果只取一个最优投影方向，分类鉴别能力是远远不够的，通常情况下会选择相互正交且极大化准则函数的一组投影向量  $X_1^k, X_2^k, \dots, X_d^k$ ，即满足如下要求：

$$\begin{cases} \{X_1^k, X_2^k, \dots, X_d^k\} = \arg \min J(X^k) \\ (X_i^k)^T (X_j^k) = 0, \quad i \neq j, \quad i, j = 1, 2, \dots, d \end{cases} \quad (k = 1, 2, \dots, 30) \quad (4-10)$$

在实际计算中，最优投影向量组  $X_1^k, X_2^k, \dots, X_d^k$  恰好就是协方差矩阵  $G_i^k$  的最大的前  $d$  个特征值所对应的特征向量，记成矩阵形式则是  $X^k = [X_1^k \quad X_2^k \quad \dots \quad X_d^k]$ ，称作最优投影矩阵。

通过上述计算步骤，我们就得到 30 个最优投影矩阵  $X^k (k = 1, 2, \dots, 30)$ ，即获得了 30 组图像块对应的特征子空间。

#### (4) 计算训练集图像的结构主元

设矩阵  $A_i^k (i = 1, 2, \dots, M; k = 1, 2, \dots, 30)$  表示训练集中第  $i$  幅人脸图像的第  $k$  个图像块所对应的二维矩阵，将其往最优投影方向上进行投影：

$$Y_{(i,l)}^k = A_i^k X_l^k \quad (i = 1, 2, \dots, M; k = 1, 2, \dots, 30; l = 1, 2, \dots, d) \quad (4-11)$$

则可以获得一组投影特征向量  $Y_{(i,1)}^k, Y_{(i,2)}^k, \dots, Y_{(i,d)}^k$ ，称为图像块  $A_i^k$  的主元向量，一个图像块的所有主元向量构成一个  $16 \times d$  的矩阵  $B_i^k = [Y_{(i,1)}^k \quad Y_{(i,2)}^k \quad \dots \quad Y_{(i,d)}^k]$ ，称为图像块  $A_i^k$  的特征矩阵，由于  $B_i^k$  反映了人脸的局部结构特征，因此将其称作人脸图像的结构主元。

训练集中第  $i$  幅人脸图像  $A_i$  的全部结构主元为  $B_i^1, B_i^2, \dots, B_i^{30} (i = 1, 2, \dots, M)$ ，将训练集中所有人脸图像的结构主元，以及对应的个人身份信息保存起来，就构成了已知人脸图像的特征数据库。

### 4.3.4 识别过程

#### (1) 图像预处理与图像分块

对每一幅待识别的测试图像，采用训练过程中步骤(1)、(2)的方法，进行图像预处理，得到尺寸为 $96 \times 80$ 的标准化图像，并分割成30个 $16 \times 16$ 的非重叠图像块。如前所述，各个图像块代表了确定的人脸局部特征。

#### (2) 计算测试图像的结构主元

假设矩阵 $A^*$ 表示一幅 $m \times n$ 维的待识别测试图像， $A^{k*}$  ( $k=1,2,\dots,30$ ) 表示图像 $A^*$ 的第 $k$ 个图像块，将 $A^{k*}$ 往最优投影方向上投影，即：

$$Y_i^{k*} = A^{k*} X_i^k \quad (k=1,2,\dots,30; i=1,2,\dots,d) \quad (4-12)$$

则可获得一组投影特征向量 $Y_1^{k*}, Y_2^{k*}, \dots, Y_d^{k*}$ ，这些投影特征向量构成一个 $16 \times d$ 的矩阵 $B^{k*} = [Y_1^{k*} \ Y_2^{k*} \ \dots \ Y_d^{k*}]$ ，称作图像 $A^*$ 的第 $k$ 个结构主元。

这样，我们就得了测试图像 $A^*$ 的30个结构主元 $B^1, B^2, \dots, B^{30}$ 。

#### (3) 判定测试图像所属的类别

将测试图像 $A^*$ 的结构主元 $B^{k*} = [Y_1^{k*} \ Y_2^{k*} \ \dots \ Y_d^{k*}]$  ( $k=1,2,\dots,30$ ) 与已知人脸图像数据库中保存的各个人脸图像的结构主元 $B_i^k = [Y_{(i,1)}^k \ Y_{(i,2)}^k \ \dots \ Y_{(i,d)}^k]$  ( $i=1,2,\dots,M$ ,  $k=1,2,\dots,30$ ) 进行比较，依照下式计算测试图像 $A^*$ 的第 $k$ 个图像块与数据库中第 $i$ 幅人脸图像 $A_i$ 的第 $k$ 个图像块的相似度：

$$T_i^k = \sum_{l=1}^d \|Y_l^{k*} - Y_{(i,l)}^k\|_2 \quad (i=1,2,\dots,M; k=1,2,\dots,30) \quad (4-13)$$

其中， $\|\cdot\|_2$ 表示两个向量的欧氏距离。

两幅人脸图像 $A^*$ 与 $A_i$ 的相似度为各个图像块对应相似度的加权和，即：

$$d(A^*, A_i) = \frac{1}{W} \sum_{k=1}^{30} w_k \cdot T_i^k \quad (i=1,2,\dots,M) \quad (4-14)$$

其中， $w_k$ 为第 $k$ 个图像块的加权系数，可以根据实际需要确定， $W = \sum_{k=1}^{30} w_k$  是加权系数的归一化因子。

由于图像块代表了确定的人脸局部特征，因此在计算相似度的时候，可以根据各局部特征的重要性来确定各图像块的加权系数。一般而言，对于正面且较为

清晰的人脸图像，加权系数可设为  $w_k = 1 (k=1,2,\dots,30)$ ，即所有图像块的重要性都相同。在某些特定情况下，如经由拼图产生的人脸图像，可能眼睛的真实程度较高，则可以提高眼睛对应的图像块的加权系数，如令  $w_i = 5, w_k = 1 (i=6,7,9,10, k=1,2,\dots,30 \text{ 且 } k \neq i)$ 。如果待识别人脸存在被遮挡的情况，则可将被遮挡区域对应图像块的加权系数设为 0，其余图像块的加权系数设为 1。

对所有的训练样本  $A_1, A_2, \dots, A_M$  ( $M$  为训练集图像的总数)，每一个样本都属于某一个确定的类别  $\omega_k (k=1,2,\dots,C)$  ( $C$  为训练集中的总人数)，对于一幅给定的测试图像  $A^*$ ，若有  $d(A^*, A_i) = \min_j d(A^*, A_j)$ ，且有  $A_i \in \omega_k$ ，则分类的结果是  $A^* \in \omega_k$ 。

### 4.3 实验结果与分析

本章的实验是在 FERET 人脸库上实现的，选取选择库中的 53 人，每人 6 幅正面人脸图像，共 318 幅人脸图像组成训练和测试人脸库，包括了不同光线、表情等情况下的人脸图像。原始图像的尺寸为  $112 \times 92$ ，经过预处理的标准化人脸图像的尺寸为  $96 \times 80$ ，将图像分割成 30 个互不重叠的图像块，每个图像块的尺寸为  $16 \times 16$ 。此外，标准化人脸图像的光线分布被统一到了标准水平，消除了光线差异对人脸识别的影响。图 4-4 给出了原始图像与标准化人脸图像的比较。



图 4-4 原始图像（上）与标准化人脸图像（下）的比较（FERET 人脸库）

#### 4.3.1 实验一

实验一比较了 2DPCA 方法与本文提出的基于结构分块的 2DPCA 方法的人脸识别率，后者在计算最终相似度的时候，将加权系数设为  $w_k = 1 (k=1,2,\dots,30)$ ，即认为所有图像块具有相同的重要性。

## (1) 第一组实验

随机地选取库中 2 幅图像 / 人用于训练, 余下的 4 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 2 = 106$  个, 测试人数: 53 人, 测试样本数:  $53 \times 4 = 212$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 4-1 识别率比较 (53×2 个训练样本, 53×4 个测试样本, FERET 人脸库)

特征向量数目 \ 识别率	2DPCA	基于结构分块的 2DPCA
1	77.48%	83.35%
2	79.58%	83.71%
3	81.87%	84.12%
4	83.51%	84.37%
5	83.82%	84.78%
6	83.69%	84.78%
7	83.95%	85.06%
8	84.48%	85.13%
9	85.02%	85.19%
10	85.27%	85.33%

下图是在第一组的实验条件下, 2DPCA 方法与基于结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

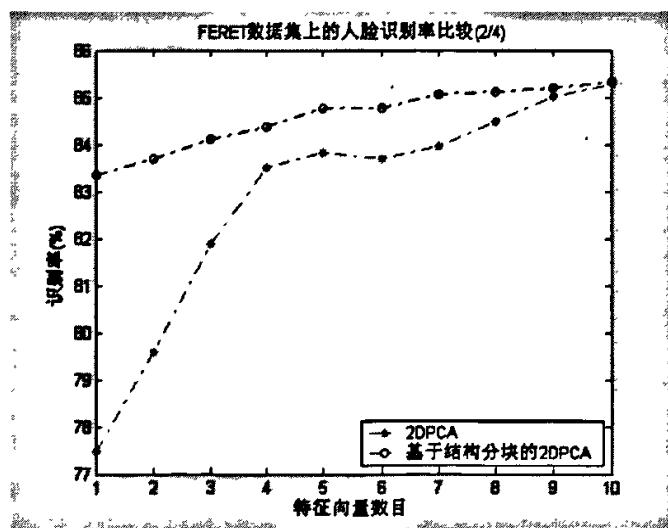


图 4-5 识别率比较 (53×2 个训练样本, 53×4 个测试样本, FERET 人脸库)

## (2) 第二组实验

随机地选取库中 3 幅图像 / 人用于训练, 余下的 3 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 3 = 159$  个, 测试人数: 53 人, 测试样本数:  $53 \times 3 = 159$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 4-2 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

特征向量数目 识别率	2DPCA	基于结构分块的 2DPCA
1	83.21%	87.82%
2	84.36%	88.01%
3	86.39%	88.55%
4	87.48%	88.78%
5	87.95%	89.01%
6	87.88%	89.18%
7	88.24%	89.39%
8	88.64%	89.41%
9	89.14%	89.66%
10	89.33%	89.62%

下图是在第二组的实验条件下, 2DPCA 方法与基于结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

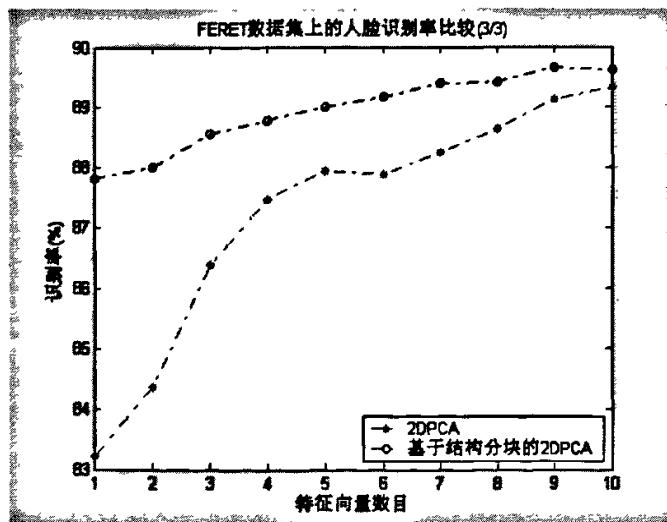


图 4-6 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

### (3) 第三组实验

随机地选取库中 4 幅图像 / 人用于训练, 余下的 2 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 4 = 212$  个, 测试人数: 53 人, 测试样本数:  $53 \times 2 = 106$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 4-3 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

特征向量数目 \ 识别率	2DPCA	基于结构分块的 2DPCA
1	87.30%	91.19%
2	88.77%	91.13%
3	89.75%	91.60%
4	90.28%	91.86%
5	90.79%	92.20%
6	90.82%	92.52%
7	91.32%	92.52%
8	91.67%	92.74%
9	92.20%	92.89%
10	92.26%	93.02%

下图是在第三组的实验条件下, 2DPCA 方法与基于结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

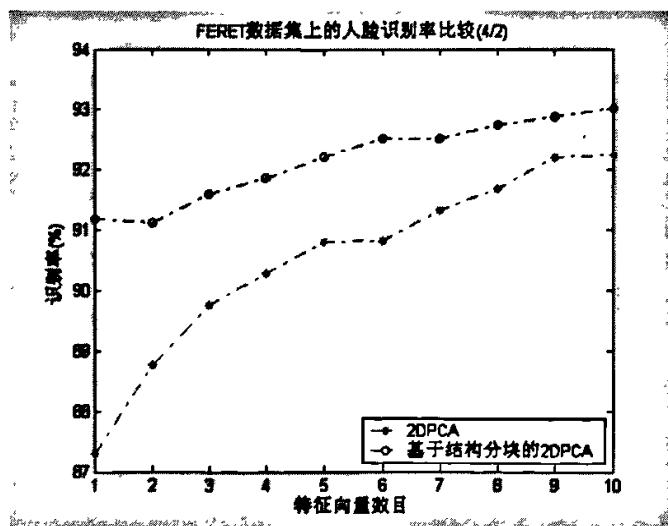


图 4-7 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

从上面三组实验的结果可以看出，在选取相同个数的2DPCA特征向量的情况下，基于结构分块的2DPCA方法的比2DPCA方法的人脸识别率要高，特别是当特征向量数目较少的时候，基于结构分块的2DPCA方法的优势更明显，人脸识别率约有3%~7%的提升。

考虑到训练样本和测试样本的选择不同，会使达到最佳识别率时的特征向量数目 $d$ 产生波动，因此，还需要比较2DPCA方法与基于结构分块的2DPCA方法的最佳识别率。

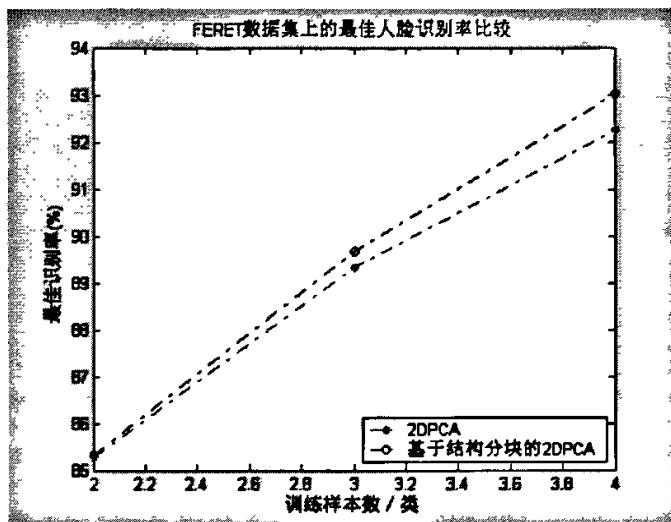


图4-8 最佳识别率比较（训练样本数 / 类  $k$ ,  $2 \leq k \leq 4$ , FERET 人脸库）

图4-8是在 $2 \leq k \leq 4$ ( $k$ 为训练样本数 / 类)的条件下，2DPCA方法与基于结构分块的2DPCA方法的最佳识别率的直观比较，可见，基于结构分块的2DPCA方法在识别率上总是优于2DPCA方法。

### 4.3.2 实验二

实验一比较了基于结构分块的2DPCA方法与基于加权结构分块的2DPCA方法的人脸识别率，二者的区别在于计算最终相似度的时候加权系数设置上的不同。前者将加权系数设为 $w_k = 1$  ( $k = 1, 2, \dots, 30$ )，即认为所有图像块具有相同的重要性；后者将对应于眼睛区域的图像块的加权系数设为3(即 $w_{6,7,9,10} = 3$ )，将对应于嘴巴(包括下巴)区域的图像块的加权系数设为2(即 $w_{22,23,24,27,28,29} = 2$ )，将其余图像

块的加权系数设为 1，即认为眼睛与嘴巴的重要性较人脸的其它部分要高。

### (1) 第一组实验

随机地选取库中 2 幅图像 / 人用于训练，余下的 4 幅图像 / 人则用于测试，即训练人数：53 人，训练样本数： $53 \times 2 = 106$  个，测试人数：53 人，测试样本数： $53 \times 4 = 212$  个。采用随机抽取多次建库的方式重复 30 次实验，取识别率的平均值作为最终识别率，下表给出了实验的结果。

表 4-4 识别率比较 (53×2 个训练样本, 53×4 个测试样本, FERET 人脸库)

识别率 特征向量数目	基于结构分块的 2DPCA	基于加权结构分块的 2DPCA
1	83.35%	83.49%
2	83.71%	84.12%
3	84.12%	84.40%
4	84.37%	84.94%
5	84.78%	85.35%
6	84.78%	85.69%
7	85.06%	86.23%
8	85.13%	86.56%
9	85.19%	86.73%
10	85.33%	86.71%

下图是在第一组的实验条件下，基于结构分块的 2DPCA 方法与基于加权结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线：

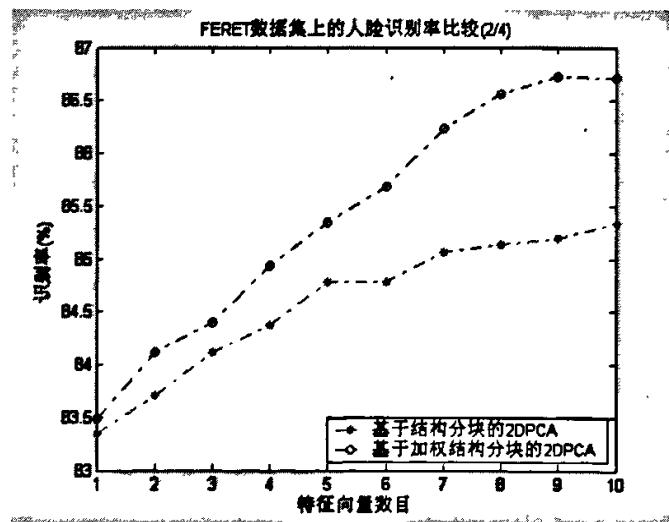


图 4-9 识别率比较 (53×2 个训练样本, 53×4 个测试样本, FERET 人脸库)

## (2) 第二组实验

随机地选取库中 3 幅图像 / 人用于训练, 余下的 3 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 3 = 159$  个, 测试人数: 53 人, 测试样本数:  $53 \times 3 = 159$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 4-5 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

特征向量数目 \ 识别率	基于结构分块的 2DPCA	基于加权结构分块的 2DPCA
1	87.82%	87.44%
2	88.01%	88.18%
3	88.55%	89.14%
4	88.78%	89.43%
5	89.01%	89.77%
6	89.18%	90.06%
7	89.39%	90.44%
8	89.41%	90.90%
9	89.66%	91.01%
10	89.62%	91.05%

下图是在第二组的实验条件下, 基于结构分块的 2DPCA 方法与基于加权结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

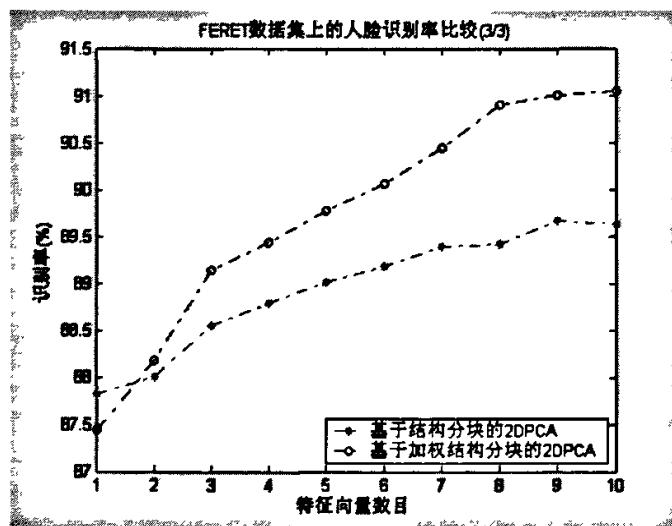


图 4-10 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

### (3) 第三组实验

随机地选取库中 4 幅图像 / 人用于训练, 余下的 2 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 4 = 212$  个, 测试人数: 53 人, 测试样本数:  $53 \times 2 = 106$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 4-6 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

特征向量数目 \ 识别率	基于结构分块的 2DPCA	基于加权结构分块的 2DPCA
1	91.19%	90.88%
2	91.13%	91.32%
3	91.60%	91.86%
4	91.86%	92.14%
5	92.20%	93.08%
6	92.52%	93.46%
7	92.52%	93.87%
8	92.74%	94.12%
9	92.89%	94.12%
10	93.02%	94.18%

下图是在第三组的实验条件下, 基于结构分块的 2DPCA 方法与基于加权结构分块的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

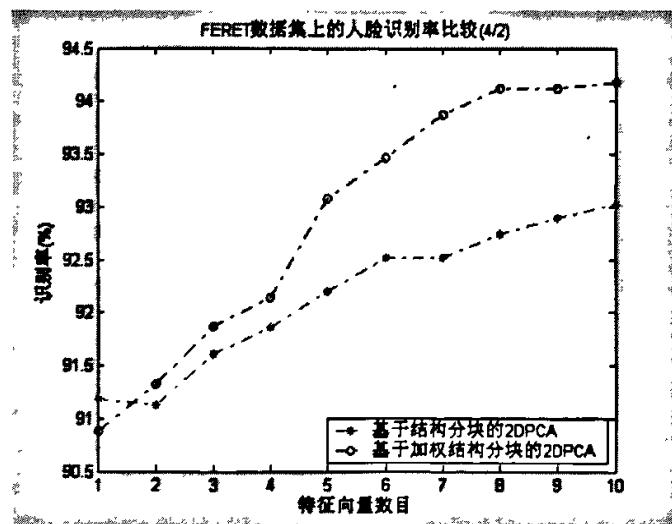


图 4-11 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

从上面三组实验的结果可以看出，在选取相同个数的2DPCA特征向量的情况下，基于加权结构分块的2DPCA方法的比基于结构分块的2DPCA方法的人脸识别率要高（除去特征向量数目为1时的例外），特别是当特征向量大于5的时候，基于加权结构分块的2DPCA方法的优势更明显。

考虑到训练样本和测试样本的选择不同，会使达到最佳识别率时的特征向量数目 $d$ 产生波动，因此，还需要比较基于加权结构分块的2DPCA方法与基于结构分块的2DPCA方法的最佳识别率。

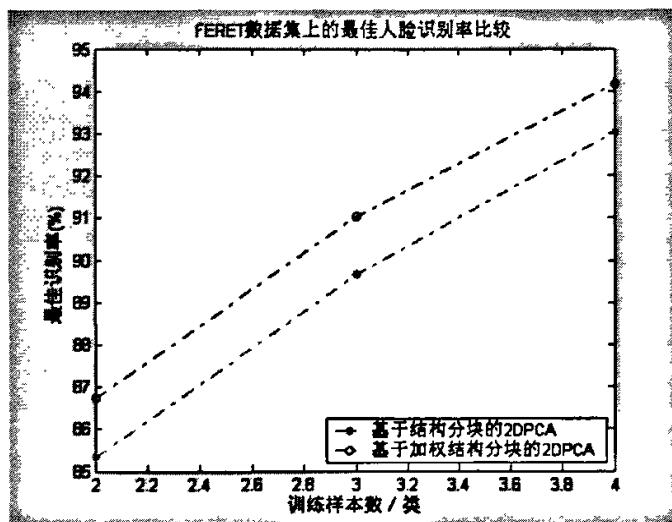


图 4-12 最佳识别率比较（训练样本数 / 类  $k$ ,  $2 \leq k \leq 4$ , FERET 人脸库）

图 4-12 是在  $2 \leq k \leq 4$  ( $k$  为训练样本数 / 类) 的条件下，基于结构分块的2DPCA方法与基于加权结构分块的2DPCA方法的最佳识别率的直观比较，可见，基于加权结构分块的2DPCA方法在识别率上总是优于基于结构分块的2DPCA方法。

## 第 5 章 基于块内相关性的 2DPCA 人脸识别方法

### 5.1 引言

2DPCA 是人脸识别中的一种新方法，可以直接对二维图像进行处理，而不需要事先将图像转化为一维向量，从而大大地降低了处理的维数，并能取得比 PCA 方法更好的识别效果。

通过对 2DPCA 的数学分析可以发现，2DPCA 方法是利用图像同一行像素之间的相关性来构造协方差矩阵的，其实质是一种特殊的分块 PCA 方法<sup>[71][72]</sup>，即每个分块是图像矩阵的一行。可见，2DPCA 方法只利用了图像矩阵同一行像素之间的相关性信息，完全抛弃了行与行之间像素的相关性信息，然而对于人脸图像来说，图像矩阵同一行像素之间的相关性没有一个局部块像素之间的相关性强，而且人脸图像一个局部块通常具有特定的语义，如鼻子、眼睛、嘴巴等等。

基于上述考虑，本文在 2DPCA 人脸识别方法的基础上，提出了基于块内相关性的 2DPCA 人脸识别方法。这种方法首先对人脸图像进行分块，然后把图像块的像素转化为行向量，再把这些行向量按图像块的顺序重排为新的图像矩阵，最后，对这些新的图像矩阵采用 2DPCA 方法进行人脸识别。基于块内相关性的 2DPCA 人脸识别方法既保留了 2DPCA 方法低计算复杂度的优点，又能够充分地利用图像行与列像素之间的相关性信息，克服了 2DPCA 方法的不足，能够取得比 2DPCA 方法更高的人脸识别率。

本章详细叙述了基于块内相关性的 2DPCA 人脸识别方法的基本思想与实现过程，并在 FERET 人脸库上做相关实验，证明本文提出的基于块内相关性的 2DPCA 人脸识别方法优于 2DPCA 人脸识别方法。

### 5.2 2DPCA 方法的实质

为了研究 2DPCA 方法的实质，以及 2DPCA 方法与 PCA 方法的区别与联系，下面我们将对 2DPCA 方法进行数学分析。

假设训练集中共有  $M$  幅  $m \times n$  维的人脸图像，矩阵  $\mathbf{A}_i$  ( $i = 1, 2, \dots, M$ ) 表示第  $i$  幅

人脸图像的二维矩阵，向量  $\mathbf{A}_i^j$  ( $i=1,2,\dots,M$ ;  $j=1,2,\dots,m$ ) 表示矩阵  $\mathbf{A}_i$  的第  $j$  行，则矩阵  $\mathbf{A}_i$  可以写成：

$$\mathbf{A}_i = \begin{bmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \cdots & a_{1n}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & \cdots & a_{2n}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}^{(i)} & a_{m2}^{(i)} & \cdots & a_{mn}^{(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_i^1 \\ \mathbf{A}_i^2 \\ \vdots \\ \mathbf{A}_i^m \end{bmatrix} \quad (i=1,2,\dots,M) \quad (5-1)$$

其中，

$$\mathbf{A}_i^j = [a_{j1}^{(i)} \ a_{j2}^{(i)} \ \cdots \ a_{jn}^{(i)}] \quad (i=1,2,\dots,M; j=1,2,\dots,m) \quad (5-2)$$

假设  $E(\mathbf{A})$  表示所有训练样本的均值矩阵，向量  $E(\mathbf{A})^j$  表示矩阵  $E(\mathbf{A})$  的第  $j$  行，再令矩阵  $\hat{\mathbf{A}}_i = \mathbf{A}_i - E(\mathbf{A})$ ，向量  $\hat{\mathbf{A}}_i^j = \mathbf{A}_i^j - E(\mathbf{A})^j$ ，则 2DPCA 方法图像协方差矩阵的样本估计为：

$$\begin{aligned} G_i &= \frac{1}{M} \sum_{i=1}^M [\mathbf{A}_i - E(\mathbf{A})]^T [\mathbf{A}_i - E(\mathbf{A})] = \frac{1}{M} \sum_{i=1}^M \hat{\mathbf{A}}_i^T \hat{\mathbf{A}}_i \\ &= \frac{1}{M} \sum_{i=1}^M \left[ (\hat{\mathbf{A}}_i^1)^T \ (\hat{\mathbf{A}}_i^2)^T \ \cdots \ (\hat{\mathbf{A}}_i^m)^T \right] \begin{bmatrix} \hat{\mathbf{A}}_i^1 \\ \hat{\mathbf{A}}_i^2 \\ \vdots \\ \hat{\mathbf{A}}_i^m \end{bmatrix} \\ &= \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^m (\hat{\mathbf{A}}_i^j)^T (\hat{\mathbf{A}}_i^j) \\ &= \sum_{j=1}^m G_i^{(j)} \end{aligned} \quad (5-3)$$

其中，

$$G_i^{(j)} = \frac{1}{M} \sum_{i=1}^M (\hat{\mathbf{A}}_i^j)^T (\hat{\mathbf{A}}_i^j) = \frac{1}{M} \sum_{i=1}^M [\mathbf{A}_i^j - E(\mathbf{A})^j]^T [\mathbf{A}_i^j - E(\mathbf{A})^j] \quad (5-4)$$

可以看出，矩阵  $G_i^{(j)}$  ( $j=1,2,\dots,m$ ) 恰好就是人脸图像  $\mathbf{A}_i$  的第  $j$  个行向量  $\mathbf{A}_i^j$  的协方差矩阵，其维数为  $n \times n$ ，而人脸图像  $\mathbf{A}_i$  的图像协方差矩阵  $G_i$  就是  $\mathbf{A}_i$  的各个行向量的协方差矩阵的和。可见，2DPCA 方法只利用了图像矩阵同一行像素之间的相关性信息，完全抛弃了行与行之间像素的相关性信息。

现在考虑 PCA 方法的情况，在进行特征提取之前，首先要将图像矩阵  $\mathbf{A}_i$  ( $i=1,2,\dots,M$ ) 转化为一维列向量  $\mathbf{X}_i$ ：

$$\mathbf{X}_i = [a_{11}^{(i)} \ a_{12}^{(i)} \ \cdots \ a_{1n}^{(i)} \ a_{21}^{(i)} \ a_{22}^{(i)} \ \cdots \ a_{2n}^{(i)} \ \cdots \ a_{m1}^{(i)} \ a_{m2}^{(i)} \ \cdots \ a_{mn}^{(i)}]^T \quad (5-5)$$

即有

$$\mathbf{X}_i = [\mathbf{A}_i^1 \quad \mathbf{A}_i^2 \quad \cdots \quad \mathbf{A}_i^m]^T \quad (i=1,2,\cdots,M) \quad (5-6)$$

假设  $E(\mathbf{X})$  表示所有训练样本的均值向量，再令向量  $\hat{\mathbf{X}}_i = \mathbf{X}_i - E(\mathbf{X})$ ，则 PCA 方法的协方差矩阵的样本估计为：

$$\begin{aligned} C &= \frac{1}{M} \sum_{i=1}^M [\mathbf{X}_i - E(\mathbf{X})][\mathbf{X}_i - E(\mathbf{X})]^T = \frac{1}{M} \sum_{i=1}^M \hat{\mathbf{X}}_i \hat{\mathbf{X}}_i^T \\ &= \frac{1}{M} \sum_{i=1}^M \begin{bmatrix} (\hat{\mathbf{A}}_i^1)^T \\ (\hat{\mathbf{A}}_i^2)^T \\ \vdots \\ (\hat{\mathbf{A}}_i^m)^T \end{bmatrix} \begin{bmatrix} (\hat{\mathbf{A}}_i^1) & (\hat{\mathbf{A}}_i^2) & \cdots & (\hat{\mathbf{A}}_i^m) \end{bmatrix} \\ &= \frac{1}{M} \sum_{i=1}^M \begin{bmatrix} (\hat{\mathbf{A}}_i^1)^T (\hat{\mathbf{A}}_i^1) & (\hat{\mathbf{A}}_i^1)^T (\hat{\mathbf{A}}_i^2) & \cdots & (\hat{\mathbf{A}}_i^1)^T (\hat{\mathbf{A}}_i^m) \\ (\hat{\mathbf{A}}_i^2)^T (\hat{\mathbf{A}}_i^1) & (\hat{\mathbf{A}}_i^2)^T (\hat{\mathbf{A}}_i^2) & \cdots & (\hat{\mathbf{A}}_i^2)^T (\hat{\mathbf{A}}_i^m) \\ \vdots & \vdots & \ddots & \vdots \\ (\hat{\mathbf{A}}_i^m)^T (\hat{\mathbf{A}}_i^1) & (\hat{\mathbf{A}}_i^m)^T (\hat{\mathbf{A}}_i^2) & \cdots & (\hat{\mathbf{A}}_i^m)^T (\hat{\mathbf{A}}_i^m) \end{bmatrix} \\ &= \begin{bmatrix} C^{11} & C^{12} & \cdots & C^{1m} \\ C^{21} & C^{22} & \cdots & C^{2m} \\ \vdots & \vdots & \ddots & \vdots \\ C^{m1} & C^{m2} & \cdots & C^{mm} \end{bmatrix} \end{aligned} \quad (5-7)$$

其中，

$$C^{jk} = \frac{1}{M} \sum_{i=1}^M (\hat{\mathbf{A}}_i^j)^T (\hat{\mathbf{A}}_i^k) \quad (j,k=1,2,\cdots,m) \quad (5-8)$$

可以看出，PCA 的协方差矩阵  $C$  是由  $m \times n$  个方阵  $C^{jk}$  ( $j,k=1,2,\cdots,m$ ) 构成的，每个方阵的维数均为  $n \times n$ 。

我们将 PCA 方法中的方阵  $C^{jk}$  ( $j,k=1,2,\cdots,m$ ) 与 2DPCA 方法中的方阵  $G_i^{(j)}$  ( $j=1,2,\cdots,m$ )，可以发现如下关系：

$$C^{jk} = G_i^{(j)} \quad (j=1,2,\cdots,m) \quad (5-9)$$

图 5-1 形象地描述了上述关系，即反映了 2DPCA 方法的协方差矩阵与 PCA 方法的协方差矩阵之间的联系，同时也说明了 2DPCA 方法的实质是基于行分块的 PCA 方法。

最后，简单地分析一下 2DPCA 方法与 PCA 方法相比的优势。首先，2DPCA 方法相当于对图像的各行实施 PCA，所处理的数据维数远远低于 PCA 方法中图像

向量的维数，因而解决了 PCA 方法高计算复杂性的难题；其次，2DPCA 方法中将图像的每一行看作是一个样本，即是极大地扩充了样本数量，因而不存在 PCA 方法中的小样本问题，可以较为精确地估计协方差矩阵；第三，考虑图 5-1 的情况，2DPCA 方法的协方差矩阵  $G_t$ ，可以由 PCA 方法的协方差矩阵  $C$  得到，但协方差矩阵  $C$  中的绝大多数信息都被抛弃了，可为什么 2DPCA 方法的识别率仍然优于 PCA 方法呢？实际上，协方差矩阵  $C$  中被抛弃的信息，多数是相隔很远的像素点之间的相关性信息，这些信息原本就没有什么利用价值（相关性几乎为零），但在样本数量较少的情况下，这些数据反而会对计算 PCA 方法的特征向量造成干扰，导致利用了更多相关性信息的 PCA 方法反而不如 2DPCA 方法。

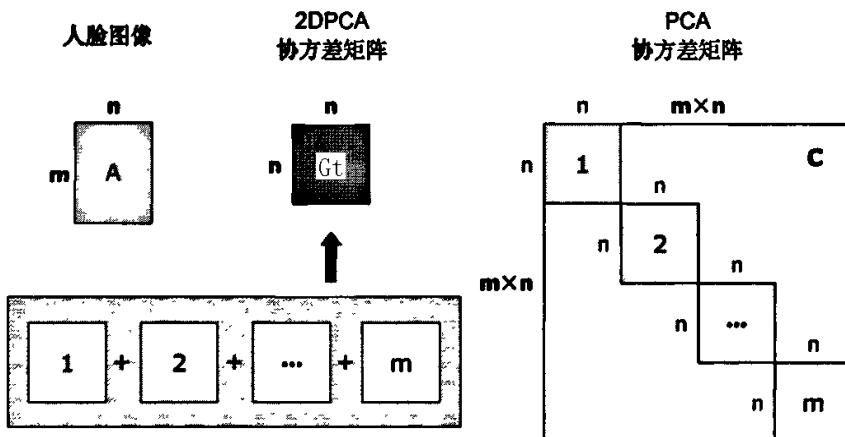


图 5-1 2DPCA 方法的实质示意图

### 5.3 基于块内相关性的 2DPCA 新方法

#### 5.3.1 基本思想

通过上一节对 2DPCA 的数学分析可知，2DPCA 方法是利用图像矩阵同一行像素之间的相关性来构造协方差矩阵的，其实质是一种特殊的分块 PCA 方法，即每个分块是图像矩阵的一行。可见，2DPCA 方法在构造协方差矩阵时，仅保留了图像中同一行像素之间的相关性信息，完全舍弃了行与行之间像素的相关性信息。再者从直观上理解，同一行内相距较远的像素之间的相关性是较弱的，如果能够抛弃这些相关性较弱的信息，用行与行之间相关性较强的信息取代之，那么就可以充分利用图像行与列像素之间的相关性信息。

此外，人脸的特征主要表现在局部区域，如人脸的主要器官眼睛、鼻子、嘴巴等，都属于人脸的局部特征，而2DPCA方法仅考虑了人脸图像同一行像素之间的相关性，割裂了人脸局部区域内的各个像素之间的相关性，很难用于表征人脸的局部特征，这是不利于进行人脸识别的因素。我们知道，对图像中的某一个像素点来说，离该像素点越近的点，与该点的相关性就越大，因此图像矩阵中同一行像素之间的相关性不如同一个局部块像素之间的相关性强。可见，如果对图像进行分块，利用块内像素的相关性来构造协方差矩阵，那么不但能够充分地利用图像行与列像素之间的相关性信息，而且构造出的协方差矩阵和计算出来的最优投影方向都包含了人脸图像的局部特征信息，因而可以获得更高的人脸识别率。

基于上述考虑，本文在2DPCA方法的基础上，提出了基于块内相关性的2DPCA人脸识别方法。这种方法先将人脸图像分割为若干个互不重叠的图像块，图像块的尺寸需要满足以下要求：每个图像块所包含的像素个数与原图像一行的像素个数相同。然后，照从左到右、从上到下的扫描方式，图像块转化为一维行向量，再将这些行向量按照图像块的先后顺序排列成新的二维矩阵。该矩阵与原图像的尺寸是一致的，但由于经过了重排，该矩阵对应的图像不再是人眼可以识别的人脸图像。最后，将经过前面的步骤处理之后的二维图像矩阵，当作是2DPCA方法的输入图像，使用2DPCA方法进行特征提取与分类识别。这种方法在算法上优点是利用了经典2DPCA方法的全部流程，只需要做简单的图像预处理（分块、重排操作）即可以使人脸识别率得到提升。

本文提出的基于块内相关性的2DPCA人脸识别方法，在保持2DPCA方法低计算复杂度的同时，最大限度地保留了图像行与列像素之间的相关性信息，克服了2DPCA方法的缺陷，而且该方法能够较好地表征人脸图像的局部特征，比2DPCA方法具有更好的人脸识别性能。

### 5.3.2 算法描述

基于块内相关性的2DPCA人脸识别方法的算法描述如下：

#### (1) 图像预处理

首先对人脸图像进行预处理，包括几何归一化、光线补偿、灰度归一化等操作，经过预处理的人脸图像称作标准化人脸图像。标准化人脸图像的尺寸固定，而且眼睛、嘴巴均处于预先指定的标准位置，这使得经过图像分块之后，相同的

人脸局部特征能落入编号相同的图像块内。此外，标准化人脸图像的光线分布被统一到了标准水平，消除了光线差异对人脸识别的影响。

## (2) 人脸图像分块与重排

对于一幅  $m \times n$  维的标准化人脸图像，若有  $m = p \times m'$  与  $n = q \times n'$ ，则可将原图像分割为  $p \times q$  个互不重叠的图像块，每个图像块的尺寸为  $m' \times n'$ 。这里的  $p, q, m', n'$  均为整数，且基于块内相关性的 2DPCA 方法要求  $m' \times n' = n$ ，因此人脸图像的分块需要图像预处理中几何归一化步骤的配合。

人脸图像分块之后，对每一个  $m' \times n'$  维的图像块，按照从左到右、从上到下的扫描方式，将二维矩阵转化为一维行向量（ $n$  维），再将这些行向量按图像块的先后顺序排列成一个新的二维图像矩阵。这个新的矩阵是由  $m$  个  $n$  维的行向量组成，尺寸与原图像一样，也是  $m \times n$  维的，但新矩阵中每一个行向量选取的是原图像中处于同一局部块内的像素。

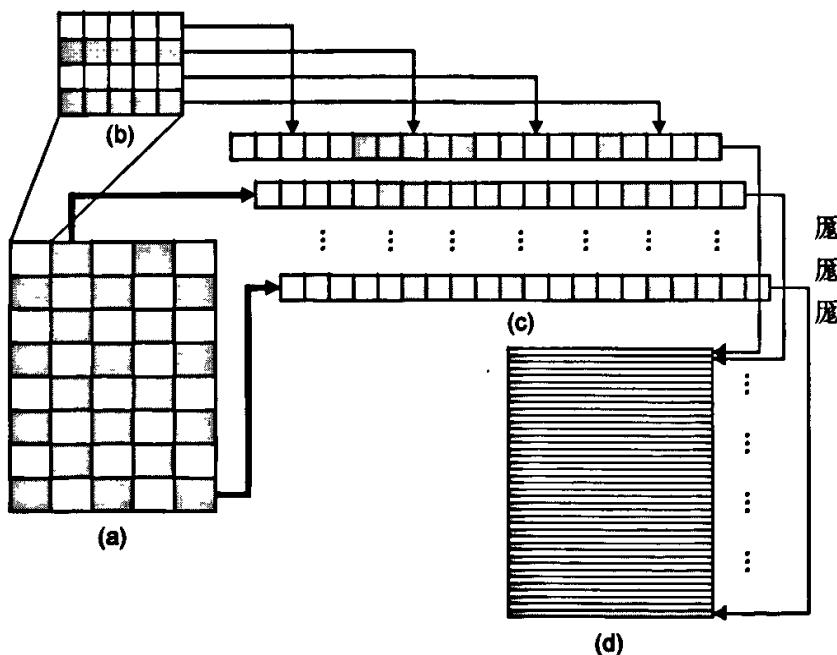


图 5-2 人脸图像分块、重排的操作示意图

图 5-2 是人脸图像分块、重排的操作示意图。其中，图(a)表示对原图像进行分块，每一个小方块代表一个图像块；图(b)表示原图像中的一个图像块，每一个小方块代表一个像素；图(c)表示将图像块按照逐行扫描的方式排列成的行向量，

每一个小方块代表一个像素；图(d)表示将所有行向量按照图像块的先后顺序排列产生的新的二维图像矩阵，每一行表示一个行向量。

### (3) 计算最优投影矩阵

假设训练集中总共有  $M$  幅人脸图像，则进行图像分块、重排的操作之后，每一幅图像  $A_i$  ( $i=1,2,\dots,M$ ) 对应于一个新的二维矩阵  $A'_i$  ( $i=1,2,\dots,M$ )，二者皆为  $m \times n$  维的矩阵。将新的二维矩阵视作 2DPCA 方法中的人脸图像矩阵（尽管这些矩阵对应的图像不再是人眼可以识别的人脸图像），使用 2DPCA 方法计算其协方差矩阵，进而求出其最优投影矩阵，即得到人脸图像的特征子空间。

### (4) 特征提取

计算出人脸图像的特征子空间后，对任意一幅人脸图像均可通过步骤(1)、(2)，将二维图像矩阵经过图像预处理和分块、重排操作，得到一个与原图像尺寸一致的新的二维矩阵，然后可以使用 2DPCA 方法中的特征投影方法，提取人脸图像的特征矩阵。

### (5) 模式分类

对待识别的人脸图像，通过步骤(1)(2)(4)计算出它的特征矩阵，然后将该特征矩阵与数据库中已知人脸图像的特征矩阵进行比对，计算出人脸图像之间的相似度，将数据库中与待识别人脸图像相似度最大的人脸图像，判定为与待识别人脸图像最相似的图像，该最相似图像对应的身份则是待识别人脸图像的人物身份。

下面详细地介绍基于块内相关性的 2DPCA 人脸识别方法的实现过程。

## 5.3.3 建库过程

### (1) 建立训练集

建立训练集即是将输入图像经过图像预处理，转化成为标准化人脸图像，并保存到数据库中的过程，图像预处理主要包括如下几个步骤：

#### ①几何归一化

几何归一化对于基于块内相关性的 2DPCA 人脸识别方法来说是非常重要的。由于基于块内相关性的 2DPCA 方法要保证每个图像分块所包含像素个数与原图像一行的像素个数相同，因此需要将图像调整为合适的尺寸。在经过几何归一化后图像中，眼睛、嘴巴均处于预先指定的标准位置，经过图像分块之后，可以保证

编号相同的图像块对应于相同的人脸局部特征。几何归一化具体实施方法如下：

我们设定标准化图像的尺寸为  $m \times n$ ，右眼中心点在  $(X_{eyeR}, Y_{eye})$ ，左眼中心点在  $(X_{eyeL}, Y_{eye})$ ，嘴巴中心点在  $(X_{mouth}, Y_{mouth})$ 。对任意一幅给定的人脸图像，假设其右眼中心点在  $(x_1, y_1)$ ，左眼中心点在  $(x_2, y_2)$ ，嘴巴中心点在  $(x_3, y_3)$ ，令

$$X = X_{eyeL} - X_{eyeR} \quad (5-10)$$

$$Y = Y_{mouth} - Y_{eye} \quad (5-11)$$

几何扩张系数分别如下：

$$C_x = X / (x_2 - x_1) \quad (5-12)$$

$$C_y = Y / (y_3 - y_1) \quad (5-13)$$

将图像的横坐标按照  $C_x$  的比例，纵坐标按照  $C_y$  的比例进行放大 ( $C > 1$ ) 或者缩小 ( $C < 1$ )。然后，将图像的原点定在  $(x_1 - 16, y_1 - 24)$ ，切割出  $96 \times 80$  的图像，并抛弃多余的部分。

### ②直方图均衡

直方图均衡可以看作是一种简单的光线补偿处理，可以提高图像的对比度，扩大灰度动态范围，使图像的灰度层次更加分明、细节更加清晰，而突显图像的局部细节则是有利于提高 2DPCA 方法的人脸识别率的。

### ③灰度归一化

灰度归一化使得所有人脸图像的灰度平均值处于同一水平，这对基于统计分析的 2DPCA 人脸识别方法也是必要的，可以避免由于图像灰度差异过大而造成的误差。灰度归一化具体实施方法如下：

选取一张灰度适中的人脸图像作为灰度归一化的标准图，计算其灰度平均值：

$$W = \bar{\sigma} = \frac{1}{N} \sum_{i=1}^N \sigma_i \quad (5-14)$$

其中， $\sigma_i$  代表人脸图像中第  $i$  个像素的灰度值， $N$  为人脸图像的像素总数。

然后计算出需要归一化的图像的灰度平均值  $\omega$ ，则灰度调整比例为：

$$C_w = \frac{W}{\omega} \quad (5-15)$$

将该图像中所有像素的灰度值按照  $C_w$  的比例进行调整，即完成了灰度归一化。

## (2) 人脸图像的分块、重排

假设训练集中共有  $M$  幅人脸图像，则任意一幅标准化人脸图像，都可以用一个  $m \times n$  的二维矩阵来表示，矩阵元素表示像素的灰度值，将图像矩阵记为  $A$ ，即：

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (5-16)$$

将矩阵  $A$  划分为  $p \times q$  个大小相同的子矩阵，即：

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix} \quad (5-17)$$

其中，每个分块  $A_{kl}$  ( $k=1,2,\dots,p$ ;  $l=1,2,\dots,q$ ) 是  $m' \times n'$  维的矩阵，即有  $p \times m' = m$ ,  $q \times n' = n$ ，并要求每个分块的尺寸为  $m' \times n' = n$ ，这需要图像预处理中几何归一化步骤的配合。

于是，每一幅图像就被分为  $p \times q = m$  个分块，按顺序记为  $D_i$  ( $i=1,2,\dots,m$ )，即：

$$D_i = \begin{bmatrix} d'_{11} & d'_{12} & \cdots & d'_{1n'} \\ d'_{21} & d'_{22} & \cdots & d'_{2n'} \\ \vdots & \vdots & \ddots & \vdots \\ d'_{m'1} & d'_{m'2} & \cdots & d'_{m'n'} \end{bmatrix} \quad (5-18)$$

对每一分块进行逐行扫描，得到每一分块对应的行向量表示  $V_i$  ( $i=1,2,\dots,m$ )：

$$V_i = [d'_{11} \ d'_{12} \ \cdots \ d'_{1n'} \ d'_{21} \ d'_{22} \ \cdots \ d'_{2n'} \ \cdots \ d'_{m'1} \ d'_{m'2} \ \cdots \ d'_{m'n'}] \quad (5-19)$$

再将这些行向量按顺序重排，构造出新的二维图像矩阵  $A'$  如下：

$$A' = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix} \quad (5-20)$$

新的二维图像矩阵  $A'$  由  $m$  个  $n$  维 ( $m' \times n' = n$ ) 的行向量组成，尺寸与原图像相同，但新矩阵中每一个行向量选取的是原图像处于同一局部块内的像素。

对训练集中的所有标准化的人脸图像，都进行上面的图像分块、重排操作，得到新的二维图像矩阵库。

### (3) 计算最优投影矩阵

进行图像分块、重排的操作之后，每一幅输入图像  $A_i (i=1,2,\dots,M)$  ( $M$  为训练集中的图像总数) 对应于一个  $m \times n$  维的新的二维图像矩阵  $A'_i (i=1,2,\dots,M)$ ，对所有新的二维图像矩阵，后使用 2DPCA 方法计算其协方差矩阵，进而求出其最优投影矩阵，即得到人脸图像的特征子空间。具体实施方法如下：

根据新的二维图像矩阵可以计算出其图像协方差矩阵，计算公式如下：

$$G_i = \frac{1}{M} \sum_{i=1}^M (A'_i - \bar{A}') (A'_i - \bar{A}')^T \quad (5-21)$$

其中， $M$  为训练集中人脸图像的总数，矩阵  $\bar{A}'$  为人脸图像的均值，即

$$\bar{A}' = \frac{1}{M} \sum_{i=1}^M A'_i \quad (5-22)$$

计算出协方差矩阵  $G_i^*$  后，可根据以下准则函数求出最优投影方向：

$$J(X) = X^T G_i X \quad (5-23)$$

其中， $X$  是归一化的列向量，称为最优投影方向。最大化上准则函数  $J(X)$  的意义是：图像矩阵在  $X$  方向上投影后得到的特征向量的类间分散程度最大。实际上，最优投影方向  $X$  即是协方差矩阵  $G_i$  的最大特征值所对应的特征向量。

在样本类别数较多的情况下，如果只取一个最优投影方向，分类鉴别能力是远远不够的，通常情况下会选择相互正交且极大化准则函数的一组投影向量  $X_1, X_2, \dots, X_d$ ，即满足如下要求：

$$\begin{cases} \{X_1, X_2, \dots, X_d\} = \arg \max J(X) \\ X_i^T X_j = 0, \quad i \neq j, \quad i, j = 1, 2, \dots, d \end{cases} \quad (5-24)$$

在实际计算中，最优投影向量组  $X_1, X_2, \dots, X_d$  恰好就是协方差矩阵  $G_i$  的最大的前  $d$  个特征值所对应的特征向量，记成矩阵的形式则是  $X = [X_1 \quad X_2 \quad \dots \quad X_d]$ ，称作最优投影矩阵。

#### (4) 计算训练集图像的特征矩阵

对任意一幅  $m \times n$  维的人脸图像  $A$ ，其对应的经过分块、重排的新的二维矩阵为  $A'$ ，将其往最优投影方向上进行投影：

$$Y_k = A' X_k \quad (k = 1, 2, \dots, d) \quad (5-25)$$

则可获得一组投影特征向量  $Y_1, Y_2, \dots, Y_d$ ，称为图像  $A$  的主元向量，一幅图像的所有主元向量构成一个  $m \times d$  的矩阵  $B = [Y_1 \quad Y_2 \quad \dots \quad Y_d]$ ，称为图像  $A$  的特征矩阵。

将训练集中所有人脸图像的特征矩阵，以及对应的个人身份信息保存起来，就构成了已知人脸图像的特征数据库。

### 5.3.4 识别过程

#### (1) 图像预处理与图像分块、重排

对每一幅待识别的测试图像，采用训练过程中步骤(1)、(2)的方法，进行图像预处理，得到尺寸为  $m \times n$  的标准化人脸图像，然后再对标准化人脸图像进行图像分块、重排操作，得到相应的  $m \times n$  维的新的二维图像矩阵。

#### (2) 计算测试图像的特征矩阵

假设矩阵  $A^*$  表示一幅  $m \times n$  维的待识别测试图像，经过图像分块、重排操作之后，得到相应的新的二维图像矩阵  $A^{**}$ ，将  $A^{**}$  往最优投影方向上投影，即：

$$Y_k^* = A^{**} X_k \quad (k=1,2,\dots,d) \quad (5-26)$$

则可获得一组投影特征向量  $Y_1^*, Y_2^*, \dots, Y_d^*$ ，这些投影特征向量构成一个  $m \times d$  的矩阵  $B^* = [Y_1^* \ Y_2^* \ \dots \ Y_d^*]$ ，称为图像  $A^*$  的特征矩阵。

#### (3) 判定测试图像所属的类别

对于任意两幅人脸图像  $A_i$  和  $A_j$ ，对应的特征矩阵分别  $B_i = [Y_1^{(i)} \ Y_2^{(i)} \ \dots \ Y_d^{(i)}]$  和  $B_j = [Y_1^{(j)} \ Y_2^{(j)} \ \dots \ Y_d^{(j)}]$ ，则特征矩阵  $B_i$  与  $B_j$  之间的距离定义为：

$$d(B_i, B_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|_2 \quad (5-27)$$

其中， $\|Y_k^{(i)} - Y_k^{(j)}\|_2$  表示两个主元向量  $Y_k^{(i)}$  与  $Y_k^{(j)}$  之间的欧氏距离。

对所有的训练样本  $A_1, A_2, \dots, A_M$  ( $M$  为训练集中的图像总数)，每一个样本都属于某个确定的类别  $\omega_k$  ( $k=1,2,\dots,C$ ) ( $C$  为训练集图像中的总人数)，它们对应的特征矩阵分为  $B_1, B_2, \dots, B_M$ 。对于某个给定的测试图像  $A^*$ ，其对应的特征矩阵为  $B^*$ ，若有  $d(B^*, B_i) = \min_j d(B^*, B_j)$ ，且有  $B_i \in \omega_k$ ，则分类的结果是  $A^* \in \omega_k$ 。

## 5.4 实验结果与分析

本章的实验是在 FERET 人脸库上实现的，选取选择库中的 53 人，每人 6 幅正面人脸图像，共 318 幅人脸图像组成训练和测试人脸库，包括了不同光线、表

情等情况下的人脸图像。原始图像的尺寸为  $112 \times 92$ ，经过预处理的标准化人脸图像的尺寸为  $96 \times 80$ ，将图像分割成 30 个互不重叠的图像块，每个图像块的尺寸为  $16 \times 16$ 。此外，标准化人脸图像的光线分布被统一到了标准水平，消除了光线差异对人脸识别的影响。图 5-3 给出了原始图像与标准化人脸图像的比较。



图 5-3 原始图像（上）与标准化人脸图像（下）的比较（FERET 人脸库）

实验比较了 2DPCA 方法与本文提出的基于块内相关性的 2DPCA 方法的人脸识别率，分为三组实验。

#### （1）第一组实验

随机地选取库中 2 幅图像 / 人用于训练，余下的 4 幅图像 / 人则用于测试，即训练人数：53 人，训练样本数： $53 \times 2 = 106$  个，测试人数：53 人，测试样本数： $53 \times 4 = 212$  个。采用随机抽取多次建库的方式重复 30 次实验，取识别率的平均值作为最终识别率，下表给出了实验的结果。

表 5-1 识别率比较（ $53 \times 2$  个训练样本， $53 \times 4$  个测试样本，FERET 人脸库）

识别率 特征 向量数目 \\	2DPCA	基于块内相关性 的 2DPCA	识别率 特征 向量数目 \\	2DPCA	基于块内相关性 的 2DPCA
1	77.48%	78.79%	9	85.02%	85.69%
2	79.58%	82.89%	10	85.27%	85.82%
3	81.87%	83.84%	11	85.27%	85.72%
4	83.51%	84.23%	12	85.33%	85.60%
5	83.82%	84.78%	13	85.19%	85.63%
6	83.69%	85.06%	14	85.38%	85.50%
7	83.95%	85.55%	15	85.42%	85.53%
8	84.48%	85.50%	16	85.39%	85.53%

下图是在第一组的实验条件下，2DPCA 方法与基于块内相关性的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线：

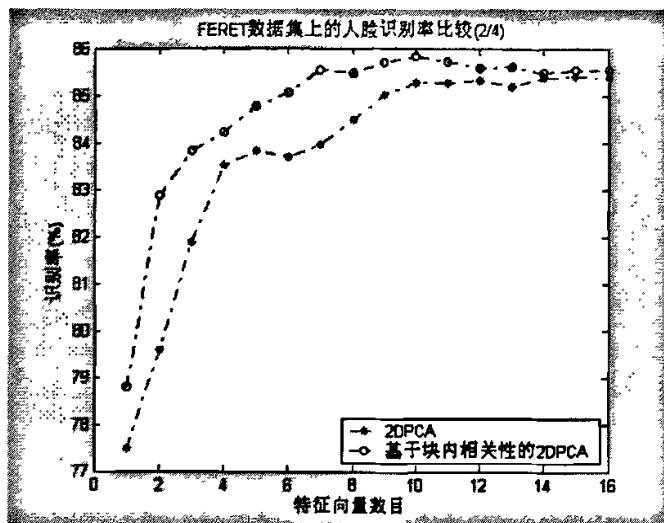


图 5-4 识别率比较 (53×2 个训练样本, 53×4 个测试样本, FERET 人脸库)

## (2) 第二组实验

随机地选取库中 3 幅图像 / 人用于训练, 余下的 3 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 3 = 159$  个, 测试人数: 53 人, 测试样本数:  $53 \times 3 = 159$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 5-2 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

识别率 特征 向量数目	2DPCA	基于块内相关性 的 2DPCA	识别率 特征 向量数目	2DPCA	基于块内相关性 的 2DPCA
1	83.21%	83.58%	9	89.14%	90.04%
2	84.36%	86.94%	10	89.33%	90.29%
3	86.39%	88.09%	11	89.29%	90.04%
4	87.48%	88.45%	12	89.33%	90.15%
5	87.95%	88.85%	13	89.35%	89.94%
6	87.88%	89.39%	14	89.54%	89.83%
7	88.24%	89.87%	15	89.50%	89.94%
8	88.64%	90.02%	16	89.58%	90.02%

下图是在第二组的实验条件下, 2DPCA 方法与基于块内相关性的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

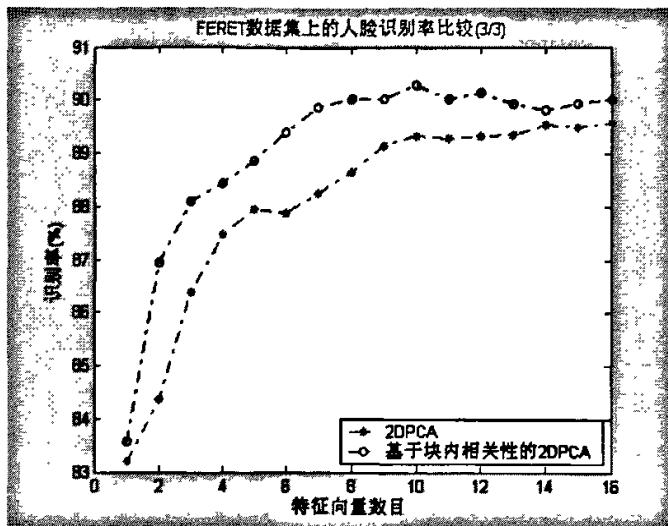


图 5-5 识别率比较 (53×3 个训练样本, 53×3 个测试样本, FERET 人脸库)

### (3) 第三组实验

随机地选取库中 4 幅图像 / 人用于训练, 余下的 2 幅图像 / 人则用于测试, 即训练人数: 53 人, 训练样本数:  $53 \times 4 = 212$  个, 测试人数: 53 人, 测试样本数:  $53 \times 2 = 106$  个。采用随机抽取多次建库的方式重复 30 次实验, 取识别率的平均值作为最终识别率, 下表给出了实验的结果。

表 5-3 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

识别率 特征 向量数目	2DPCA	基于块内相关性 的 2DPCA	识别率 特征 向量数目	2DPCA	基于块内相关性 的 2DPCA
1	87.30%	87.67%	9	92.20%	92.99%
2	88.77%	90.25%	10	92.26%	93.33%
3	89.75%	91.19%	11	92.11%	93.24%
4	90.28%	91.45%	12	92.11%	93.27%
5	90.79%	91.64%	13	92.04%	92.83%
6	90.82%	92.48%	14	92.20%	92.83%
7	91.32%	93.02%	15	92.20%	93.05%
8	91.67%	93.08%	16	92.52%	93.33%

下图是在第三组的实验条件下, 2DPCA 方法与基于块内相关性的 2DPCA 方法的人脸识别率随投影向量的数目而变化的曲线:

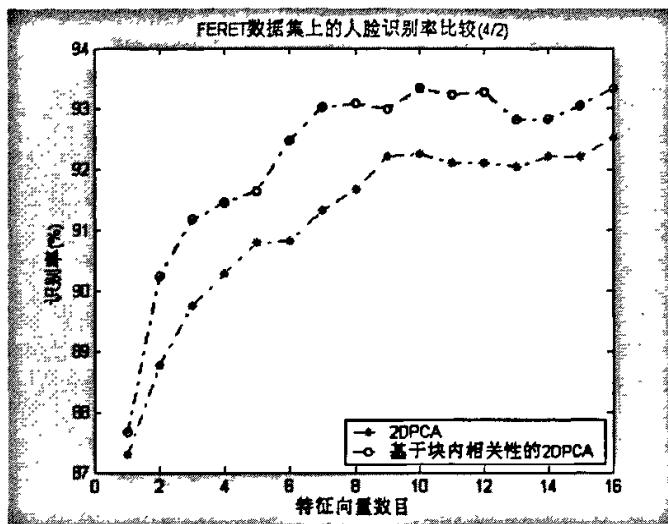


图 5-6 识别率比较 (53×4 个训练样本, 53×2 个测试样本, FERET 人脸库)

从上面三组实验的结果可以看出, 在选取相同个数的 2DPCA 特征向量的情况下, 基于块内相关性的 2DPCA 方法的比 2DPCA 方法的人脸识别率要高, 特别是当特征向量数目较少的时候, 基于块内相关性的 2DPCA 方法的优势更明显, 人脸识别率约有 2%~4% 的提升。从另一个角度看, 基于块内相关性的 2DPCA 方法与 2DPCA 方法相比, 即仅需要较少的投影向量就可以达到更高的识别率。

从实验结果的曲线还可以看出, 随着特征向量数目的增加, 基于块内相关性的 2DPCA 方法与 2DPCA 方法在识别率上的差别逐渐缩小, 这是因为对于各种基于子空间方法的人脸识别方法 (如 PCA、2DPCA) 来说, 随着主元数目的增加, 主元构成的子空间会向相邻图像空间扩张, 各种子空间方法的识别效果也将逐渐趋同。因此, 各种子空间方法的识别性能, 主要是看能否在低维子空间上获得较高的判别率, 这正是基于块内相关性的 2DPCA 的优势所在。

表 5-4 最佳识别率比较 (训练样本数 / 类  $k$ ,  $2 \leq k \leq 4$ , FERET 人脸库)

训练样本数 / 类 ( $k$ )	最佳识别率	2DPCA	基于块内相关性的 2DPCA
2		85.42%	85.82%
3		89.58%	90.29%
4		92.52%	93.33%

考虑到训练样本和测试样本的选择不同, 会使达到最佳识别率时的特征向量

数目  $d$  产生波动，因此，还需要比较 2DPCA 方法与基于结块内相关性的 2DPCA 方法的最佳识别率（如表 5-4 所示）。

图 5-7 是在  $2 \leq k \leq 4$  ( $k$  为训练样本数 / 类) 的条件下，2DPCA 方法与基于块内相关性的 2DPCA 方法的最佳识别率的直观比较，可见，基于块内相关性的 2DPCA 方法在识别率上总是 2DPCA 方法。

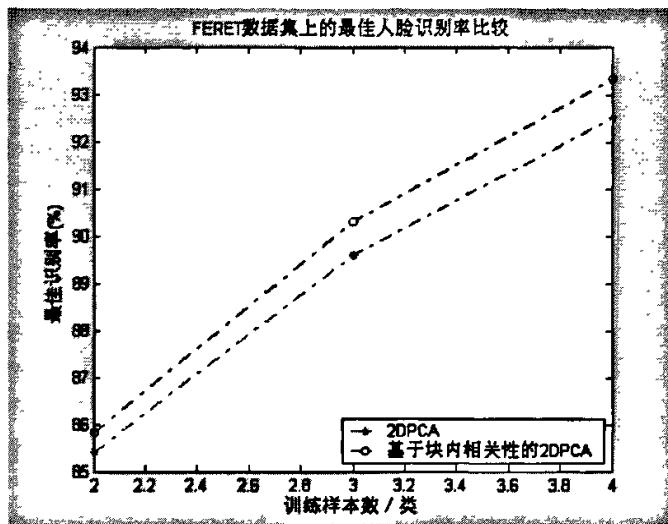


图 5-7 最佳识别率比较 (训练样本数 / 类  $k$ ,  $2 \leq k \leq 4$ , FERET 人脸库)

## 第 6 章 结束语

### 6.1 本文总结

人脸识别技术是利用计算机分析人脸图像，进而从中提取出有效的识别信息，用来识别的身份的一门技术。人脸识别涉及到图像处理、模式识别、人工智能、生理学、心理学等多个学科，是一个相当困难而又极富挑战性的前沿课题。

人脸识别的研究主要分为人脸检测、特征提取和人脸识别三个部分，本文以人脸的特征提取与识别为工作重点，对基于二维主元分析（2DPCA）的人脸识别方法进行了深入的研究，主要工作可以分为下面三个部分：

#### （1）对 PCA 与 2DPCA 的基本原理及相互关系的研究

本文从 PCA 人脸识别方法着手，叙述了主元分析的基本原理及相关的数学理论知识，介绍了经典 PCA 算法中的特征脸（Eigenface）方法的原理与实现过程。然后，在 PCA 方法的基础上，叙述了 2DPCA 方法的基本思想与实现过程，并给出了文献中 2DPCA 方法与 PCA 方法的性能比较，证实了 2DPCA 方法不但避免了庞大的计算量，而且能获得优于 PCA 方法的人脸识别率。此外，还介绍了 2DPCA 人脸识别的一些改进方法，如 2DPCA+PCA、K2DPCA、B2DPCA 等。

在第 5 章中，对 2DPCA 方法进行了数学分析，给出了 2DPCA 方法的协方差矩阵与 PCA 方法的协方差矩阵之间的联系，证明了 2DPCA 方法的实质是基于行分块的 PCA 方法，同时指出了 2DPCA 方法抛弃了行与行之间像素的相关性信息的缺陷，为后续的研究工作奠定了一定的理论基础。

#### （2）提出了基于结构分块的 2DPCA 新方法

针对不完整人脸图像的识别问题，本文在 2DPCA 人脸识别方法的基础上，考虑了人脸图像所具有局域特性（即人脸图像一个局部块往往表示一个完整的语义，如鼻子、眼睛、嘴巴等），提出了基于结构分块的 2DPCA 人脸识别方法，该方法结合了统计模式识别与结构模式识别的优点，通过调整各图像块的加权系数，可应用于各种不完整人脸图像的识别。其基本思想是：首先把人脸图像分割为若干个互不重叠的图像块，然后对每一个图像块应用 2DPCA 方法，求出其特征矩阵（本文中称之为结构主元），在计算图像相似度的时候，先计算出各图像块对应的相似

度，然后用加权和的方法计算出最终的相似度。

为了验证本文所提出方法的有效性，在 FERET 人脸库上进行了相关实验，对 2DPCA 方法和基于结构分块的 2DPCA 方法进行比较。实验结果表明，基于结构分块的 2DPCA 方法在人脸识别率上优于 2DPCA 方法，而且可以通过调整各图像块的加权系数，获取更好的识别性能。在选取较少主元数目的情况下，与 2DPCA 方法相比，基于结构分块的 2DPCA 方法的识别率大约有 3%~7% 的提升。

### (3) 提出了基于块内相关性的 2DPCA 新方法

通过对 2DPCA 的数学分析可以发现，2DPCA 方法是利用图像矩阵同一行像素之间的相关性来构造协方差矩阵的，其实质是基于行分块 PCA 方法。2DPCA 方法只利用了图像矩阵同一行像素之间的相关性信息，完全抛弃了行与行之间像素的相关性信息，然而对于人脸图像来说，图像矩阵同一行像素之间的相关性没有一个局部块像素之间的相关性强，而且人脸图像一个局部块通常具有特定的语义。基于上述考虑，本文提出了基于块内相关性的 2DPCA 人脸识别方法，该方法不但保留了 2DPCA 方法低计算复杂度的优点，而且能够充分地利用图像行与列像素之间的相关性信息，克服了 2DPCA 方法的不足。其基本思想是：首先对人脸图像进行分块，然后把图像块的像素转化为行向量，再把这些行向量按图像块的顺序重排为新的图像矩阵，最后对这些新的图像矩阵采用 2DPCA 方法进行人脸识别。

为了验证本文所提出方法的有效性，在 FERET 人脸库上进行了相关实验，对 2DPCA 方法和基于块内相关性的 2DPCA 方法进行比较。实验结果表明，基于块内相关性的 2DPCA 方法的人脸识别率优于 2DPCA 方法，在选取较少主元数目的情况下，识别率大约有 2%~4% 的提升。而且，基于块内相关性的 2DPCA 方法仅需要较少的投影向量就可以达到比 2DPCA 方法更高的识别率，从而降低了对数据存储方面的需求。

## 6.2 进一步的工作

在 2DPCA 方法中，图像特征矩阵的维数是  $m \times d$  ( $m$  为图像行数， $d \ll m$ )，与 PCA 相比，2DPCA 需要更多的系数才能表征一幅图像，从压缩存储的角度考虑，2DPCA+PCA 方法不仅保持了 2DPCA 方法的高识别率，同时也拥有 PCA 方法所需系数较少的特点，是一种较好的人脸识别方法。本文提出的两种新方法：基于

结构分块的2DPCA方法和基于块内相关性的2DPCA方法，均可利用2DPCA+PCA方法的思想，使用PCA对图像特征矩阵进行降维，以降低对数据存储的需求。

本文的第4章中，对基于结构分块的2DPCA方法进行了实验证，结果表明各图像块加权系数的设置，会对最终人脸识别率造成影响，但加权系数是通常根据经验进行设置的，并没有相关理论上的支持，如果能自适应地找出一种最优的加权系数设置方法，使得人脸识别率总是能达到最大，那么基于结构分块的2DPCA人脸识别方法将具有更好的适应性。

本文提出两种人脸识别新方法，都使用了图像分块这一步骤，但基于块内相关性的2DPCA人脸识别方法中，由于利用了2DPCA方法的全部流程，只是事先对图像进行分块、重排操作，因而并没有引入加权系数。可以考虑借鉴基于结构分块的2DPCA人脸识别方法的思想，通过设置加权系数来获得更优的性能。

上述设想，均是本文后续的工作，但由于个人能力以及时间上的关系，还有待于进一步的研究与验证。

## 参考文献

- [1] P.J. Phillips1, W.T. Scruggs, A.J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, M. Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Results," March 2007, <http://www.frvt.org/FRVT2006/Results.aspx>
- [2] P. Juell, R. Marsh, "A Hierarchical Neural Network for Human Face Detection," *Pattern Recognition*, Vol. 29(5), pp. 781~787, 1996.
- [3] A.J. Schofield, P.A. Mehta, T.J. Stonham, "A System for Counting People in Video Images Using Neural Networks to Identify the Background Scene," *Pattern Recognition*, Vol. 29(8), pp. 1425~1428, 1996.
- [4] K.S. Yoon, Y.K. Ham, R.H. Park, "Hybrid Approaches to Frontal View Face Recognition Using the Hidden Markov Model and Neural Network," *Pattern Recognition*, Vol. 31(3), pp. 283~293, 1998.
- [5] H.A. Roudey, "Neural Network-Based Face Detection," in *Proceedings of Image Understanding Workshop*, pp. 725~735, 1996.
- [6] S. Ranganath, K. Arun, "Face Recognition Using Transform Features and Neural Networks," *Pattern Recognition*, Vol. 30(10), pp. 1615~1622, 1997.
- [7] H.A. Rowley, S. Baluja, T. Kanade, "Neural Network-Based Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20(1), pp. 23~28, 1998.
- [8] M. Turk, A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3(1), pp. 71~86, 1991.
- [9] M. Turk, A. Pentland, "Face Recognition using Eigenfaces," in *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586~591, 1991.
- [10] A. Pentland, B. Moghaddam, T. Starner, "View-based Modular Eigenspaces for Face Recognition," in *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 84~91, 1994.
- [11] G.Z Yang, T.S. Huang, "Human Face Detection in a Complex Background," *Pattern Recognition*, Vol. 27(1), pp. 53~63, 1994.
- [12] K.C Yow, R. Cipolla, "Feature-Based Human Face Detection," *Image and Vision Computing*, Vol. 15(9), 713~735, 1997.

- [13] A.J. Baddeley, H. Heijmans, “Incidence and Lattice Calculus with Applications to Stochastic Geometry and Image Analysis,” *Applicable Algebra in Engineering, Communication and Computing*, Vol. 3(3), pp. 129~146, 2005.
- [14] A. Tanku, Y. Yeshurun, N. Intrator, “Face Detection by Direct Convexity Estimation,” *Pattern Recognition Letters*, Vol. 18(9), pp. 913~922, 1997.
- [15] C.H. Lee, J.S. Kim, K.H. Park, “Automatic Human Face Location in a Complex Background Using Motion and Color Information,” *Pattern Recognition*, Vol. 29(11), pp. 1877~1899, 1996.
- [16] Y. Dai, Y. Nakano, “Face Texture Model Based on SGLD and Its Application in Face Detection in a Color Scene,” *Pattern Recognition*, Vol. 29(16), pp. 1007~1017, 1996.
- [17] M.H. Yang, D.J. Kriegman, N. Ahuja, “Detecting Faces in Images: A Survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24(1), pp. 34~58, 2002.
- [18] R. Brunelli, T. Poggio, “Face Recognition: Features versus Templates,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15(10), pp. 1042~1052, 1993.
- [19] A. Samal, P.A. Iyengar, “Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey,” *Pattern Recognition*, Vol. 25(1), pp. 65~77, 1992.
- [20] A.V. Nefian, M.H. Hayes, “Hidden Markov Models for Face Recognition,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 2721~2724, 1998.
- [21] A.P. Dempster, N.M. Laird, D.B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society, Series B: Methodological*, Vol. 39(1), pp. 1~38, 1977.
- [22] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, “Active Shape Models: Their Training and Application,” *Computer Vision and Image Understanding*, Vol. 61(1), pp. 38~59, 1995.
- [23] T.F. Cootes, G.J. Edwards, C.J. Taylor, “Active Appearance Models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23(6), pp. 681~685, 2001.
- [24] I. Matthews, S. Baker, “Active Appearance Models Revisited,” *International*

- Journal of Computer Vision*, Vol. 60(2), pp. 135~164, 2004.
- [25] M. Kirby, L. Sirovich, “Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12(1), pp. 103~108, 1990.
- [26] P.J. Philips, H. Moon, S. Rizvi, P. Rassu, “The FERET Evaluation Methodology for Face Recognition Algorithm,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19(7), pp. 673~686, 1997.
- [27] B. Moghaddam, A. Pentland, “Probabilistic Visual Learning for Object Representation,” *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 19(7), pp. 696~710, 1997.
- [28] C.J. Liu, H. Wechsler, “An Unified Bayesian Framework for Face Recognition,” in *Proceedings of International Conference on Image Processing*, Vol. 1, pp.151~155, 1998.
- [29] C.J. Liu, H. Wechsler, “Robust Coding Schemes for Indexing and Retrieval from Large Face Database,” *IEEE Transactions on Image Processing*, Vol. 19(1), 132~137, 2000.
- [30] A. Yilmaz, M. Gokmen, “Eigenhill vs Eigenface and Eigenedge,” in *Proceedings of International Conference Pattern Recognition*, Vol. 2, pp. 827~830, 2000.
- [31] J. Yang, D. Zhang, A.F. Frangi, J.Y. Yang, “Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26(1), pp. 131~137, 2004.
- [32] J. Yang, J.Y. Yang, “From Image Vector to Matrix: A Straightforward Image Projection Technique – IMPCA vs. PCA,” *Pattern Recognition*, Vol. 35(9), pp. 1997~1999, 2002.
- [33] K. Liu, Y.Q. Cheng, J.Y. Yang, “Algebraic Feature Extraction for Image Recognition Based on an Optimal Discriminant Criterion,” *Pattern Recognition*, Vol. 26(6), pp. 903~911, 1993.
- [34] D.L. Swets, J. Weng, “Using Discriminant Eigenfeatures for Image Retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18(8), pp. 831~836, 1996.
- [35] P.N. Belhumeur, J.P. Hespanha, “Eigenface vs. Fisherface: Recognition Using Class Specific Linear Projection,” *IEEE Transactions on Pattern Analysis and*

- Machine Intelligence*, Vol. 19(7), pp. 711~720, 1997.
- [36] C.J. Liu, H. Wechsler, “Enhanced Fisher Linear Discriminant Models for face Recognition,” in *Proceedings of International Conference Pattern Recognition*, Vol. 2, pp. 1368~1372, 1998.
- [37] R. Huang, Q. Liu, H. Lu, S. Ma, “Solving the Small Sample Size Problem of LDA,” in *Proceedings of 16th International Conference on Pattern Recognition*, Vol. 3, pp. 29~32, 2002.
- [38] H. Yu and J. Yang, “A Direct LDA Algorithm for High-Dimensional Data with Application to Face Recognition,” *Pattern Recognition*, Vol. 34, pp. 2067~2070, 2001.
- [39] H. Cevikalp, M. Neamtu, M. Wikes, A. Barkana, “Discriminative Common Vectors for Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27(1), pp. 4~13, 2005.
- [40] M. Li, B.Z. Yuan, “2D-LDA: A Statistical Linear Discriminant Analysis for Image Matrix,” *Pattern Recognition Letters*, Vol. 26(5), pp. 527~532, 2005.
- [41] H. Xiong, M.N.S. Swamy, M.O. Ahmad, “Two-dimensional FLD for Face Recognition,” *Pattern Recognition*, Vol. 38(7), pp. 1121~1124, 2005.
- [42] P. Comon, “Independent Component Analysis: A New Concept,” *Signal Processing*, Vol. 36(3), pp. 287~314, 1994.
- [43] A. Hyvarinen, “Survey on Independent Component Analysis,” *Neural Computing Surveys*, Lawrence Erlbaum Associates Inc., Vol. 2, pp. 94~128, 1999.
- [44] Y. Li, D. Powers, J. Peach, “Comparison of Blind Source Separation Algorithms,” *Advances in Neural Networks and Applications*, World Scientific Engineering Society, pp. 18~23, 2000.
- [45] A. Hyvarinen, “Fast and Robust Fixed-Point Algorithm for Independent Component Analysis,” *IEEE Transactions on Neural Networks*, Vol. 10(3), pp. 626~634, 1999.
- [46] M.S. Bartlett, H.M. Lades, T.J. Sejnowski, “Independent Component Representation for Face Recognition,” in *Proceedings of the SPIE*, Vol. 2339(3), pp. 528~539, 1998.
- [47] O. Deniz, M.C. Santana, M. Hernandez, “Face Recognition Using Independent Component Analysis and Support Vector Machines,” *Pattern Recognition Letters*, Vol. 24(13), pp. 2153~2157, 2003.

- 
- [48] D. Valentin, H. Abdi, A.J. O'Toole, "Categorization and Identification of Human Face Images by Neural Networks: A Review of the Linear Autoassociative and Principal Component Approaches," *Journal of Biological Systems*, Vol. 2(3), pp. 413~429, 1994.
  - [49] T. Kohonen, *Self-Organization and Associative Memory*, 2nd Edition, Berlin, German: Springer-Verlag, 1988.
  - [50] M. Lades, J.C. Vobruggen, J. Buhmann, etc., "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Transactions on Computers*, Vol. 42(3), pp. 300~311, 1992.
  - [51] S. Lawrence, C.L. Giles, A.C. Tsoi, etc. "Face Recognition: A Convolution Neural Network Approach," *IEEE Transactions on Neural Networks*, Vol. 8(1), pp. 98~113, 1997.
  - [52] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys (CSUR)*, Vol. 35(4), pp. 399~458, 2003.
  - [53] T. Fromherz, P. Stucki, M. Bichsel, "A Survey of Face Recognition," *MML Technical Report No. 97.01, Department of Computer Science, University of Zurich, Zurich*, 1997.
  - [54] W.A. Barrett, "A Survey of Face Recognition Algorithms and Testing Results," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, Vol. 1, pp. 301~305, 1997.
  - [55] W.Y. Zhao, R. Chellappa, "Image-based Face Recognition: Issues and Methods," In B. Javid and Mercel Dekker, editors, *Image Recognition and Classification*, pp. 375~402, 2002.
  - [56] R. Chellappa, C.L. Wilson, S. Sirohey, "Human and Machine Recognition of Faces: A Survey," in *Proceedings of the IEEE*, Vol. 83(5), pp. 705~741, 1995.
  - [57] X.Y. Tan, S.C. Chen, Z.H. Zhou, F.Y. Zhan, "Face Recognition from A Single Image per Person: A Survey," *Pattern Recognition*, Vol. 39(9), pp. 1725~1745, 2006.
  - [58] J. Ruiz-del-Solar, P. Navarrete, "Eigenspace-based Face Recognition: A Comparative Study of Different Approaches," *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, Vol. 35(3), pp.315~325, 2005.
  - [59] 张翠平、苏光大, 人脸识别技术综述, *中国图象图形学报*, Vol. 5A(11), pp.

- 885~894, 2000.
- [60] 李武军、王崇骏、张炜、陈世福, 人脸识别研究综述, *模式识别与人工智能*, Vol. 19(1), pp. 58~66, 2006.
- [61] 刘青山、卢汉清、马颂德, 综述人脸识别中的子空间方法, *自动化学报*, Vol. 29(6), pp. 900~911, 2003.
- [62] J.M. Lattin, J. Douglas Carroll, Paul E Green, *Analyzing Multivariate Data*, 北京: 机械出版社, 2003.
- [63] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd Edition, New York: Academic Press, 1990.
- [64] S.S. Wilks, *Mathematical Statistics*, New York: Wiley, 1963.
- [65] G. Shakhnarovich, B. Moghaddam, "Handbook of Face Recognition," *MITSUBISHI ELECTRIC RESEARCH LABORATORIES* (<http://www.merl.com>), TR2004-041, May 2004.
- [66] H. Moon, P.J. Phillips, "Computational and Performance Aspects of PCA-Based Face Recognition Algorithms," *Perception*, Vol. 30, pp. 303~321, 2001.
- [67] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 2nd Edition, The Johns Hopkins University Press, Baltimore, 1989.
- [68] 罗小桂, 何雁, 矩阵奇异值分解在计算技术中的应用, *计算机与现代化*, Vol. 6, pp. 67~68, 2006.
- [69] G.W. Stewart, *Introduction to Matrix Computations*, New York: Academic Press, 1973.
- [70] 北京大学数学系, 高等代数(第二版), 北京: 高等教育出版社, 1988.
- [71] L.W. Wang, X. Wang, X.R. Zhang, J.F. Feng, "The Equivalence of Two-Dimensional PCA to Line-Based PCA," *Pattern Recognition Letters*, Vol. 26(1), pp. 57~60, 2005.
- [72] H. Kong, L. Wang, "Generalized 2D Principal Component Analysis for Face Image Reorientation and Recognition," *Neural Networks*, Vol. 18(5-6), pp. 585~594, 2005.
- [73] L.W. Wang, X. Wang, J.F. Feng, "On Image Matrix Based Feature Extraction Algorithms," *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, Vol. 36(1), pp. 194~197, 2006.
- [74] W.M. Zuo, K.Q. Wang, D. Zhang, "Assembled Matrix Distance Metric for

- 2DPCA-Based Face and Palmprint Recognition,” in Proceedings of 2005 International Conference on Machine Learning and Cybernetics, Vol. 8, pp. 4870~4875, 2005.**
- [75] L.J. Rodríguez-Aragon, C. Conde, A. Serrano, E. Cabello, “Face Verification Advances Using Spatial Dimension Reduction Methods: 2DPCA & SVM,” *Image Analysis and Processing ? ICIAP 2005*, pp.978~985, 2005.
- [76] 陈伏兵、陈秀宏、高秀梅、杨静宇, 二维主成分分析方法的推广及其在人脸识别中的应用, *计算机应用*, Vol. 25(8), pp. 1767~1770, 2005.
- [77] 高全学、张洪才、潘泉、程咏梅, 基于肤色、知识和 2DPCA 的人脸检测, *计算机工程与应用*, Vol. 34, pp. 54~57, 2004.

## 致谢

首先要感谢我的导师马争鸣教授在我攻读硕士期间，对我一丝不苟的指导以及无微不至的关怀。马老师严谨的治学作风、渊博的学术背景、丰富的实践经验、热诚的处事态度，都给我留下了深刻的印象，也深深地影响着我的学习与工作。本论文是在马老师的悉心指导下完成的，在论文完成之际，我谨向马老师表示衷心的感谢和深深的敬意！

感谢李莹师姐在理论、学业上对我的指导，从她身上我学到了很多科研和理论方面的知识。与她的多次讨论交流，使我对人脸识别理论有了更加清晰的认识。

感谢师姐周映红博士，与她的讨论与畅谈，使我获益良多。此外，她在我的工作与生活方面，也给予了诸多宝贵的帮助和建议，使我受益匪浅。

感谢实验室的各位同窗好友，大家共同创造了一个良好的学习氛围，在一次次的学术讨论中点燃了思想的火花，与他们在一起的时光，总是充满快乐。特别感谢同一课题组的邓娜同学，与她的讨论帮助我拓展了思路。

最后，特别感谢我的父母长久以来给予我最大的支持与关怀，让我在求学过程中能专心于学业上而无后顾之忧，对他们的感激之情无法用言语来表达，他们的关怀鼓励是我前进的最大动力。