

学校 曲阜师范大学

参赛类别 自然科学类

第十届挑战杯大学生课外学术科技作品

基于支持向量机的人脸识别算法研究及其仿真实现  
Face Recognition Algorithm Research and Simulation Based on  
Support Vector Machine

院    系：物理工程学院

指导老师：王玉德（副教授）

队    员：张学志 封玲娟

专    业：物理学



# 目录

参赛作品信息.....	III
摘要 .....	1
ABSTRACT.....	2
第一章 绪论.....	3
1.1 人脸识别综述.....	3
1.2 人脸识别的研究历史及现状.....	3
1.3 人脸识别的优势及应用 .....	3
1.4 本文的主要内容.....	4
第二章 人脸图像预处理.....	6
2.1 滤波.....	6
2.2 图像的几何变换.....	6
2.3 直方图均衡化.....	6
2.4 离散余弦变换 (DCT) .....	7
第三章 特征提取.....	9
3.1 主成分分析法的原理.....	9
3.2 传统的 PCA 方法 (Eigenfaces) .....	9
3.3 基于类内平均脸的 PCA.....	12
第四章 支持向量机.....	16
4.1 支持向量机的基本思想.....	16
4.2 主要的核函数.....	17
4.3 最优分类面.....	17
第五章 人脸识别算法实验研究.....	19
5.1 实验数据.....	19
5.1.1 ORL 人脸库的简介.....	19
5.1.2 人脸库图像的重新编号.....	19
5.2 程序运行流程图展示.....	20
5.3 人脸识别的算法的实现.....	20
5.3.1 导入 ORL 人脸库图像.....	20
5.3.2 缩小图像的尺寸.....	21
5.3.3 离散余弦变换 (DCT) 处理.....	22
5.3.4 计算类内均值脸.....	22
5.3.5 用 PCA 和 SVD 提取特征向量 .....	23
5.3.6 计算图像在新的特征空间中的坐标.....	24
5.3.7 SVM 训练.....	25
5.3.8 SVM 分类预测 .....	26
5.3.9 GUI 界面设计.....	26
5.4 影响人脸识别准确率的相关因素探究.....	31
5.4.1 图像尺寸的影响.....	31
5.4.2 离散余弦变换变换的影响.....	33
5.4.3 类内均值脸的影响.....	35
5.4.4 特征向量维数的影响.....	36

5.4.5 支持向量机 (SVM) 的影响 .....	38
5.5 影响人脸识别速度的相关因素的探究 .....	41
5.5.1 “一次性导入 ORL 人脸库” 的影响 .....	41
5.5.2 图像尺寸的影响 .....	42
5.5.3 缩小尺寸和 DCT 顺序的影响 .....	43
5.5.4 DCT、PCA、SVD 和 SVM 的影响 .....	45
5.6 结果与分析 .....	45
第六章 总结与展望 .....	47
参考文献 .....	48
致谢 .....	49

## 参赛作品信息

### 一、项目的立项依据及当前国内外同类课题研究水平概述

在这日新月异的信息时代里，社会不断进步，各方面对快速有效地自动身份验证的要求日益迫切。传统的个人身份鉴别其安全性能低，伪造性高，越来越不适合当今高速发展的社会。人们在不断地研究中发现，生物特征识别技术是一种方便可靠的身份鉴别方法。

生物特征识别技术是 21 世纪国际上十大最具有前景的高新技术之一，是利用人体特有的生理或行为特征来进行身份识别，并综合了模式识别、神经网络、图像处理、数字信号处理等几大学科。从理论上说，生物特征之所以能够作为个人身份鉴别与识别的有效手段，这是由其自身所具有的四个特点所决定的：普遍性、唯一性、稳定性和不可复制性。

在生物验证方面，指纹识别和声音识别等常见的生物特征识别技术已较为成熟，但这些识别技术仍存在一些不足的地方，相比之下，人脸识别则具备方便、自然，使用者无任何心理障碍，并且通过人脸的表情与姿态分析，还能获得其它识别系统难以得到的一些信息等优势。

目前，国外对人脸识别问题研究已有很多，如美国的 MIT，英国的 Department of Engineering in University of Cambridge 等研究机构，主要集中研究模板匹配、神经网络和基于隐马尔可夫模型的方法等方面。

因此，人脸识别将在国家安全、身份认证、机器智能、人机交互、智能信息处理、以及工作和生活方式等方面有广泛的应用前景。

## 二、项目的研究目标、研究内容以及拟解决的关键问题

### (1) 总体目标

- 1.探究影响人脸识别速率的相关因素。
- 2.探究影响人脸识别准确率的相关因素。
- 3.寻找最佳算法和参数，实现快速、高准确率的人脸识别系统的构建
- 4.建立 GUI 界面，实现程序的集成，方便用户的使用。

### (2) 研究内容

- 1.在人脸图像预处理过程中，采用尺寸变换和离散余弦变换(DCT)等方法，对 ORL 人脸库中的图像进行预处理。
- 2.在人脸图像特征提取过程中，运用基于类内平均脸的主成分分析法(PCA)进行人脸图像的特征提取和降维处理，采用奇异值分解(SVD)法求解特征向量。
- 3.在人脸图像鉴别过程中，进行支持向量机(SVM)训练和分类预测。
- 4.建立 GUI 界面，实现程序的集成。

### (3) 拟解决的关键问题

- 1.对人脸图像本身所带有的光照不均匀，表情复杂和尺寸不一等问题进行处理。
- 2.通过试验寻找合适的图像尺寸和离散余弦变换（DCT）的截止阈值。
- 3.采用类内平均脸方法来缩小类内样本识别距离，扩大类间样本识别距离。

- 4.如何选择特征向量来构建特征空间。
- 5.选择核函数和寻找 SVM 分类器的最佳参数( $-c$ 和 $-g$ )。

### 三、拟采取的研究方法、技术路线和实验方案及可行性分析

#### (1) 研究方法

- 1.缩小图像尺寸，减少程序运行时间，提高识别准确率。
- 2.进行离散余弦变换 (DCT)，降低光照等噪声对识别结果的影响。
- 3.运用基于类内平均脸的主成分分析法 (PCA)，提取特征向量。
- 4.运用奇异值分解(SVD)法，求解特征向量。
- 5.建立支持向量机(SVM)模型，对人脸图像进行分类。
- 6.建立 GUI 界面，实现程序运行，显示识别结果与进度。

#### (2) 技术路线和实验方案

研究方案的流程如图 I 所示。

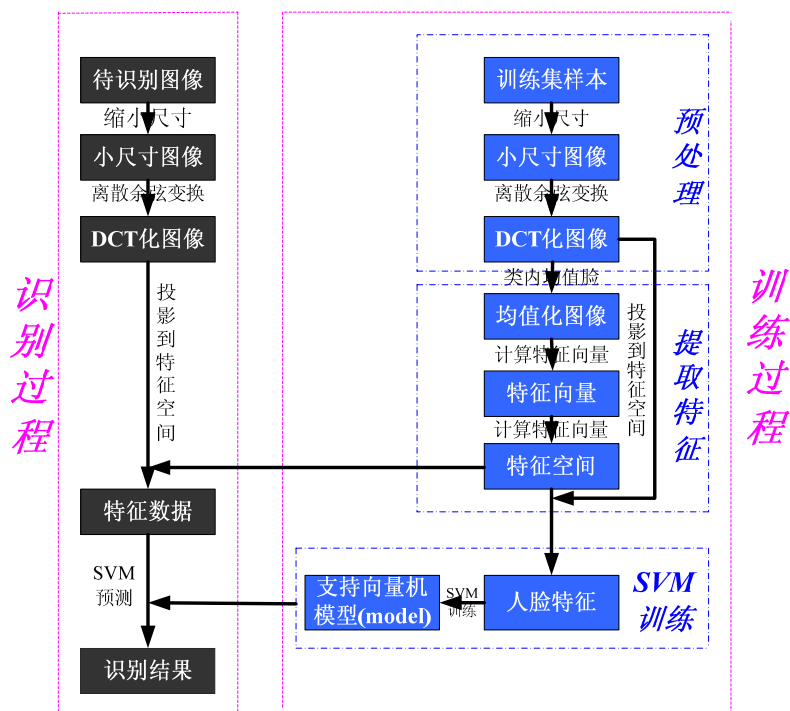


图 I 流程图  
FigI Flowchart

### (3) 可行性分析

- 1.缩小图像尺寸,可以减少计算量,提高数据运算速度。
- 2.离散余弦变换(DCT)技术是一种在数据压缩中常用的变换编码方法,可去除图像中引入的光照噪声信息,改善图像质量。
- 3.类内平均脸是每一个人的所有训练图像的平均值,它可以缩小类内图像之间的识别距离,扩大类间图像之间的识别距离。
- 4.主成分分析法(PCA)是一种正交变换。它变换后产生的新的分量正交,减少数据冗余,使得数据在一个低维的特征空间被线性表出,同时保持原始数据的绝大部分的信息。
- 5.支持向量机方法能够较好地解决小样本,非线性,高维数等问题,克服了传统的统计模式识别方法的一些缺点,可以有效地进行图像识别。
- 6.我们课题组组成人员合理,指导老师王玉德老师在支持向量机、图像处理等方面有较深入地研究。

### 四、本项目的特色和创新之处

(1)采用目前广泛使用的英国 ORL 人脸库,将其进行一次性全部导入,然后建立数据库存储人脸图像,可以避免将人脸图像分为训练集和测试集两部分导入,从而节约了时间。

(2)对于人脸图像预处理步骤进行分析,选择了尺寸归一化将所有的图像都缩小为 $8 \times 7$ 矩阵,实验表明在选择这样的尺寸时,人脸识别的准确率最高。

(3)采用离散余弦变换(DCT)方法进行压缩图像和降噪实验,在



参数和处理都相同的条件下,与运用小波变换方法相比,所运行的时间明显缩短。

(4)运用传统的混合平均脸方法来均值化原图像的,会出现类内样本的差距扩大和类间样本的差距缩小的问题,因此,本文使用基于类内平均脸的 PCA,从而很好地解决此问题。

(5)寻找到最佳参数:特征向量的维数是 48, SVM 核函数参数  $-g$  和最优惩罚参数  $-c$  分别是 0.4 和 3。

(6)设计 GUI 界面,显示人脸识别的各个过程。

(7)本文实验结果表明对 ORL 人脸库进行仿真实验,识别率达 99.5%,高于传统算法,其运行时间是 3.5 秒左右,从而真正实现人脸快速、高效的识别。

# 基于支持向量机的人脸识别算法研究及其仿真实现

张学志 封玲娟

(曲阜师范大学物理工程学院, 山东曲阜, 273165)

**摘要:** 本文基于支持向量机对人脸识别的算法进行研究, 并通过仿真实验验证算法的有效性。通过缩小图像的尺寸和调整离散余弦变换(DCT)与尺寸变换等方法节省了大量的时间, 实验表明整个程序运行时间是 3.5 秒左右; 通过 DCT、基于类内平均脸的主成分分析(PCA)和支持向量机(SVM)等方法结合使用使其识别准确率明显提高, 实验表明准确率达到 99.5%。同时, 本文还重点对影响人脸识别系统速度和识别准确率的相关因素进行了研究, 为研究人脸识别技术的应用提供理论依据。

**关键词:** 人脸识别; 尺寸缩小, 离散余弦变换; 类内平均脸; 主成分分析; 支持向量机

# Face Recognition Algorithm Research and Simulation Based on Support Vector Machine

Xuezhi Zhang, Lingjuan Feng

(College of Physics Engineering, Qufu Normal University, Shandong Qufu 273165)

**Abstract:** This paper studies face recognition algorithm based on support vector machine, and validates the validity of the algorithm by the simulation experiment. By reducing the image size and adjusting discrete cosine transform (DCT) and dimension transformation method can save a lot of time. And the experiment shows that the whole program's running time is 3.5 seconds; by DCT, the principal component analysis (PCA) of class-based average face, and support vector machine (SVM) method, it makes its identification accuracy improve obviously. The accuracy rate reaches 99.5%. Meanwhile, this paper mainly studies the impact of face recognition system on speed and the accuracy, which provides the theory basis for face recognition technology.

**Key words:** Face Recognition; Reduce the Size; Discrete Cosin Transform;

Within-Class Average Face; Principal Component Analysis; Support Vector Machine

# 第一章 绪论

在这日新月异的信息时代里，社会不断进步，各方面对快速有效地自动身份验证的要求日益迫切。传统的个人身份鉴别依靠 ID 卡、磁卡和标识号码等方式，其安全性能低，伪造性高，越来越不适合如今高速发展的社会。于是，人们希望有一种更加方便可靠的方法来进行身份鉴别，生物特征识别技术给这一愿望带来了实现的可能。

生物特征识别技术是 21 世纪国际上十大最具有前景的高新技术之一，是利用人体特有的生理或行为特征来进行身份识别。在生物特征识别技术中，人脸识别具有广泛的应用前景。

## 1.1 人脸识别综述

人脸识别(Face Recognition)是利用计算机分析人脸图像，从中提取有效地识别信息，进而来“辨识”身份的一种生物特征识别技术。

人脸识别的研究范围广义上分为以下几方面<sup>[1]</sup>：人脸检测、人脸表征、人脸鉴别、表情/姿态分析、生理分类等。

## 1.2 人脸识别的研究历史及现状

人脸识别的研究已有很长的历史，大致可分为三个阶段：第一阶段是研究人脸识别所需要的面部特征；第二阶段是人机交互式识别阶段；第三阶段是真正的机器自动识别阶段<sup>[1]</sup>。

目前，人脸识别领域是国内外研究的热点，其研究方法主要集中在模板匹配、示例学习、神经网络和隐马尔可夫模型等。

## 1.3 人脸识别的优势及应用

人脸识别作为一种新兴的生物特征识别技术 (Biometrics)，与虹

膜识别、指纹识别、掌纹扫描等技术相比，人脸识别技术在应用方面具有独特优势<sup>[2]</sup>：

(1)直观自然性：指人脸识别方式同人类进行个体识别时所利用的生物特征相同，人脸识别主要是通过观察比较人脸来区分和确认身份。

(2)非接触性：人脸识别是一种非接触的生物特征识别技术，用户友好性强，系统的使用寿命长。

(3)不被察觉性：不被察觉性可使识别者不会产生反感，不易引起人的注意而不容易被欺骗。

(4)事后性好：由于人脸识别采用人脸图像进行识别，在识别的同时，可以存储人脸图像，便于事后追查取证。

因此，人脸识别在不同的领域得到广泛应用，如身份确认、身份鉴别、访问控制(门禁系统)、安全监控(银行、海关监控)、人机交互(虚拟现实、游戏)等。

国际生物特征组织(IBG)对生物特征识别技术做了较详细的市场分析和预测，其结果显示全球生物特征识别技术 2009 年的产值达到 34 亿美元，预计 2014 年将超过 93 亿美元，市场潜力非常巨大<sup>[3]</sup>。

### 1.4 本文的主要内容

论文的内容安排如下：

第一章 绪论，对人脸识别概念，人脸识别研究历史、现状，以及其优势和广泛应用进行介绍。

第二章 人脸图像预处理，介绍了几何变换、直方图均衡化和离

散余弦变换 (DCT) 等相关内容。

第三章 人脸图像特征提取, 介绍了主成分分析法(PCA), 基于类内平均脸的 PCA 和奇异值定理(SVD)的相关内容。

第四章 支持向量机, 介绍了支持向量机的基本思想, 主要核函数和最优分类面的相关内容。

第五章 人脸识别算法实验研究, 介绍了 ORL 人脸库, 人脸识别算法及实现, 并对影响人脸识别的准确率和速率的相关因素进行了探究, 最后, 以 GUI 界面展示人脸识别程序运行过程。

第六章 结论与展望, 对全文进行总结并提出进一步需要改善的地方。

## 第二章 人脸图像预处理

预处理是模式识别过程中的一个重要的步骤。人脸图像在采集过程中，大多数的情况下会受到客观因素(如光照、设备性能)和主观因素(如表情、姿势)的影响，从而产生噪声，灰度分布不均和人脸位置和大小不确定等问题。因此，为了能稳定地进行特征提取等处理，有必要对人脸图像进行预处理。

### 2.1 滤波

图像滤波具有良好的降噪平滑作用，主要分为线性滤波和非线性滤波。线性滤波对高斯噪声有较好的平滑作用；非线性滤波能抑制非高斯噪声，保护图像边缘和细节。常用方法：均值滤波法、中值滤波法、低通滤波法等<sup>[4]</sup>。

### 2.2 图像的几何变换

对人脸图像进行几何变换，可使不同尺寸的人脸图像都统一到同样的大小，并且人脸关键部位也尽量保持一致。几何变换主要包括：缩放、平移、旋转、翻转。本文将重点介绍缩放方法。

缩放：指将给定的图像在  $x$  方向和  $y$  方向按相同的比例  $a$  缩放，从而获得一幅新的图像<sup>[4]</sup>。

由于 ORL 人脸库中的原始图像的大小为  $112 \times 92$ ，这幅图像所对应的离散方差矩阵大小为  $10304 \times 10304$ ，计算时非常费时。所以很自然想到通过缩小图像的尺寸的方法来减少计算量。

### 2.3 直方图均衡化

灰度级的直方图就是反映一幅图像中的灰度级与出现这种灰度

的概率之间的关系的图形。直方图均衡化的基本思想是把原始图像的直方图变化为均匀分布的形式,增加像素灰度值的动态范围从而达到增强图像整体对比度的效果<sup>[4]</sup>。

## 2.4 离散余弦变换 (DCT)

离散余弦变换 (DCT) 是正交变换的一种,其变换前后的信号熵和能量不变,以去除图像中的冗余信息,具有较好的压缩和降噪的作用。

二维离散余弦正变换公式<sup>[4]</sup>:

$$F(u,v) = \frac{2}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (2.2.1)$$

二维离散余弦反变换公式:

$$\begin{aligned} f(x,y) &= \frac{1}{N} F(0,0) + \frac{\sqrt{2}}{N} \sum_{v=1}^{N-1} F(0,v) \cos \frac{(2x+1)v\pi}{2N} \\ &+ \frac{\sqrt{2}}{N} \sum_{u=1}^{N-1} F(u,0) \cos \frac{(2x+1)u\pi}{2N} \\ &+ \frac{2}{N} \sum_{u=1}^{N-1} \sum_{v=1}^{N-1} F(u,v) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \end{aligned} \quad (2.2.2)$$

其中,  $f(x,y)$  是空间域二维向量的元素,  $F(u,v)$  是变化系数阵列的元素, 式中表示的阵列为  $N \times N$ 。

经过 DCT 变换后,其低频分量都集中在矩阵左上角,高频分量分布在矩阵右下角。由于低频分量包含了图象的主要信息,而高频分量



与之相比,不是很重要,所以可以把高频分量置零,只保留低频分量,从而达到压缩和降噪的目的<sup>[5]</sup>。

## 第三章 特征提取

图像在计算机中是以矩阵的形式存储的，其维数通常超出了计算机的能力，解决这个问题常用两个办法，一个是通过缩小图像的尺寸，另一个是使用特征提取的方法。常用的特征提取方法大体可以分为几何特征法、变换系数特征法和代数特征法等几类。本文所运用是代数特征方法中的“特征脸”法。

“特征脸”法，即主成分分析法，具有简单、快速、实用识别率较高的特点，而且有很好的稳定性，位移不变性，特征向量与图像高度成比例变化以及转置不变性。正是由于这些优点，主成分分析方法在人脸识别中应用很广泛，是人脸识别领域中研究的热点。

### 3.1 主成分分析法的原理

主成分分析法是由Turk和Pentlad提出来的一种最好的特征压缩的降维方法。此方法是在数据空间中找出一组向量，用这组向量尽可能的表达数据的方差，从而将数据从高维降到低维。

主成分分析方法利用K-L变换获得逼近原图像空间的最低维识别空间，它把变换人脸的图像区域看作是一种随机向量，采用K-L变换获得其正交K-L基底，其中对应较大特征值的基底具有与人脸相似的形状，因此又称之为特征脸（Eigenfaces）<sup>[6]</sup>。利用这些基底的线性组合可以描述、表达和逼近人脸图像，因此可以进行人脸的识别与合成。

### 3.2 传统的 PCA 方法（Eigenfaces）

为了行文的方便，首先介绍一下矩阵向量化的概念和奇异值分解

定理。

矩阵向量化:

设  $X = (X_1, X_2, \dots, X_n) \in R^{m \times n}$ ,  $Vec(X) \in R^{mn \times 1}$ , 其中,

$X_i$  是  $n$  维列向量, 有:

$$Vec(X) = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \quad (3.2.1)$$

这是把矩阵  $X$  按列向量依次排成一个大的列向量, 常常把这个过程称为矩阵  $X$  的向量化。

(SVD)<sup>[7,8]</sup>:

设  $A \in R_r^{m \times n} (r > 0)$ , 则存在正交矩阵  $U \in R^{m \times m}$  和  $V \in R^{n \times n}$ , 以及对角矩阵  $\Sigma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k, 0, \dots, 0)$ , 满足以下关系:

$$A = U \Sigma V^T \quad (3.2.2)$$

其中,

$$U = [u_1, u_2, \dots, u_k, u_{k+1}, \dots, u_m]$$

$$V = [v_1, v_2, \dots, v_k, v_{k+1}, \dots, v_n].$$

$\lambda_i^2 (i=1, 2, \dots, K)$  是  $AA^T$  的特征根,  $\lambda_i$  称为  $A$  的奇异值。

$u_i, v_i (i=1, 2, \dots, k)$  分别是  $AA^T$  和  $A^T A$  对应于  $\lambda_i^2$  的特征向量;

$u_i (i=k+1, \dots, m)$ ,  $v_i (i=k+1, \dots, n)$  是为了表达上的方便分别引入

的 $(m-k)$ 个列向量和 $(n-k)$ 个行向，可以把它们分别看作 $AA^T$ 和 $A^T A$ 对应于 $\lambda=0$ 的特征向量。这样矩阵 $\Sigma$ 中主对角线上的奇异值元素 $\lambda_i$ 连同 $(n-k)$ 个0构成了 $n$ 维列向量，称为矩阵 $A$ 的奇异值特征向量。

对于一幅大小为 $m \times n$ 的人脸图像，将其每列依次相连则构成一个大小为 $D(D=m \times n)$ 维的列向量，这里的 $D$ 就是人脸图像的维数。设 $M$ 是训练样本的数目， $X_i$ 为第 $i$ 幅人脸图像形成的人脸向量，

则所有训练样本的平均图像向量：

$$u = \frac{1}{M} \sum_{i=1}^M X_i \quad (3.2.3)$$

每个人脸 $X_i$ 与平均脸 $u$ 的差值向量是：

$$\Psi_i = X_i - u, i = (1, 2, \dots, M) \quad (3.2.4)$$

如令： $A = (\Psi_1, \Psi_2, \dots, \Psi_n)$ ，则有：

$$S_T = AA^T, (S_T \in R^{D \times D}) \quad (3.2.5)$$

特征脸由协方差 $S_T$ 的正交向量组成。对于 $m \times n$ 的人脸图像，协方差矩阵 $S_T$ 的维数 $(m^2 \times n^2)$ 很大，直接对它求解特征值和特征向量是很困难的，一种取而代之的方法是解 $A^T A (M \times M)$ 特征值和特征向量。

$$R_T = A^T A \in R^{M \times M} \quad (3.2.6)$$

令 $\lambda_i (i=1, 2, \dots, \gamma)$ 为矩阵 $R_T$ 的 $\gamma$ 个非零特征值， $v_i$ 为 $R_T$ 对应

$\lambda_i$  的特征向量。

根据奇异值分解(SVD)法, 容易得  $S_T$  的特征向量为:

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad (3.2.7)$$

实际上, 选择特征值大的前  $d(d < M)$  个特征向量足够用于人脸识别。因此, 仅取  $U$  的前  $d$  个最大特征值对应的特征向量构造特征空间即可,  $d$  的大小取决于公式(3.2.8)

$$\alpha = \frac{\sum_{i=0}^d \lambda_i}{\sum_{i=0}^r \lambda_i} \quad (3.2.8)$$

$\alpha$  是特征值的累积贡献率。

则在新的特征正交子空间  $U$  中, 人脸样本  $X_i$  就可以表示为

$$Y_i = U^T X_i \quad (3.2.9)$$

由于这些特征向量所对应的图像很像人脸, 所以他们被称为“特征脸”(Eigenfaces), 故通常 PCA 变换进行人脸识别的方法被称为“特征脸”方法。有了这样一个由“特征脸”组成的降维子空间, 任何一幅图像都可以向其投影并获得一组坐标系数, 这组系数表明了该图像在子空间的位置, 从而可以作为人脸识别的依据。由此, 就可以设计分类器对人脸进行分类。

### 3.3 基于类内平均脸的 PCA

传统的 PCA, 是在整个训练集中求取平均脸, 不仅缩小了类间样本的差别, 而且扩大了类内样本的差别, 这样就不利于对人脸进行

分类。为了解决这个问题，本文提出了类内平均脸的方法。类内平均脸的方法不仅有效地扩大了类间样本的识别距离，而且也很好地缩小了类内样本的识别距离，有利于提高人脸识别的准确率。

设样本的维数为  $m \times n$ ，共有  $L$  类， $N_1, N_2, \dots, N_L$  分别表示每一类训练样本的数目， $N$  为训练样本总数。第  $i$  类训练样本集合表示为<sup>[9]</sup>:

$$X_i = \{X_{i1}, X_{i2}, \dots, X_{iN_i}\} \quad (3.3.1)$$

其中  $X_{ij} \in \mathbb{R}^{m \times n}$ ， $N_i$  表示第  $i$  类训练样本的个数。所有训练样本集合是:

$$X = \{X_1, X_2, \dots, X_L\} \quad (3.3.2)$$

第  $i$  类训练样本的类内平均脸定义为:

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij}, \dots, i=1, 2, \dots, L \quad (3.3.3)$$

因此，对第  $i$  类训练样本进行规范化:

$$\Psi_{ij} = X_{ij} - \mu_i, \dots, i=1, 2, \dots, L \quad j=1, 2, \dots, N_i \quad (3.3.4)$$

此时，协方差矩阵定义为:

$$S_T = \sum_{i=1}^L \sum_{j=1}^{N_i} \Psi_{ij} \Psi_{ij}^T \quad (3.3.5)$$

$$R_T = \sum_{i=1}^L \sum_{j=1}^{N_i} \Psi_{ij}^T \Psi_{ij} \quad (3.3.6)$$

令  $\lambda_i (i=1,2,\dots,\gamma)$  为矩阵  $R_T$  的  $\gamma$  个非零特征值,  $v_i$  为  $R_T$  对应  $\lambda_i$  的特征向量。

根据奇异值分解(SVD)定理, 容易得  $S_T$  的特征向量为:

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \quad (3.3.7)$$

从矩阵  $S_T$  的特征向量中, 选取  $m$  个较大特征值对应的特征向量, 即  $u_i, i=1,2,\dots,m$ , 从而构造特征脸空间  $U = [u_1, u_2, \dots, u_m]^T$  其中  $m < n$ 。

为了使测试样本与训练样本具有可比性, 必须用同一个平均脸对他们进行规范化, 为此必须计算所有训练样本的混合平均脸, 即

$$\mu = \frac{1}{N} \sum_{i=1}^L \sum_{j=1}^{N_i} X_{ij} \quad (3.3.8)$$

然后, 对训练样本和测试样进行规范化:

$$\text{训练样本} \quad \Psi_{ij} = X_{ij} - \mu \quad (3.3.9)$$

$$\text{测试样本} \quad \Psi_{ij} = X_{test} - \mu \quad (3.3.10)$$

将  $\Psi_{ij}$  投影到特征脸空间, 即可获得训练样本和测试样本的投影特征为:

$$Y = U^T \Psi_{ij} \quad (3.3.11)$$

以上就是基于类内平均脸的 PCA 方法的介绍, 从中, 可以看出这种方法很类似于传统的 PCA, 不同点只有平均脸的计算, 传统方法计算的是所有混合样本的平均脸, 而基于类内平均脸的 PCA 计算

的是每一个人的训练样本的平均脸。需要强调的一点是，基于类内平均脸的方法也需要计算混合平均脸，不过这里的作用是使训练样本和测试样本具有可比性。



## 第四章 支持向量机

20 世纪 60 年代, Vapnik 等人提出了 VC 维和结构风险最小化归纳原则, 随后经过几十年的研究, 形成了较为完善的理论体系—统计学习理论(Statistical Learning Theory,SLT)<sup>[10]</sup>。

统计学习理论是一种专门研究小样本情况下机器学习规律的理论。在它的基础上建立了一种新的机器学习方法, 即支持向量机(Support Vector Machine,SVM)<sup>[11]</sup>, 它能有效地解决小样本、非线性和过学习等问题, 具有良好的推广性能和较好的分类精确性。

### 4.1 支持向量机的基本思想

支持向量机方法是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的一种小样本学习方法, 在训练样本数目很少的情况下取得很好的推广能力。

支持向量机的基本思想: 在低维样本空间(非线性可分)情况下, SVM 首先选择某种非线性变换, 再将输入的向量映射到一个高维特征空间(Hibert 空间), 其目的是把线性不可分数据集映射到高维特征空间以变为线性可分的数据集, 随后在特征空间建立一个具有最大隔离距离的最优分离超平面<sup>[11]</sup>, 如图 4.1 所示, 其非线性变换是通过核函数的方法来实现的。

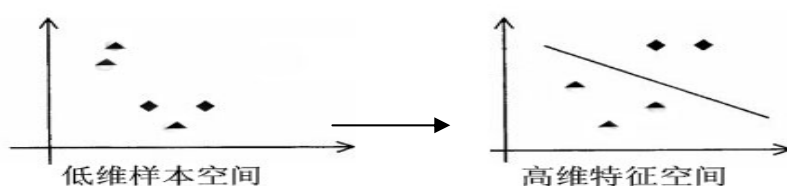


图 4.1 支持向量机的基本思想

Fig4.1 The basic thought of Support Vector Machine

## 4.2 主要的核函数

选择不同形式的核函数就可以生成不同的支持向量机。一般核函数有以下几种形式<sup>[12]</sup>

(1)线性核函数

$$K(x_i, x_j) = x_i^T x_j \quad (4.2.1)$$

(2)多项式核函数

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (4.2.2)$$

(3)径向基核函数

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (4.2.3)$$

(4)两层前馈神经网络核函数

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + \gamma) \quad (4.2.4)$$

## 4.3 最优分类面

支持向量机理论是从线性可分情况下的最优分类面提出的。最优分类面是要求分类线不但能将两类样本无错误的分开，而且要使之间的距离最大<sup>[12]</sup>。其基本思想如图 4.2 所示。

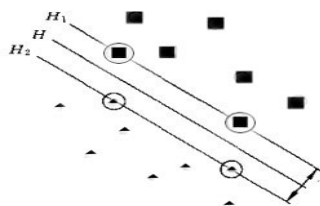


图 4.2 线性可分情况下的两类最优分类线

Fig4.2 The optimal separating line in the separable case

图中两类样本分别用实心圆和正方形代表训练样本， $H$  为样本之间的分类线， $H_1, H_2$  分别为过各类样本中离分类线最近的点且平行于分类线，它们之间的距离叫做分类间隔。所谓最优分类线就是要求分类线不但能将两类正确分开 (训练错误率为 0)，而且使分类间隔最大就是使推广性的界中置信范围最小，从而使真实风险最小。推广到高维空间，最优分类线就成为最优分类面。

设线性可分的<sup>[6,10]</sup>： $(x_i, y_i), i=1, 2, \dots, n, x \in R^d, y \in \{+1, -1\}$ 。

$d$  维空间中的线性判别函数：

$$g(x) = wx + b \quad (4.3.1)$$

分类面方程为：

$$wx + b = 0 \quad (4.3.2)$$

我们将离分类面最近的样本满足  $g(x) = 1$ ，则  $H_1$  和  $H_2$  之间的间隔就是  $\frac{2}{\|w\|}$ ，再要求分类线对所有样本正确分类，就是要求满足：

$$y_i(wx_i + b) - 1 \geq 0, i = 1, 2, \dots, n \quad (4.3.3)$$

因此，满足上述条件且使  $\|w\|^2$  最小的分类面就是最优分类面。

过两类样本中离分类面最近的点且平行于最优分类面的超平面

$H_1, H_2$  上的训练样本，就是使上式等号成立的样本，称作支持向量机。

## 第五章 人脸识别算法实验研究

### 5.1 实验数据

#### 5.1.1 ORL 人脸库的简介

人脸库的选择对人脸识别系统的识别效率影响是很重要的,需要考虑到多方面因素,人脸库的数据既要全面,又要符合人脸识别系统的研究目的,本文采用 AT&T 人脸库又称 ORL (Olivetti Research Laboratory) 人脸库。ORL 人脸库是由英国剑桥大学实验室建立,它包含 40 个人的 400 幅图像,其中每人 10 幅。这些图像都是尺寸为  $112 \times 92$  的灰度图像,它们具有不同的光照、表情、发型、转角,以及睁/闭眼和有无眼镜等区别。ORL 人脸库能够很好的验证人脸识别系统的性能。如图 5.1 所示,展示了 ORL 人脸库中的部分图像。



图 5.1 ORL 中的人脸图像

Fig 5.1 Face samples at ORL database

#### 5.1.2 人脸库图像的重新编号

下载的 ORL 人脸库中,图像的名字不方便使用,因此本文在调用 ORL 人脸库前,重新命名了图像。命名规则是:  $i$  是人脸身份的编

号，表示第  $i$  个人， $j$  是图像的序列号，表示第  $j$  幅图像，“ $i\_j$ ”表示第  $i$  个人的第  $j$  幅图像。

## 5.2 程序运行流程图展示

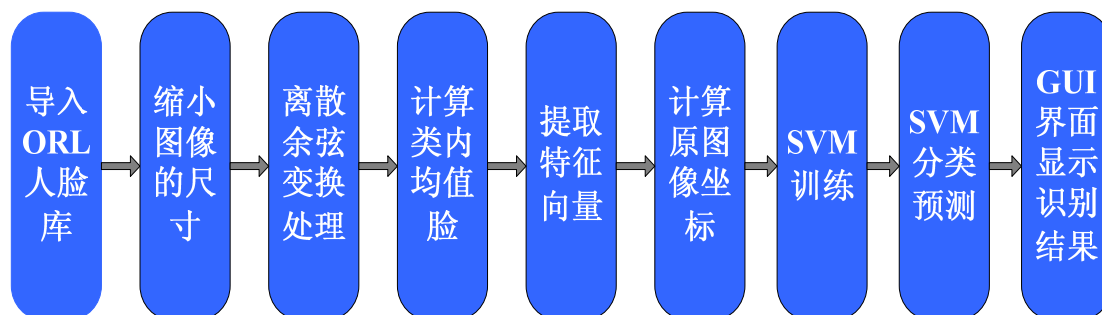


图 5.2 程序运行流程图

Fig5.2 The flowchart of the program running

## 5.3 人脸识别算法的实现

### 5.3.1 导入 ORL 人脸库图像

本文采用的方法是一次将图像全部导入，然后由程序建立数据库存储，其中在使用训练样本时，直接从数据库中调用即可。这样的处理方法，不再需要将图像分为训练集和测试集两部分导入，因此程序更加简洁，运行时间更短。算法是：首先由程序建立数据库，然后按次序导入人脸图像，并将其存储到数据库中，最后当使用训练集时，直接从数据库中提取即可。

流程图如下：

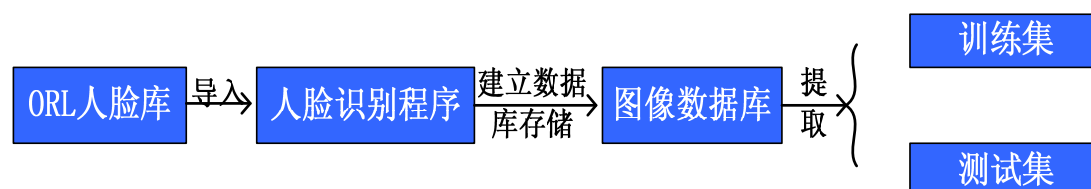


图 5.3 导入 ORL 人脸库图像的流程图

Fig5.3 The flowchart of importing ORL database

程序实现如下:

```
%导入 ORL 人脸库图像
picture=zeros(row,clum,400); %定义数据库 picture
t=1; %循环变量 t 设置为 1
for i=1:40
    for j=1:10
        str=strcat('filepath','filename');%获得图像的路径和名字
        picture(:,t)=imread(str); %读入图像并存储到数据库中
        t=t+1; %循环变量 t 加 1
    end
end
%从 picture 中分离出训练集样本，并存入 photo 中
photo=zeros(row,clum,200);%定义数据库 photo
for k=0:39
    %提取每个人的前 5 幅图像做训练集
    photo(:,:(5*k+1):(5*k+5))=picture(:,:(10*k+1):(10*k+5));
end
```

### 5.3.2 缩小图像的尺寸

ORL 人脸库中的图像尺寸为 $112 \times 92$ ，维数太大，超出了计算机的计算能力。为了解决这个问题，本文在提取特征向量前，通过尺寸变换缩小了图像的尺寸，减少了计算量，节省了大量的时间。算法是：直接调用 MATLAB 中的函数 `imresize()` 调整图像的尺寸到设定的值。

流程图如下:

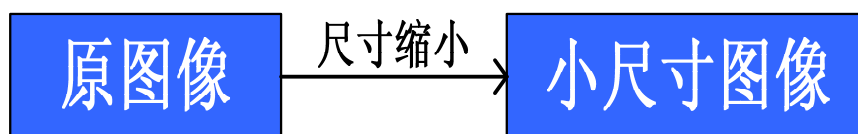


图 5.4 缩小图像的尺寸的流程图

Fig5.4 The flowchart of reducing size

程序实现如下:

```
%缩小图像的尺寸
J=imresize(I,[row,clum]);%把图像 I 的尺寸变换为 row*clum 存入 J 中
```

### 5.3.3 离散余弦变换 (DCT) 处理

DCT 图像压缩的方法在图像处理中应用很广泛。本文采用 DCT 来压缩图像和降低图像的噪声。在图像降噪方面,应用最多的方法是小波变换,但是它用的时间要比 DCT 用的时间多很多,为了节省时间,本文选用了 DCT。算法是:首先调用 MATLAB 中的离散余弦变换函数,选定参数处理图像,得到变换系数,然后调用 MATLAB 中的反离散余弦变换函数恢复图像,并将图像的灰度级化为[0,1]。

流程图如下:



图 5.5 离散余弦变换的流程图

Fig5.5 The flowchart of discrete cosine transform

程序实现如下:

```

%离散余弦变换 (DCT) 处理
J = dct2(I);%对图像 I 进行离散余弦变换
J(abs(J) < 4) = 0;%把绝对值小于 4 的像素的灰度值置零
K = idct2(J)/255;%反离散余弦变换, 恢复图像, 灰度级化为[0,1]
  
```

### 5.3.4 计算类内均值脸

类内均值脸方法是传统平均脸的改进,该方法有效地缩小了类内样本的识别距离,增大了类间样本的识别距离。对于提高人脸识别的准确率起了很大的作用。算法是:将图像矩阵向量化为列向量,然后在训练集样本中计算每个人的均值脸,最后用图像列向量减去各自的类内均值脸,就得到了训练样本的均值化列向量。

流程图如下:

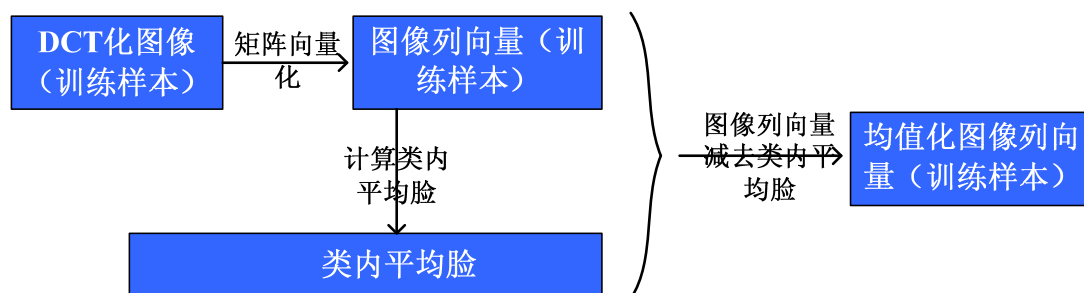


图 5.6 计算类内平均脸的流程图

Fig5.6 The flowchart of calculating the within class average face

程序实现如下:

```

%计算类内均值脸
X=zeros(row*clum,200);%定义矩阵 X, 用来存储经过类内均值脸处理的图像的数据
images=zeros(row*clum,200);%定义矩阵 images, 用来存储向量化后的图像数据
for i=1:200
    images(:,i)=reshape(pho(:, :, i),[row*clum,1]);%变换图像矩阵向量化
end
mn=zeros(row*clum,40);%定义矩阵 mn。用来存储类内均值脸
for i=0:39
    X(:,5*i+1)=images(:,5*i+1)-mn(:,i+1);%类内均值脸处理第 i 个人的第 1 幅图
    X(:,5*i+2)=images(:,5*i+2)-mn(:,i+1);%类内均值脸处理第 i 个人的第 2 幅图
    X(:,5*i+3)=images(:,5*i+3)-mn(:,i+1);%类内均值脸处理第 i 个人的第 3 幅图
    X(:,5*i+4)=images(:,5*i+4)-mn(:,i+1);%类内均值脸处理第 i 个人的第 4 幅图
    X(:,5*i+5)=images(:,5*i+5)-mn(:,i+1);%类内均值脸处理第 i 个人的第 5 幅图
end
  
```

### 5.3.5 用 PCA 和 SVD 提取特征向量

主成分分析方法 (PCA) 是一种最优的信息压缩方法, 奇异值理论 (SVD) 是一种快速计算高维矩阵特征向量的方法。本文采用 PCA 提取图像的特征向量, 采用 SVD 来计算特征向量。算法是: 首先把训练样本的均值化列向量组合成离散方差矩阵, 然后调用 MATLAB 中的 SVD 函数求解离散方差矩阵的特征向量, 最后按照特征值的大小选择特征向量, 构造特征空间。



流程图如下:

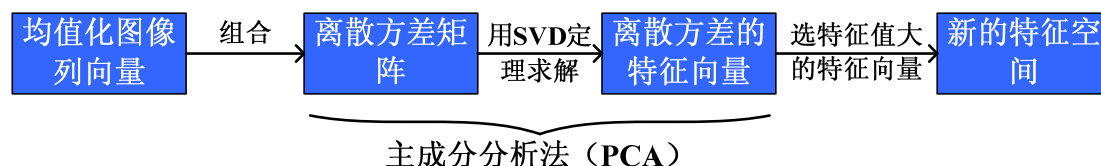


图 5.7 用 PCA 和 SVD 提取特征向量的流程图

Fig5.7 The flowchart of using PCA and SVD to extrat characteristic vector

程序实现如下:

```

%提取特征向量
[ev,ed,v]=svd(X);%用 SVD 函数求特征向量
EV=ev(:,1:48);%提取前 48 维的特征向量
  
```

### 5.3.6 计算图像在新的特征空间中的坐标

为了得到图像的特征需要将图像的数据投影到新的特征空间中。算法是：首先要计算所有训练样本的混合平均脸，并用它把所有的图像都均值化，然后用选取的特征向量线性表出原图像，其中这些线性表出系数就是原图像的新的特征，最后为得到的新的人脸特征数据添加标签。

流程图如下:

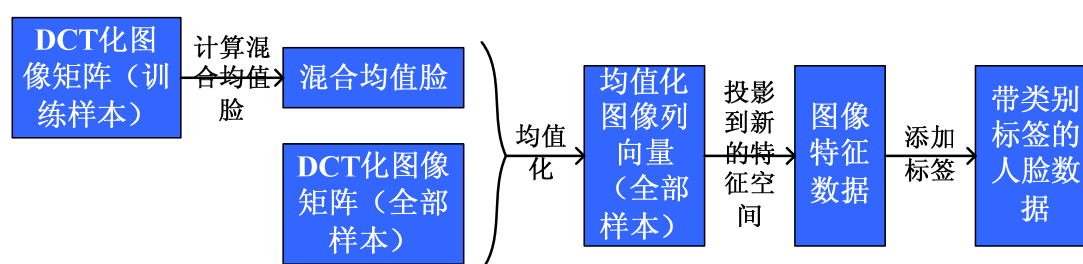


图 5.8 计算线性标出系数的流程图

Fig5.8 The flowchart of calculating coefficient linear

程序实现如下:

```

%计算图像在新的特征空间里的坐标
Y=zeros(row*clum,400); %定义矩阵 Y
date=zeros(48,400); %定义矩阵 date
mn2=mean(images,2);%求训练样本的传统平均脸
for i=1:400
  
```

```

Y(:,i)=reshape(picture(:,i),[row*clum,1]);%原图像矩阵向量化
Z=Y(:,i)-mn2;%原图像数据均值化
date(:,i)=EV'*Z;%用特征向量线性表出原图像
end
Date=date';%矩阵变形
renlian_date=zeros(400,49);%定义矩阵 renlian_date
for i=0:39
    renlian_date((10*i+1):(10*i+10),1)=i+1;%添加类别标签
end
renlian_date(:,2:49)=Date;%把 Date 中的人脸的数据存储到 renlian_date 中

```

### 5.3.7 SVM 训练

本文选择的分类器是支持向量机（SVM）。算法是：首先从人脸特征数据中提取出人脸标签和人脸特征数据，并把人脸标签分为训练集人脸标签和测试集人脸标签，把人脸特征数据分为训练集人脸特征数据和测试集人脸特征数据，再把训练集人脸特征数据和预测集人脸特征数据归一化到[0,1]；然后用训练集人脸标签和训练集人脸特征数据训练 SVM，并选择合适的参数 $-c$ 和 $-g$ ，训练完成后得到了支持向量机模型（model）<sup>[13]</sup>。

流程图如下：

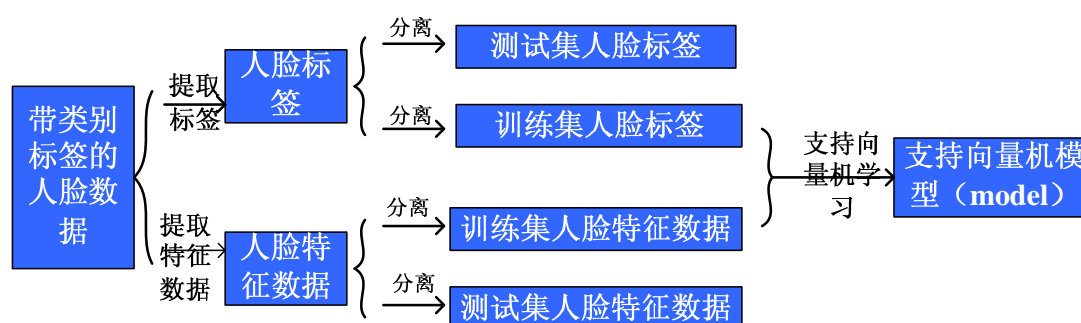


图 5.9 SVM 训练的流程图

Fig5.9 The flowchart of the SVM training

程序实现如下：

```

% SVM 训练
renlian_labels=renlian_date(:,1);%提取人脸标签
renlian=renlian_date(:,2:49);%提取人脸特征数据

```

```

%提取训练集人脸特征数据
train_renlian= [renlian(1:5,:);...; renlian(391:395,:)];
%提取训练集人脸标签
train_renlian_labels=[renlian_labels(1:5,:);...;renlian_labels(391:395,:)];
%提取测试集人脸特征数据
test_renlian= [renlian(6:10,:);...;renlian(396:400,:)];
%提取测试集人脸标签
test_renlian_labels=[renlian_labels(6:10,:);...;renlian_labels(396:400,:)];
%把训练集人脸特征数据和测试集人脸特征数据归一化到[0,1]区间
[train_renlian1,test_renlian]=scaleForSVM(train_renlian,test_renlian,0,1);
%支持向量机的学习，并得出支持向量机模型（model）
model = svmtrain(train_renlian_labels, train_renlian1, '-c 3 -g 0.4');

```

### 5.3.8 SVM分类预测

这是整个程序的最后一步，也是检验支持向量机模型性能优劣的一步。算法是：首先把测试集人脸标签和测试集人脸特征数据的导入到上一步中得到的支持向量机模型（model）中，然后进行分类预测，最后得到人脸识别的准确率。

流程图如下：

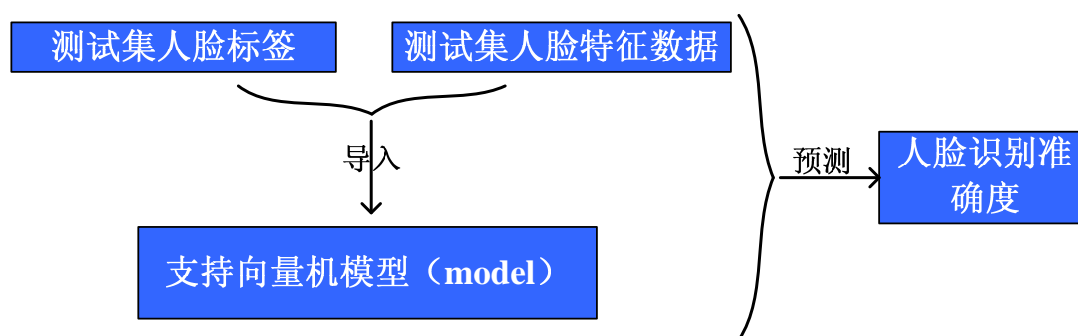


图 5.10 SVM 预测的流程图

Fig5.10 The flowchart of the SVM predicting

程序实现如下：

```

%SVM 分类预测
%调用函数 svmpredict()实现
[predict_label, accuracy] = svmpredict(test_renlian_labels, test_renlian, model);

```

### 5.3.9 GUI 界面设计

为了使用的方便，本文在探究实验的基础上建立了 GUI 界面<sup>[14]</sup>，

并把程序集成到 GUI 中。GUI 界面如图 5.10 所示。



图 5.11 GUI 界面“开始”界面图

Fig5.11 The GUI interface of the "start" interface diagram

设计 GUI 的步骤:

(1)添加“左运行提示”、“右运行提示”“选择人脸编号”、“识别结果”、“准确度显示”和图像显示窗口。它们的功能分别是，显示左边按钮的操作提示，显示右边按钮的操作提示，显示使用者选定的待识别的人脸图像的编号，显示待识别的人脸图像的预测编号，显示程序总的识别准确率和显示选定的待识别的人脸图像。

(2)添加“开始”按钮，写入程序。它的功能是，当点击它时，程序正式开始工作，左提示框显示“导入数据”。

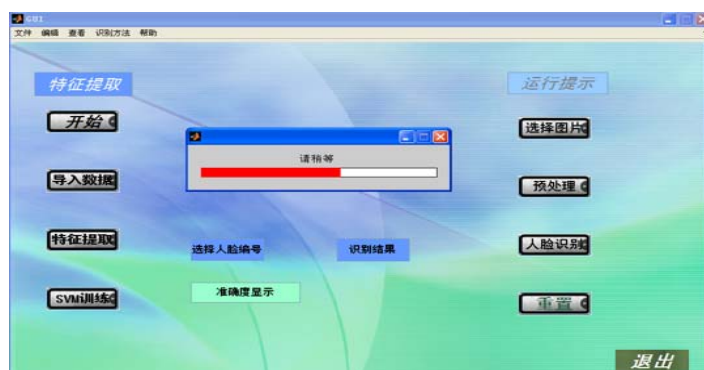
点击“开始”按钮，展示运行结果图



(3)添加“导入数据”按钮，写入程序。它的功能是，点击它时，开始导入 ORL 人脸库，尺寸缩小，DCT 变换，并显示运行进度框，

左提示框显示“特征提取”。

点击“导入数据”按钮，展示运行过程图



点击“导入数据”按钮，展示运行结果图



(4)添加“特征提取”按钮，写入程序。它的功能是，点击它时，进行主成分分析，提取特征向量，并计算出 ORL 人脸库图像的线性特征，左提示框显示“SVM 训练”

点击“特征提取”按钮，展示运行结果图



(5)添加“SVM 训练”按钮，写入程序。它的功能是，点击它时，进行 SVM 训练和 SVM 测试，左提示框显示“运行提示”，右提示框

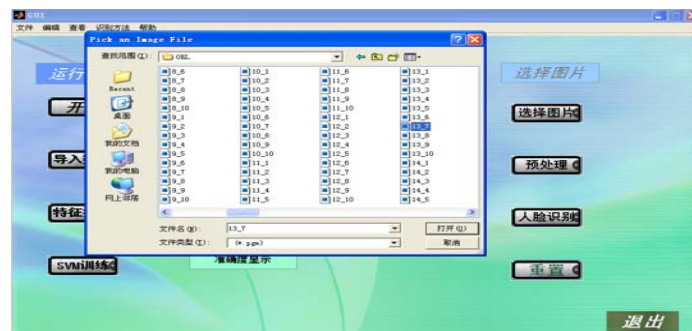
显示“选择图片”

点击“SVM 训练”按钮，展示运行结果图



(6)添加“选择图片”按钮，写入程序。它的功能是，点击它时，打开选择图片的窗口，选择图片后，显示选定的人脸图像和它的编号，右提示框显示“预处理”

点击“选择图片”按钮，展示运行过程图



点击“选择图片”按钮，展示运行结果图



(7)添加“预处理”按钮，写入程序。它的功能是，点击它时，把选定的待识别的人脸图像投影到特征空间里，得到线性表出系数，右提示框显示“人脸识别”



点击“预处理”按钮，展示运行结果图



(8)添加“人脸识别”按钮，写入程序。它的功能是，点击它时，对待识别的人脸图像编号进行预测，显示预测的结果、识别准确度，并弹出“评价”窗口，右提示框显示“重置”。

点击“人脸识别”按钮，展示运行结果图



(9)添加“重置”按钮，写入程序。它的功能是，点击它时，清空“选择人脸编号”、“识别结果”、“准确度显示”和图片显示窗口，为下一次人脸识别做好准备，右提示框显示“选择图片”。

点击“重置”按钮，展示运行结果图



(10)添加“退出”按钮，写入程序。它的功能是，点击它时，关闭 GUI 界面。

(11)美化 GUI 界面。添加背景和按钮图案，使其更加美观。

GUI 运行流程图如下所示：

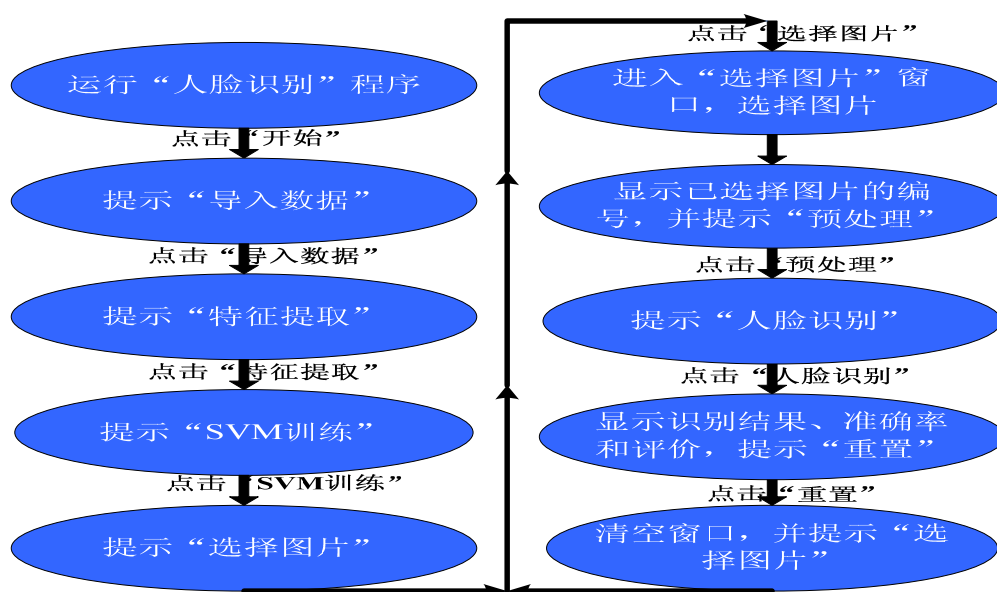


图 5.12 GUI 的工作流程图

Fig5.12 The flowchart of the GUI

## 5.4 影响人脸识别准确率的相关因素探究

### 5.4.1 图像尺寸的影响

图像的尺寸有大有小，如图 5.13 所示，展示了同一个人的不同尺寸的图像。图像的尺寸越大，包含的信息量越大；尺寸越小，包含的信息量越少。

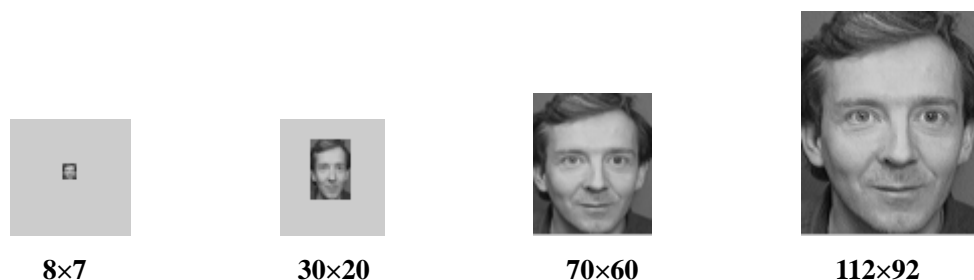


图 5.13 同一个人的不同尺寸的图像

Fig5.13 The same people of the different sizes



虽然在人脸识别的相关文献中并未查到相关资料来证明图像的尺寸的变化会影响识别的准确率，但是我们在实验中得到了表 5.1 所示的数据。从表中，我们可以看出通过改变图像的尺寸，可以影响识别的准确率。

表 5.1 不同尺寸的图像对识别准确度的影响

Tab5.1 Different sizes of the image recognition effecting accuracy

图 像 尺 寸（像 素）	DCT 的 截 止 阈 值	类内平均 脸	特征向量 维数	SVM 参 数		训练样 本（幅/ 人）	测试样 本（幅/ 人）	人脸识别 的准确度
				$-C$	$-g$			
8×7	4	有	48	3	0.4	5	5	99.5%
20×15	4	有	48	3	0.4	5	5	94%
30×25	4	有	48	3	0.4	5	5	93%
40×35	4	有	48	3	0.4	5	5	92%
50×45	4	有	48	3	0.4	5	5	92%
60×55	4	有	48	3	0.4	5	5	93%
70×60	4	有	48	3	0.4	5	5	91%

一般而言，图像的尺寸变换作为一种预处理的方法常用在人脸识别中，但通过查阅文献，我们知道图像变换在预处理中的作用是将不同尺寸的图像归一化。并没有提到图像的尺寸对识别准确率的影响。

在实验的过程中，当 DCT 的截止阈值为 4，基于类内平均脸的 PCA，特征向量维数为 48，SVM 的参数  $-C$  和  $-g$  分别为 3、0.4，每个人的训练样本和测试样本都是 5 个时，缩小图像的尺寸时，发现了识别准确率反而在升高。我们假设：在一定的范围内，缩小图像的尺寸可以提高识别的准确度。按照这个假设，我们在图像尺寸是 8×7 时找到了本实验中的最高准确度 99.5%。

图像尺寸影响准确度的原理：

在人脸图像中，人脸的大小、表情、姿势是不同的，而且在所有的图像中都存在着噪声。图像的尺寸越大，图像包含更多的信息，其中也包括较大的噪声等。我们推测，在一定的范围内缩小图像的尺寸，不仅减少了对识别有利的信息，还减少了对识别不利的信息，而且不利信息减少的速率更快。所以，缩小图像尺寸可以减少各种噪声对识别的干扰，有利于识别准确率的提高。

表 5.2 不同尺寸的图像对识别准确度的影响

Tab5.2 The different sizes of the image recognition effecting accuracy

图像尺寸 (像素)	DCT 的 截止阈值	类内平均 脸	特征向量 维数	SVM 参数		训练样 本(幅/ 人)	测试样 本(幅/ 人)	人脸识别 的准确度
				$-C$	$-g$			
5×5	4	有	48	3	0.4	5	5	90.5%
6×5	4	有	48	3	0.4	5	5	95%
7×6	4	有	48	3	0.4	5	5	95.5%
8×7	4	有	48	3	0.4	5	5	99.5%

然而，随着缩小图像尺寸，图像的信息大量流失，当有利信息的流失大于各种噪声流失时，就会对识别产生不利的影响，所以过度的缩小图像的尺寸，又会降低识别的准确度。如表 5.2 所示，随着图像尺寸的缩小，识别准确率下降。

所以，我们可以得出这样的结论：适当的缩小图像的尺寸，有利于提高人脸识别准确率；但是过度的缩小图像的尺寸，又会降低识别的准确率。

#### 5.4.2 离散余弦变换的影响

离散余弦变换(DCT)本身是一种无损的正交变换，可以除去图

像中的冗余信息,是一种常用的数据压缩方法。在压缩过程中,首先DCT将整体图像分成 $N \times N$ 像素块,再对每一块逐一进行DCT变换;然后设定阈值,舍弃高频的系数,并对余下的系数进行量化以进一步减少数据量;最后再通过反离散余弦变换(IDCT)恢复图像。

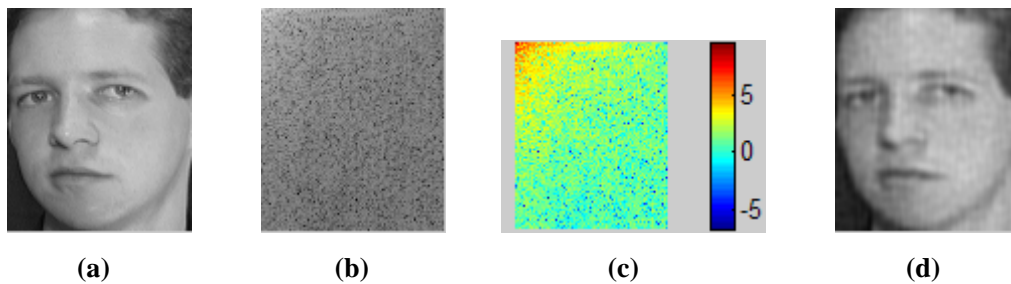


图 5.14 离散余弦变换 (DCT)

Fig5.14 Discrete Cosine Transform (DCT)

本文在预处理过程中,用到了离散余弦变换 (DCT)。如图 5.14 所示,图 a 是 ORL 人脸库中一个人的图像;图 b 显示了离散余弦变换后的系数,可以发现左上角最亮,能量最高;图 c 在前一幅图的基础上用不同的颜色显示了能量分布,其中红色的能量最高,分布在左上角,蓝色的能量最低,分布在右下角;图 d 是用反离散余弦变换恢复的图像。

在人脸识别过程中,光照等其它的噪声会影响识别的结果。本文中,DCT 的作用就是降低这种负影响。

表 5.3 DCT 对识别准确率的影响

Tab5.3 The influence of the DCT for recognition accuracy

离散余弦变换 (DCT)	图像尺寸 (像素)	类内 平均 脸	特 征 向 量 维数	SVM 参数		训 练 样 本 (幅/ 人)	测 试 样 本 (幅/ 人)	人脸识 别的准 确度
				$-C$	$-g$			
有	8×7	有	48	3	0.4	5	5	95.5%
无	8×7	有	48	3	0.4	5	5	97%

DCT 影响识别准确度的原理:

DCT 变换后,数据分为低频区域和高频区域,其中,低频区域的值都较大,分布在左上角,高频区域的值都较小,分布在右下方。而图像中的高亮区和部分噪声分布在高频区,系数都很小,所以可以通过设定阈值,将这部分的系数化为零,除去了光照和其它一些噪声对识别结果的影响。如表 5.3 所示,没有经过 DCT 处理的程序的最高识别率只有 97%;但是,经过 DCT 处理的程序的准确度提高到了 99.5%。由此,证明了 DCT 降低了图像一些噪声和光照的负影响,提高了人脸识别的准确率。

### 5.4.3 类内均值脸的影响

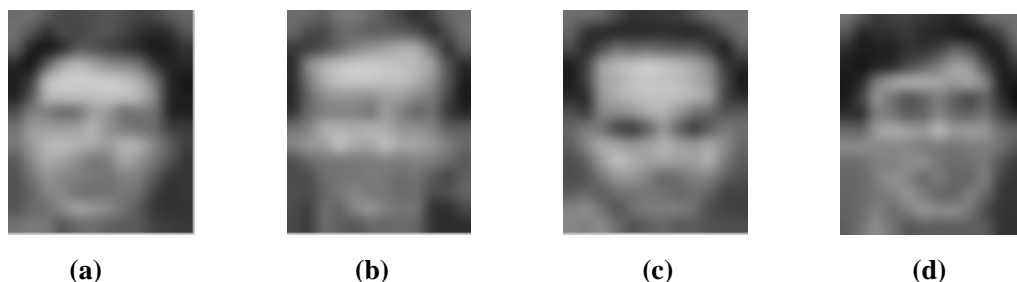


图 5.15 四个人的类内平均脸

Fig5.15 Four men of the within class average face

本文在提取特征向量前,对原图像数据做了均值化处理,用到的平均脸是每一个人的训练样本的类内平均脸。如上图 5.15 所示,展示了四个人的类内平均脸。

传统的方法是用所有训练样本的混合平均脸来均值化原图像的。这样做的好处是使图像的数据以均值为标准分布;但是,这种处理方法不仅扩大了类内样本的差距,而且缩小了类间样本的差距,使人脸识别变得更难。

而类内均值脸处理方法，不仅解决了上述的问题，还保留了传统方法的优点。如表 5.4 所示，不使用类内均值脸，人脸识别的准确率只有 94.5%，然而，使用类内均值脸后，人脸识别的准确率提高到了 99.5%。

表 5.4 类内均值脸对识别准确度的影响

Tab5.4 The influence of the within class average face for recognition accuracy

类内平均脸	图像尺寸 (像素)	离散余弦变换 (DCT)	特征向量维数	SVM 参数		训练样本 (幅/人)	测试样本 (幅/人)	人脸识别的准确度
				$-C$	$-g$			
有	8×7	有	48	3	0.4	5	5	99.5%
无	8×7	有	48	3	0.4	5	5	94.5%

类内平均脸有利于人脸识别的原理：

类内平均脸是指每一个人的所有训练图像的平均值。类内平均脸也是平均脸，所以也使图像的数据分布在均值的附近，不同的是，这里的参考系是每一个人自己的平均脸，而不再是混合平均脸。很显然地类内平均脸缩小了类内样本之间的识别距离，扩大类间样本之间的识别距离，因此，有利于提高人脸识别的准确率，正如表 5.4 所示。

所以，可以得出结论：使用类内均值脸的方法，有利于提高识别的准确率。

#### 5.4.4 特征向量维数的影响

PCA 是人脸识别中的研究热点。PCA 有很多的优点，比如它有简单、快速、识别度较高的特点，而且很好的稳定性，位移不变性，特征向量与图像高度成比例变化以及转置不变性。

如下图所示，就是通过主成分分析方法得到的部分特征脸：



图 5.16 特征脸

Fig5.6 Eigenfaces

主成分分析方法就是求一组相互正交的特征向量，并且选择较大特征值所对应的特征向量来构造特征空间，将原图像投影到这个特征空间里，就得到一组线性表出系数，这些系数就是原图像在该特征空间里的坐标，也是图像的特征。

表 5.5 特征向量的维数对识别准确度的影响

Tab5.4 The influence of the eigenvector dimension for recognition accuracy

特 征 向 量 维 数	图像尺寸 (像素)	DCT 的 截 止 阈 值	类 内 平 均 脸	SVM 参 数		训 练 样 本 ( 幅 / 人 )	测 试 样 本 ( 幅 / 人 )	人 脸 识 别 的 准 确 度
				$-C$	$-g$			
1	8×7	4	有	3	0.4	5	5	6%
8	8×7	4	有	3	0.4	5	5	58.5%
16	8×7	4	有	3	0.4	5	5	81%
24	8×7	4	有	3	0.4	5	5	93%
32	8×7	4	有	3	0.4	5	5	96%
40	8×7	4	有	3	0.4	5	5	97%
48	8×7	4	有	3	0.4	5	5	99.5%

通过查阅文献，我们知道不同的特征向量的组合对原图像的表达程度不同，从而最终影响人脸识别的结果，如表 5.5 所示。在特征向量的选择上，有两种观点，一种观点认为，较大的特征值对应的特征向量对原图像的概括能力更强，所以应该选择那些大的特征值对应的特征向量来构造特征空间；另一种观点认为，有些特征值虽然较大，但是却不利于人脸识别，应该挑选那些有利的特征向量，而不是仅仅

看特征值的大小。

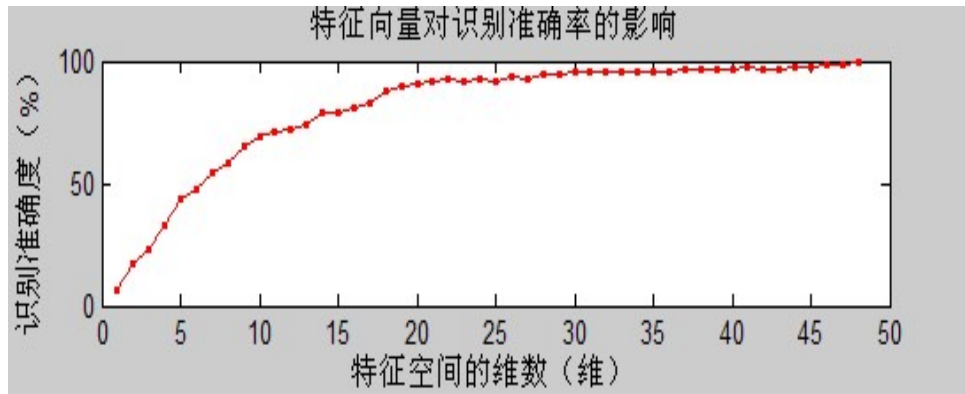


图 5.17 特征向量的维数对识别准确度的影响

Fig5.17 The influence of the eigenvector dimension for recognition accuracy

本文中采用的是第一种方法，即按照特征值的大小选择特征向量。如图 5.17 所示，图中按照特征值的大小，选择了前 48 个较大的特征向量构造特征空间。从图中，不难发现，随着特征向量的个数的增加，识别的准确率在提高。当特征向量的个数是 48 时，准确率达到最大值 99.5%。

所以，我们可以得出结论：特征向量的维数可以影响识别准确率；并且，在一定的范围内随着特征向量数目的增加，识别准确率总体上在提高。

#### 5.4.5 支持向量机 (SVM) 的影响

人脸识别的两个关键的问题就是特征提取方法和分类器的设计，可见分类器的重要性。本文选择的分类器是支持向量机 (SVM)。

支持向量机是在统计学习的理论的基础上发展出来的一种新的机器学习方法。它具有理论完备、适应性强、全局优化、训练时间短、泛化性能好等优点，尤其在解决小样本、非线性及高维模式识别的问题中表现很好。在实际的应用中，SVM 也有很好地表现。如表 5.6

所示，贝尔实验室对美国邮政手写数据字库进行的实验，对四种分类方法进行了比较，我们可以看到，在机器识别中，SVM 是表现的最好的。

表 5.6 几种分类器的对比  
Tab.5.6 Several classifier contrast

分类器	错误率
人工表现	2.5%
决策树 C4.5	16.2%
最好的两层神经网络	5.9%
SVM	4.0%

一方面，人脸识别问题是小样本问题，样本数目有限使很多分类器在人脸识别中表现不是很好，然而，支持向量机很好地解决了小样本的问题。另一方面，支持向量机是非线性的，这样就能解决那些线性不可分的问题。所以本文选择了支持向量机来作为分类器。试验表明，支持向量机在人脸识别中表现的很好：每人 5 个训练样本时，最高的识别准确率达到 99.5%。

#### (1) 参数 $-c$ 和 $-g$ 的影响

本文中采用的是基于径向基核函数的多分类支持向量机分类器。其中，核函数中的  $C$  和  $\text{gamma}$  是两个必选的参数，分别为惩罚参数和间隔，其取值的好坏直接影响识别的准确率。如表 5.7 所示，显示了不同的  $-c$  和  $-g$  组合对应的识别准确度。



表 5.7 核函数参数对识别准确度的影响

Tab5.7 The influence of the kernel parameters for recognition accuracy

$\begin{matrix} -g \\ -c \end{matrix}$	0.1	0.3	0.5	0.7	0.9	1.1	1.3	1.5
0.0001	95.5%	95.5%	95.5%	96%	95.5%	93.5%	90.5%	87%
0.001	95.5%	95.5%	95.5%	96%	95.5%	93.5%	90.5%	87%
0.01	95.5%	95.5%	95.5%	96%	95.5%	93.5%	90.5%	87%
0.1	95.5%	95.5%	95.5%	96%	95.5%	93.5%	90.5%	87%
1	95.5%	97%	98%	97%	96%	96%	95.5%	92.5%
10	97.5%	99%	99.5%	98%	97%	96.5%	96%	94%
100	97.5%	99%	99.5%	98%	97%	96.5%	96%	94%
1000	97.5%	99%	99.5%	98%	97%	96.5%	96%	94%

从表中，我们不难看出，不同的 $-c$ 和 $-g$ 组合有不同的识别准确度。所以，选择合适的参数 $-c$ 和 $-g$ 也是提高识别准确度的途径。但是，目前并没有统一的方法来选择参数，一般而言，可以通过经验选择参数，或者采用基于交叉验证网格搜寻的方法，可惜两种方法都很难找到最好的参数。

本文在实践经验的基础上，结合网格搜寻的思想，找到了很好的参数 $-c$ 和 $-g$ ，将准确率提高到了99.5%。这也是本文对 SVM 应用到人脸识别领域的一个贡献。

## (2) SVM 训练集样本数的影响

在使用 SVM 进行分类时，需要将 ORL 人脸库中的人脸样本分为训练集和测试集两部分。其中，训练集用来训练支持向量机；测试集用来分类测试，从而反映支持向量机的性能。

从分类器的角度讲，参加训练的样本数对最终识别的结果影响是很大的，其中 SVM 也不例外。如图 5.18 所示，反映了训练样本数对

识别准确率的影响。当选择每个人的第一幅图像作为训练样本，其它 9 幅图像作测试样本时，准确率是 89%；当选择每人的前 5 幅图像作为训练样本，其它 5 幅图像作为测试样本时，准确率是 99.5%。

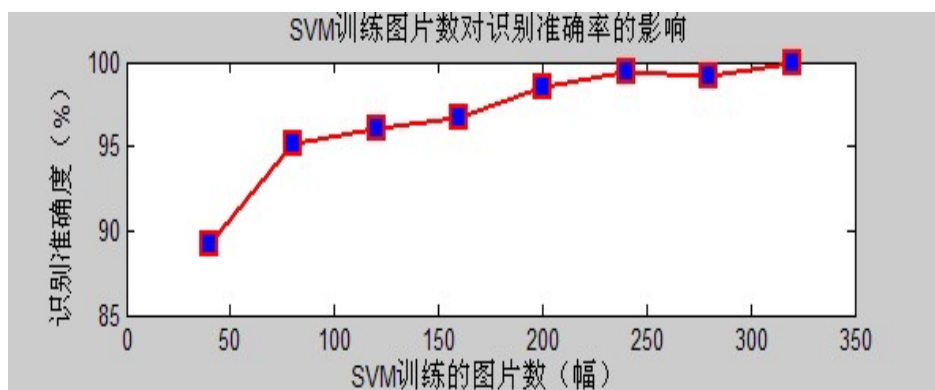


图 5.18 训练样本数目对识别准确率的影响

Fig5.18 The influence of the training sample number for recognition accuracy

人脸识别问题是一个小样本问题，在样本数目较少时，训练样本的数目对最终识别结果影响就会较大。从图 5.18 中的折线图中，我们可以看到随着训练样本数目的增加，识别准确率总体上在提高。所以，我们可以得出这样的结论：训练样本的数目影响识别准确度；在一定的范围内，随着训练数目的增加，识别准确率不断增加。

## 5.5 影响人脸识别速度的相关因素的探究

### 5.5.1 “一次性导入 ORL 人脸库”的影响

查阅有关人脸识别文献，我们发现都是把 ORL 人脸数据库分为训练集和测试集两部分导入：首先在提取特征向量时，需要将训练集导入以计算特征向量，然后在计算样本在特征空间里的线性表出系数时，又需要将训练集和测试集都导入。这样的处理方法，一方面让程序变得复杂，另一方面浪费了很长的时间。

表 5.8 导入 ORL 人脸库的方法对识别速度的影响

Tab5.8 The influence of importing ORL face database method for recognition speed

方法	ORL 人脸库图像数 (幅)	训练集图像数(幅)	导入图像时间 (s)
传统的方法	400	200	5.01
一次性导入	400	200	3.27

本文提出了“一次性将人脸库导入”的方法：把 ORL 人脸库中图像一次全部导入，由程序建立数据库存储，然后当需要训练集时，再直接从数据库中调用。这样处理方法的好处是避免了训练集样本的重复导入，因而可以节省一些时间。

如表 5.8 所示，传统的方法所用的时间是 5.01 秒，而本文提出的方法只有 3.27 秒，比传统的方法省了 1.74 秒——整个程序的运行时间只有 3.5 秒，所以节省了 50% 的时间，极大地提高了程序的效率。

### 5.5.2 图像尺寸的影响

ORL 人脸数据库中的人脸图像的尺寸都是  $112 \times 92$ ，求解特征向量时，需要先计算训练集的离散方差，然后再计算出特征向量。如果直接在原图像上计算，它的离散方差的维数是  $112^2 \times 92^2$ ，超出了计算机的计算能力。所以，在人脸识别中，一般需要缩小原图像。

表 5.9 求解不同尺寸图像的特征向量的时间表

Tab5.9 The time table of solving different size of the image feature vector

图像的尺寸 (像素)	8×7	30×20	70×60	112×92
计算特征向量的时间 (s)	0.0206	0.3084	11.5873	54.0999

图像尺寸的大小对识别的速率影响很大，如表 5.9 所示，展示了求解不同尺寸图像的特征向量所用的时间。从表中不难看出，图像的尺寸越小，计算特征向量所用的时间越短，人脸识别程序的速率越快；相反，图像尺寸越大，计算特征向量所用的时间越长，人脸识别程序

的速率越慢。

表 5.10 图像尺寸对识别速度的影响

Tab5.10 The influence of the image size for recognition speed

图 像 尺 寸 ( 像 素)	DCT 的 截 止 阈 值	类 内 平 均 脸	特 征 向 量 维 数	SVM 参 数		训 练 样 本 ( 幅 / 人)	测 试 样 本 ( 幅 / 人)	程 序 运 行 时 间 (s)
				$-C$	$-g$			
8×7	4	有	48	3	0.4	5	5	3.54
20×15	4	有	48	3	0.4	5	5	3.96
30×25	4	有	48	3	0.4	5	5	4.3
40×35	4	有	48	3	0.4	5	5	5.37
50×45	4	有	48	3	0.4	5	5	7.78
60×55	4	有	48	3	0.4	5	5	12.03
70×60	4	有	48	3	0.4	5	5	16.39

如表 5.10 所示, 当 DCT 的截止阈值、特征向量的维数、SVM 参数  $-C$  和  $-g$ 、训练样本数、测试样本数都相同, 有类内平均脸处理时, 随着图像尺寸的减小, 程序运行的时间在减少, 程序运行的效率也在提高。

缩小图像尺寸可以提高人脸识别的速度原理是:

图像的尺寸越小, 它的离散方差矩阵的维数越小, 计算特征向量时的计算量越小, 计算用的时间越少, 速率也就越快; 相反, 当图像的尺寸比较大时, 它的离散方差也就比较大, 相应的就算特征向量的计算量也就越大, 计算时间自然多, 速率也就越慢。另外, 图像的尺寸越大, 在进行离散余弦变换和计算类内平均脸时, 计算量也会相应的增加, 程序运行的时间增加, 程序效率降低。

### 5.5.3 缩小尺寸和 DCT 顺序的影响

上节讨论了缩小图像尺寸可以提高人脸识别程序的运行效率, 本节将要讨论缩小图像尺寸和 DCT 顺序对程序运行效率的影响。

离散余弦变换 (DCT) 是一种无损的正交变换, 可以除去图像中的冗余信息, 是一种常用的数据压缩方法。在压缩过程中, 首先 DCT 把整体图像分成  $N \times N$  像素块, 再对每一块逐一进行 DCT 变换; 然后设定阈值, 舍弃高频的系数, 并对余下的系数进行量化以进一步减少数据量; 最后再通过反离散余弦变换(IDCT)恢复图像。

表 5.11 缩小图像和 DCT 顺序对识别速度的影响

Tab5.11 The influence of reducing images and DCT sequence for recognition speed

缩小尺寸和 DCT 顺序	图像尺寸 (像素)	类内平均脸	特征向量维数	SVM 参数		训练样本 (幅 / 人)	测试样本 (幅 / 人)	程序运行 时间 (s)
				$-C$	$-g$			
先进行 DCT	8×7	有	48	3	0.4	5	5	7.94
后进行 DCT	8×7	有	48	3	0.4	5	5	3.54

如表 5.11 所示, 当图像的尺寸是  $8 \times 7$ , 有类内平均脸, 特征向量维数是 48, SVM 参数  $-C$  和  $-g$  分别是 3 和 0.4, 训练样本和测试样本都是 5 时, 先进行 DCT 再缩小图像的尺寸, 程序运行的时间是 7.94 秒; 先缩小图像的尺寸后进行 DCT, 程序运行的时间是 3.54 秒, 比前一种方法节省了 4.4 秒, 比程序运行的时间 (3.5 秒) 还要长。

缩小尺寸和 DCT 顺序影响人脸识别效率的原理是:

进行离散余弦变换时, 需要将图像分为  $N \times N$  的像素块, 图像尺寸较小时, 分块数量较少, 计算量较小; 随着图像尺寸增大, 分块数量增加, 计算量也是直线增加, 程序运行时间增加, 人脸识别的效率降低。先进行离散余弦变换时, 图像的尺寸大, 计算量大; 后进行离散余弦变换时, 图像的尺寸已经减小了, 计算量也减小了, 人脸识别的效率自然提高了。

### 5.5.4 DCT、PCA、SVD 和 SVM 的影响

DCT 是一种常用的数据压缩方法，本文中 DCT 的作用是图像压缩和降噪。一方面，图像压缩后，图像的信息量减少，计算量减小，程序运行的时间减少；另一方面，DCT 是一种快速的变换方法，这也是本文选择 DCT，而不是选择小波变换来降低图像的噪声的一个原因。所以，DCT 的使用也对提高程序效率有所贡献。

PCA 是一种最优压缩的变换方法。PCA 的使用减少了图像的信息量，因而对于提高程序效率也有所贡献。

SVD 是一种快速解矩阵的特征向量的方法。本文使用 SVD 解离散方差的特征向量，减少了很大的计算量，节省了很多的时间，对于提高程序效率也有所贡献。

SVM 是近年发展起来的一种机器学习的方法，除了具有全局优化、泛化能力强等优点外，还有训练时间短的特点。选择 SVM 作为分类器，也有利于提高程序效率。

### 5.6 结果与分析

如表 5.12 所示，在各项参数为表中所示的值时，准确率达到了本实验的最高识别准确率 99.5%，程序的运行时间只有 3.54 秒。

表 5.12 最高识别率的各项参数

Tab5.12 The parameters of the highest recognition rate

图 像 尺 寸 ( 像 素 )	DCT 的 截 止 值	类 内 平 均 脸	特 征 向 量 维 数	SVM 参 数		训 练 样 本 (幅/ 人)	测 试 样 本 (幅/ 人)	程 序 运 行 时 间 (s)	人 脸 识 别 的 准 确 度
				-c	-g				
8×7	4	有	48	3	0.4	5	5	3.54	99.5%

如图 5.19 所示，展示了人脸识别的实际效果。从图中，我们可

以看到，在 200 个测试样本中，只有一个样本的识别结果出现错误，正确率达到了 99.5%。

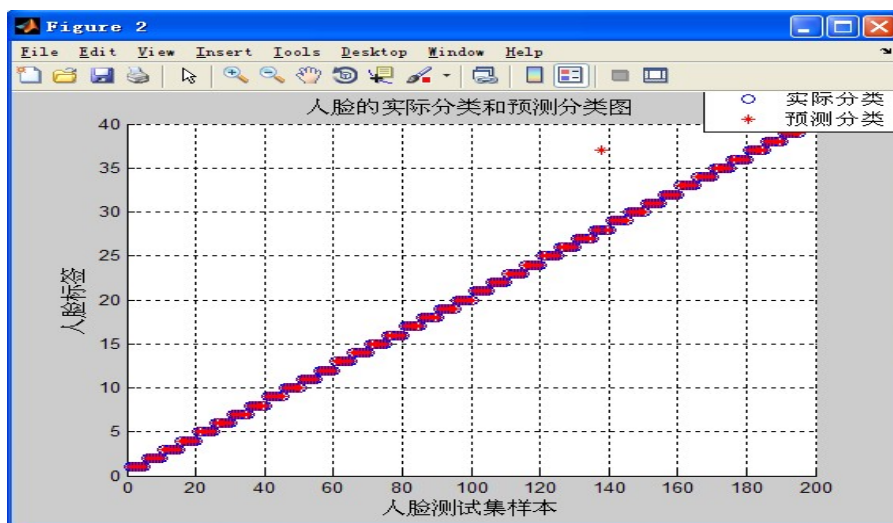


图 5.19 人脸实际分类和预测分类预测图

Fig5.19 the actual and predict classification figure of the face

本文研究了影响识别准确率和效率的相关因素，如表 5.13 所示。通过缩小图像的尺寸、离散余弦变换(DCT)、基于类内平均脸的 PCA 和支持向量机(SVM)等方法的使用，极大地提高了人脸识别的准确率；通过一次导入人脸库和缩小图像的尺寸缩短了程序的运行时间，提高了程序效率。

表 5.13 各种方法对人脸识别速度和准确率的影响

Tab5.13 A kind of methods to speed and accuracy for face recognition of influence

方法 影响	缩小尺寸	DCT	类内均值脸	PCA	SVD	SVM	一次导入人脸库
提高准确率	有	有	有	有	无	有	无
减少时间	有	有	无	有	有	有	有

本文在提高人脸识别的准确率和效率上取得的成绩是很好的。并且对影响识别准确率和效率的相关因素进行了探究，促进了人脸识别在生活中的应用。

## 第六章 总结与展望

人脸识别技术作为一门新兴的生物特征识别技术,有着非常广阔的研究前景。本文着重研究了如何实现人脸的高速、高准确率的识别。

我们结合了缩小图像尺寸、DCT、类内平均脸、PCA、SVM 等方法的优势,成功的实现了这一目标:最高准确率达到 99.5%,程序运行时间是 3.5 秒左右。所以,本文的研究成果将有利于人脸识别在实际中更加广泛的应用。

人脸识别是一个跨学科的前沿课题,相对比较复杂,涉及的内容很多,因此,本文研究在很多方面还需要在今后研究中继续完善和发展。

(1)本文只是对英国 ORL 人脸数据库进行测试,并没有结合别的人脸库(如 PIE 数据库)进行试验,因此,此试验不具备很高的代表性。

(2)每种人脸识别方法都各有优缺点,如何充分利用现有的各种人脸识别方法,将它们进行有效的综合和组合,也是以后一个探索的方向。

(3)本文所研究的人脸识别方法在应用方面有很广泛的前景,希望日后能制造出相应的低成本的人脸识别机器,用于学校或公司等地方。



## 参考文献(References):

- [1]周激流,张晔.人脸识别理论研究进展[J].四川联合大学无线电系.1999(3)
- [2]王振奇.人脸识别技术在安防行业的飞跃与突破[J].技术与应用.2010(8)
- [3]聂祥飞.人脸识别综述[J].重庆三峡学院学报.2009(3)
- [4]杨杰.数字图像处理及 MATLAB 实现[M].北京: 电子工业出版社.2010(2)
- [5]王云平.DCT 算法在图像处理中的应用[J].辽宁工学院学报.2003(2)
- [6]伍威,李晋惠.基于 PCA 和 LDA 的人脸识别技术的研究[J].科技信息(学术版).2008
- [7]胡乡峰,卫金茂.基于奇异值分解(SVD)的图像压缩[J].东北师大学报(自然科学版).2006(9)
- [8]田印中,董志学,黄建伟.基于 PCA 的人脸识别算法研究及实现[J].内蒙古科技与经济.2010(3)
- [9]李晓东,费树岷,张涛.一种改进的模块 PCA 方法及其在人脸识别中的应用[J].复杂工程系统测量与控制教育部重点实验室
- [10]边肇祺,张学工.模式识别[M].北京: 清华大学出版社, 1999.
- [11]张浩然,韩正之,李昌刚.支持向量机[J].计算机科学.2002(1)
- [12]Chih-Chung Chang,Chih-Jen Lin.a Library for Support Vector Machines[J], 2009(2).
- [13]MATLAB 中文论坛.Matlab 神经网络 30 个案例分析[M].北京: 北京航空航天大学出版社, 2010.
- [14]李显洪.Matlab7.x 界面设计与编译技巧[M].北京: 电子工业出版社, 2001.

## 致谢

本论文是在王玉德老师的悉心指导下完成的，从论文的选题、具体研究方法，直到论文的评审等等，王老师都给我们提供很多帮助和指导。他的严谨的治学态度、渊博的专业知识，求实的作风和勤奋忘我的工作都深深激励着我们，在此向王老师表示深深的敬意和由衷的感谢。

同时，要感谢岳光和刘从壮同学为我们的论文提供了帮助和支持。另外，还要感谢我们的父母的大力支持和关心。

最后，感谢所有帮助过我们的人，谢谢你们的无私奉献和大力支持。