

# 基于支持向量机的人脸识别方法<sup>\*</sup>)

崔国勤 李锦涛 高文 焦锋

(中国科学院计算技术研究所数字化技术研究室 北京 100080)

## Face Recognition Using Support Vector Machines

CUI Guo-Qin LI Jin-Tao GAO Wen JIAO Feng

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

**Abstract** Support Vector Machines are a binary classification method and have demonstrated excellent results in pattern recognition. Face recognition is a multi-class problem, where the number of classes is of the known individuals. In this paper we use face data extracted from Eigenfeatures and develop a method to extend SVM to using in multi-class. The training set consists of 5 images of each of the 50 persons equally distributed among frontal, approximately 15° rotated respectively, and the test set consists of 10 images each of the 50 persons. In the ICT-YC face gallery, the proposed system obtains competitive results highly: a correct recognition rate of 94.8% for all the 50 persons, to the less number of the persons and to the famous ORL face gallery we also get good face recognition rate.

**Keywords** Face recognition, Support vector machines, Multi-class problem

## 1. 引言

人脸是人类视觉中的常见模式,人脸识别在安全验证系统、公安(犯罪识别等)、医学、视频会议、交通量控制等方面有着广阔的应用前景<sup>[26]</sup>。现有的基于生物特征的识别技术,包括语音识别、虹膜识别、指纹识别等,都已用于商业应用。然而最吸引人的还是人脸识别,因为从人机交互的方式来看,人脸识别更符合人们的理想。虽然人能毫不费力地识别出人脸及其表情,但人脸的机器自动识别仍然是一个具挑战性的研究领域。由于人脸结构的复杂性以及人脸表情的多样性、成像过程的光照、图像的尺寸、旋转及姿势的变化等,即使同一个人,在不同的环境下拍摄得到的人脸图像也可能不同,所以,虽然人脸识别的研究已有 20 多年的历史,但至今还没有通用成熟的人脸自动识别系统出现。

从算法的实践看,人脸识别不同于很多经典的识别问题,经典的模式识别,譬如文字识别等要处理的是相对较少的类,同时每个类有大量的训练样本,人脸识别中通常处理的是有相当多的类,对于每个类则存在很少的样本<sup>[27]</sup>(譬如身份照),识别算法必须在很少的样本中提取特征,通过训练进行人脸图像的匹配。

统计学习理论(Statistical Learning Theory, SLT)是一种专门研究小样本情况下机器学习规律的理论,该理论针对小样本统计问题建立了一套新的理论体系<sup>[1,2,7,12]</sup>。支持向量机是建立在统计学习理论基础上的解决两类问题的学习方法,由于其快速性和有效性,近年来得到了广泛的研究和应用<sup>[13,20,21]</sup>。我们的人脸识别系统采用 Eigenface 技术<sup>[3]</sup>得到的人脸特征向量表示人脸,然后转化为两类问题,利用支持向量机进行分类,在实验室内部的 ICT-YC 图库和 ORL (the Olivetti Oracle Research Laboratory) 图库<sup>[11,31]</sup>中进行了实验。

## 2. 支持向量机

### 2.1 概述

支持向量机(Support Vector Machines, SVM)是在结构风险最小化<sup>[2]</sup>基础上,对两种不同类别的样本数据找到一个最优分类面的方法。下面简单介绍获取支持向量的过程,希望进一步了解的读者可参考文[1,22,25]。

SVM 算法的输入是一个集合:

$$\{(x_i, y_i) | x_i \in R^d, y_i \in \{-1, 1\}, \text{其中 } i=1, 2, \dots, N\}$$

其中  $y_i$  是样本  $x_i$  的所属类的标志,  $d$  是样本的维数,  $N$  是训练样本的个数。

SVM 首先将输入空间  $R^d$  中的元素  $x$ , 依映射  $\Phi$  投影到一个特征空间  $\Gamma = R^m$  ( $m \geq d$ ), 对应元素  $z = \Phi(x)$ , 使训练样本在特征空间  $\Gamma$  中线性可分。SVM 的目标是在空间  $\Gamma$  中构造一个超平面  $\omega^T z + b$ , 满足:

$$\text{若 } y_i = 1, \text{ 则 } \omega^T z_i + b \geq 1$$

$$\text{若 } y_i = -1, \text{ 则 } \omega^T z_i + b \leq -1$$

其中  $z_i = \Phi(x_i)$ ,  $\omega = \sum_{i=1}^N a_i y_i z_i$ 。

$a^T = (a_1, a_2, \dots, a_N)$  可以通过解下面的二次规划(Quadratic programming, QP)问题得到:

$$W(a) = a^T 1 - \frac{1}{2} a^T Q a \quad (1)$$

$$a \geq 0, a^T y = 0$$

$y^T = (y_1, y_2, \dots, y_N)$ ,  $1^T = (1, 1, \dots, 1) \in R^N$ ,  $Q$  是一个  $N \times N$  半正定对称矩阵, 其元素  $Q_{ij} = y_i y_j z_i^T z_j$ ,  $i, j = 1, 2, \dots, N$ , 为了得到  $Q_{ij}$ , 通常并不需要使用映射  $\Phi$  得到具体的值  $z_i$  和  $z_j$ 。在 Mercer 条件<sup>[2,5,18]</sup>下, 某个特征空间的内积可用一个核函数表示, 即  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) = z_i^T z_j$ , 这时  $Q_{ij} = y_i y_j K(x_i, x_j)$ 。例如, 对于  $d$  次方的多项式分类器  $K(x_i, x_j) = (x_i^T x_j + 1)^d$ 。对于满足  $a_i > 0$  的点, 根据 Karush-Kuhn-Tucker 条件<sup>[14]</sup>, 对应的训练样本应该在高分类面最近的各类样本集合的边界上, 这些样本称为支持向量。

对于一个测试样本  $x \in R^d$ , 首先计算

$$\sigma(x, \omega) = \omega^T z + b = \sum_i a_i y_i K(x, x_i) + b \quad (2)$$

<sup>\*</sup>) 本文工作得到国家 863 计划“生物特征识别核心技术与关键问题研究”(项目编号: 2001AA114190)课题资助。崔国勤 副研究员, 主要研究领域为图像识别、最优化理论和方法。李锦涛 研究员, 博士生导师, 主要研究领域为多模式接口、计算机网络、虚拟现实。高文 教授, 博士生导师, 主要研究领域为多媒体数据压缩、图像处理、计算机视觉。

可得到对应的类别标志

$$\theta(x, \omega) = \begin{cases} +1, & \sigma(x, \omega) > 0 \\ -1, & \text{其他} \end{cases}$$

当训练样本在特征空间  $\Gamma$  中不可分时, SVM 算法引入非负的松弛变量  $\xi_i$ , 其中  $i=1, 2, \dots, N$ , 这时问题转化为求

$$\text{Min } \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (3)$$

$$\text{s. t. } y_i \sigma(x_i, \omega) \geq 1 - \xi_i, i=1, \dots, N$$

这里的  $C$  是用户定义的参数, 用来实现在错分样本的比例和算法复杂度之间的折衷。  $\xi_i$  度量  $\sigma(x_i, \omega)$  和  $y_i$  的差异, (3)式可转化为一个 QP 问题: 在  $0 \leq \xi_i \leq C$  和  $a^T y = 0$  条件下求(1)式的最大值。

### 2.2 多类(k-类)问题的SVM算法

多类问题的模式识别系统可以转化为两类问题, 从而可以应用 SVM 方法<sup>[6,10]</sup>。  $k$ -类 ( $k > 2$ ) 问题就是对于  $l$  类样本  $(x_1, y_1), \dots, (x_l, y_l)$ , 其中  $x_i \in R^d, y_i \in \{0, \dots, k-1\}, 0, \dots, k-1$  为类的标志, 构造一个决策函数, 这个函数能将  $k$  类样本两两分开。

最直观的方法是将多类问题看成一系列两类问题。可作两种考虑, 一是“一对一”算法, 即将每一类和其余的  $k-1$  类中的每一类构成两类, 显然可以得到  $\frac{k(k-1)}{2}$  个分类面。二是

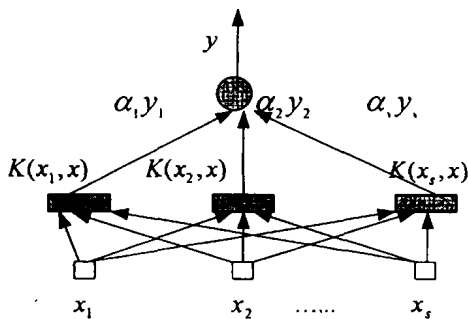


图1 支持向量机分类示意图

“一对多”算法, 即第  $n$  个分类器为解由第  $n$  类样本集合和剩下的  $k-1$  类样本集合全体构成两类问题得到, 这样可得到  $k$  个分类面。

系统采用了后一种方法。为了方便记第  $k$  个算法为  $a_k$ , 对于这  $m$  类样本集合, 可找到一个二类问题的 SVM 分类器  $u_k(x)$  (即由 SVM 得到的第  $k$  个分类函数) 将  $a_k$  和其他算法分开。可得到多类问题的分类器算法  $L(x)$ , 对于任意一个输入  $x$ :

$$L(x) = \arg \max_i \{u_i(x)\} \quad (4)$$

$$\text{其中: } u_i(x) = \sum_{j=1}^l y_{ij} a_{ij} K(x_{ij}, x) - b_i \quad (5)$$

这个算法的缺点是一些测试数据可能被分到好几个类中。为了解决这个问题, 系统对每个测试样本给出前 5 个候选结果。

### 2.3 支持向量机的构造

支持向量机用于分类, 构造的复杂程度取决于支持向量的数目, 而不是特征空间的维数, 这就有效地解决了机器学习中非线性与维数灾难问题, 图 1 就是 SVM 用于分类的构造示意图<sup>[2]</sup>。

输出结果 (决策规则)

$$y = \sum_{i=1}^s \alpha_i y_i K(x_i, \cdot) + b$$

权值  $w_i = \alpha_i y_i$

基于  $s$  个支持向量  $x_1, x_2, \dots, x_s$  的非线性变换 (内积)

输入向量  $x = (x_1, x_2, \dots, x_d)$

## 3. 人脸特征表示和提取

### 3.1 图像库



图2 ICT-YC 人脸库(部分)

都有所不同。图 2 是其中用于实验的一部分图片。



图3 (a)原图 (b)预处理后的图

### 3.2 人脸图像特征数据的提取

人脸识别中应用最广的技术是由 A. Pentland 和 M. Turk 提出的 Eigenface 技术<sup>[3]</sup>, 思路是通过 Karhunen-Loève (简称 K-L) 变换<sup>[4]</sup> 将图像数据映射到一个特征空间上, 即通过主成分分析 (Principle Components Analysis, PCA) 的方法达到降低维数的目的, 这里称降维后的图像数据为人脸的特征数据 (Eigenfeatures), 下面就来讨论人脸识别系统中通过 PCA 提取数据的方法。

#### 3.2.1 预处理过程 在人脸识别系统中, 要达到正确识

别的目的,预处理过程是非常重要的。系统采用了模板匹配的方法,先根据图像中人脸不同器官色素的特殊性,定位出瞳孔的中心,然后将图像缩放旋转使眼睛处于固定的位置,并采用直方图匹配技术使图像具有统一的均值和方差,以部分消除光照强度的影响,同时屏蔽头发和背景。图 3(b)是图 3(a)经过预处理后得到的结果。

以下的算法均基于预处理后的图像。

**3.2.2 特征的抽取** 在统计学习中常需要降低输入变量的维数,具体地说,设原始变量  $x$  为  $m$  维,希望压缩到  $n$  维,  $m > n$ 。如果简单地对  $x$  进行截断(舍去  $x$  的后  $m-n$  个数据),所带来的均方误差等于舍弃的各分量的方差之和,这不能保证得到最优的结果。我们希望找到一个可逆的线性变换  $T$ ,使得对  $Tx$  的截断误差在均方误差意义下为最优,这要求变换后的  $Tx$  的某些分量具有低的方差,主成分分析<sup>[18,29]</sup> 满足了这一要求。

对于一幅  $M \times M$  大小的人脸图像(系统中  $M=64$ ),  $p_{ij}$  表示对应于图的第  $i$  行和第  $j$  列的灰度值,这样人脸图像  $x$  可以用向量表示为:

$$x = (p_{11}, p_{12}, \dots, p_{1M}, p_{21}, p_{22}, \dots, p_{2M}, \dots, p_{M1}, p_{M2}, \dots, p_{MM}) \quad (6)$$

对于一个图像集合  $\{x_1, x_2, \dots, x_N\}$ , 设向量  $x_k \in R^m$  ( $m = M \times M$ ) 为  $c$  类(以区别不同的人脸)人脸图像  $\{\chi_1, \chi_2, \dots, \chi_c\}$  中的一类,通过变换  $Q \in R^{m \times n}$ , 将其从原来的  $m$  维空间映射到  $n$  ( $n < m$ ) 维特征空间  $R^n$ , 即:  $y_k = Q^T x_k, k=1, 2, \dots, N, y_k$  为  $n$  维向量,以达到降低维数的目的。下面讨论  $Q$  的计算过程。

**3.2.3 变换矩阵的计算** 总体散布矩阵  $\Psi \in R^{m \times m}$  为:

$$\Psi = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (7)$$

其中  $N$  为训练样本的个数,  $x_k$  为第  $k$  个图像的向量表示(形如式(6)所示),  $\mu = \frac{\sum_{k=1}^N x_k}{N} \in R^m$  为训练图集的均值。

显然  $\Psi$  为实对称矩阵,其特征值均为非负实数,为求解  $\Psi$  的特征值方程:

$$\Psi \xi = \lambda \xi \quad (8)$$

并对得到的特征向量标准化,用  $\lambda_1, \lambda_2, \dots, \lambda_m$  (非零个数为  $\Psi$  的秩)表示解式(8)可以得到的特征值,则有

$$\Psi \xi_j = \lambda_j \xi_j, j=1, 2, \dots, m \quad (9)$$

$\xi_j$  为  $\lambda_j$  对应的特征向量,设  $\lambda_j$  的排序为  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ , 对应的特征向量构成一个矩阵

$$U = [\xi_1, \xi_2, \dots, \xi_m] \quad (10)$$

则有  $\Psi U = U \Lambda$

其中  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m), U$  为正交阵,即  $U^T = U^{-1}$ , 所以有

$$U^T \Psi U = \Lambda$$

$$\text{展开后有 } \xi_i^T \Psi \xi_j = \begin{cases} \lambda_i, & i=j \\ 0, & i \neq j \end{cases} \quad (11)$$

希望将  $x$  由  $m$  维降到  $n$  维,取式(10)中的前  $n$  维向量,它们对应于  $\Psi$  的前  $m$  个最大的特征值,构成一个  $m \times n$  矩阵  $Q = [\xi_1, \xi_2, \dots, \xi_n]$

就是所需要的  $Q$ ,此时,对一个图像  $x \in R^m$ ,经过变换后得到的  $y = Q^T x$ ,显然

$$y = (x^T \xi_1, x^T \xi_2, \dots, x^T \xi_n)^T \in R^n \quad (13)$$

容易证明这个变换使得经过变换后的  $x$  的均方误差为最小。

**3.2.4 计算特征值和特征向量的方法** 总体散布矩阵<sup>[20]</sup>  $\Psi$  是一个  $m \times m$  维的实对称矩阵,由于  $m \times m$  通常是很

大的(在本系统中为  $4096 \times 4096$ ),直接求其特征向量和特征值计算比较困难,必须寻求一种快速算法。

事实上,由于式(7)中的  $\Psi$  可表示为

$$\Psi = X X^T \quad (14)$$

其中  $X = [x_1 - \mu, x_2 - \mu, \dots, x_N - \mu]$ 。

$N$  的大小通常远小于  $m$ ,求  $\Psi \in R^{m \times m}$  的非零特征值和特征向量,可通过计算矩阵  $\Phi = X^T X \in R^{N \times N}$  的特征值和特征矩阵得到<sup>[32,16]</sup>,显然求  $\Phi$  的特征值和特征向量降低了计算的复杂度。

## 4. 识别系统

从 3.2 节的讨论可以看出,对人脸图像数据,要降低其维数,必须先找到式(10)或式(12)所表示的特征矩阵。为此,必须先对一系列人脸图像数据进行处理,得到 PCA 分析需要的特征空间和特征矩阵。以下的讨论均基于这种特征空间。

### 4.1 SVM 的识别方法

SVM 的构造(见 2.3 节)主要依赖于核函数的选择。由于不适当的核函数可能导致很差的分类结果,并且目前尚没有有效的“学习”使用何种核函数的方法,只能通过实验结果确定采用何种核函数进行训练。系统中采用了两种核函数即 Gauss 函数(Radial Basis Function, RBF)<sup>[23]</sup> 和多项式函数(Polynomial Function, Poly):

$$K_{poly}(x_1, x_2) = (x_1 \cdot x_2 + 1)^d$$

$$K_{Gauss}(x_1, x_2) = e^{-\rho \|x_1 - x_2\|^2} \quad (15)$$

开始实验时  $K_{poly}$  的训练结果没有  $K_{Gauss}$  理想,因此采用 RBF(即  $K_{Gauss}$ )作为整个系统采用的核函数,对应于(2.3 节)的输出,得出的决策面为:

$$y = \sum_{i=1}^n \alpha_i y_i \exp(-\rho \|x - x_i\|_2^2) + b \quad (16)$$

### 4.2 识别过程

对于训练和测试,首先都是将经过预处理的人脸图像投影到经过 K-L 变换得到的原有的人脸图像特征子空间中,系统中特征子空间维数是 64。因此一个人脸图像可用特征空间中的一个 64 维的向量表示,这样的数据即为人脸特征数据。

在训练过程中先对不同人脸的特征数据标上不同的标号,如 0, 1, 2, ..., 同一人脸则标以同样的标号。由(2.2 节)的叙述得知,对一个  $k$ -类人脸特征数据通过支持向量机训练可以产生  $k$  个不同的决策面。可用两类问题的支持向量机得到分类函数  $u_j(x)$ ,并将把第  $j$  类的人脸特征数据和其他类分开。

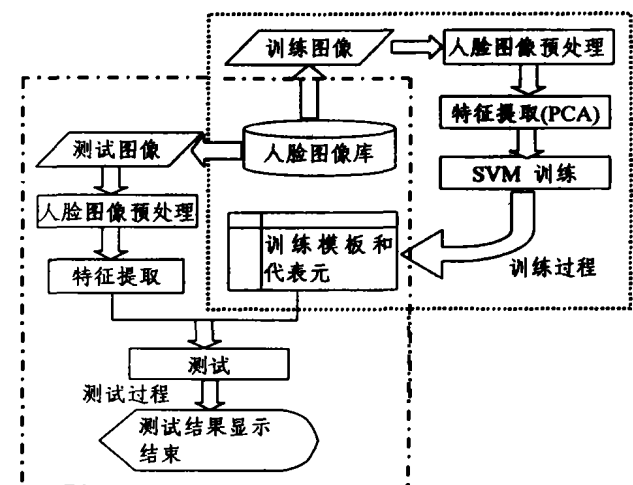


图 4 人脸识别系统框架图

当测试样本  $z$  决定它属于那一个类时,需要计算(4)式中  $u_l(z)$  的值,  $l$  从 1 到  $k$ , 在我们得到  $L(z)$  和相应的类标号  $l_0$  后, 如(2.2 节)所述, 因不能保证所有第一选择都是正确的, 我们选择在  $\{u_l(z)\}$  中离  $l_0$  最近的 5 或多个值对应图像(即类别)的人脸作为候选者。

### 4.3 基于 SVM 的人脸识别框图

识别问题可以被简单地定义为, 给定一系列人的训练图片和一个未标明的人脸测试图片, 给出这个图片所在的类别。分为训练和测试两个模块。图 4 为识别系统框图, 包含了前面讨论的各个子系统。

SVM 算法的很多<sup>[17,19,24,26]</sup>系统中采用了 J. C. Platt 的快速算法<sup>[8,9]</sup>, 其中参数的选择参考了文[22, 25, 30], 并根据实验结果调整。

## 5. 实验结果和分析

我们用 C 语言在 Visual C++ 6.0 平台上实现了通过基于 SVM 的人脸识别系统, 同时在 ICT-YC 人脸库、ORL 人脸库上进行了实验。

### 5.1 实验一

在 ICT-YC 人脸库中使用 SVM 识别人脸图像的实验, 所有的实验对每个人都使用 5 个训练样本, 10 个作为测试。训练样本和测试样本分开, 随机选择。表 1 和图 5 是任意选择 2 到 50 种不同类别的人脸的实验结果。

表 1 对于不同数目的人脸的实验结果(ICT-YC 人脸库)

样本类数	实验次数	平均识别率(%)
2	6	99.4
6	4	96.7
10	4	98.0
11	2	96.4
15	2	97.0
20	2	96.0
30	2	95.4
32	2	94.0
35	2	94.6
40	2	96.0
45	1	95.6
50	1	94.8

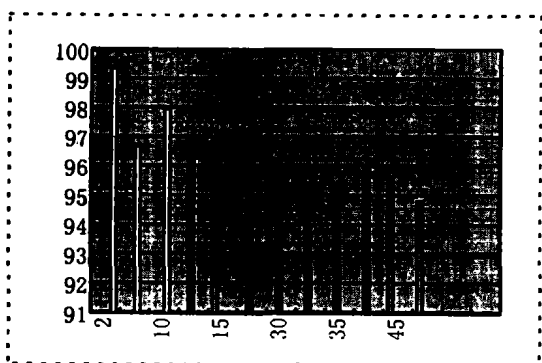


图 5 不同类数的识别率(ICT-YC 人脸库)

### 5.2 实验二

在 ORL 人脸库中, 对其 40 人, 每人 5 张人脸图像作为训练, 5 张作为测试, 实验结果如表 2, 图 6 是 ORL 人脸库和 ICT-YC 人脸库的识别率比较, 从中可以看出, SVM 在这两

种人脸库中都得到了很好的识别效果。从同数目类的识别率来看 ORL 人脸库的实验结果要好于 ICT-YC 人脸库, ICT-YC 人脸库开始必须用比较繁琐的预处理, 其中对眼睛等的定位有些误差。而 ORL 人脸库的图像<sup>[11]</sup>(部分见图 7, 8), 背景单一不需作预处理直接用其 Eigenfeatures 代表人脸图像进行实验, 克服了预处理带来的误差, 这可能是导致两种识别效果的原因。

表 2 ORL 人脸库的实验结果

样本类数	识别率(%)
≤15	100
16	98.75
20	98.0
24	98.4
28	98.57
30	98
35	98.3
40	97

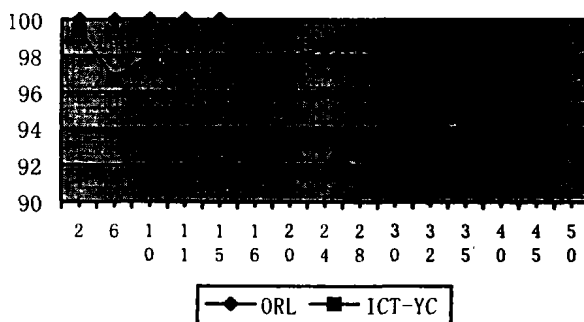


图 6 ORL 人脸库和 ICT-YC 人脸库的识别率比较



图 7 ORL 库中的部分人脸图像

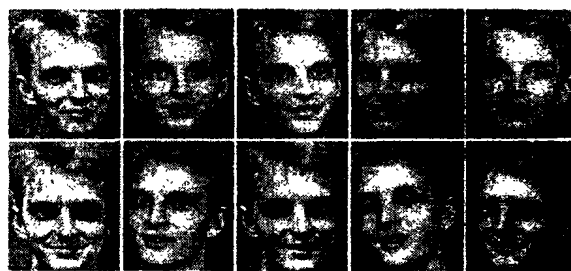


图 8 ORL 人脸库的一个人的 10 张图片

### 5.3 实验结果分析

1998 年 P. J. Phillips 采用“作差法”, 实现了用 SVM 识别人脸图像的一个系统<sup>[27]</sup>, 这种算法建立在不同图像的差异基础上, 即同一个人的不同图像数据之间的差异, 不同人的图像数据之间的差异, 由于这两种不同的差异可以将人脸识别转化为两类问题, 从而得到一个分类器, 我们在系统中实现过这种方法, 在 ICT-YC 人脸库中, 采用 50 人每人 5 张图片训练, 另取每人 5 张图片测试, 识别率在 64% 左右。P. J. Phillips 的

论文<sup>[27]</sup>的测试结果在取 Eigenfeatures 数为 40~80 之间的识别率在 77~78% 之间,这说明我们采取的将多类问题转化为两类问题的方法要优于“作差法”。另外,经过主成分分析后

仅取前面若干数据,同一人的图像的差异(特别是有表情变化时)和不同人的图像之间的差异似乎也没有所期望的明显,足以将它们区分开来(见图 9)。

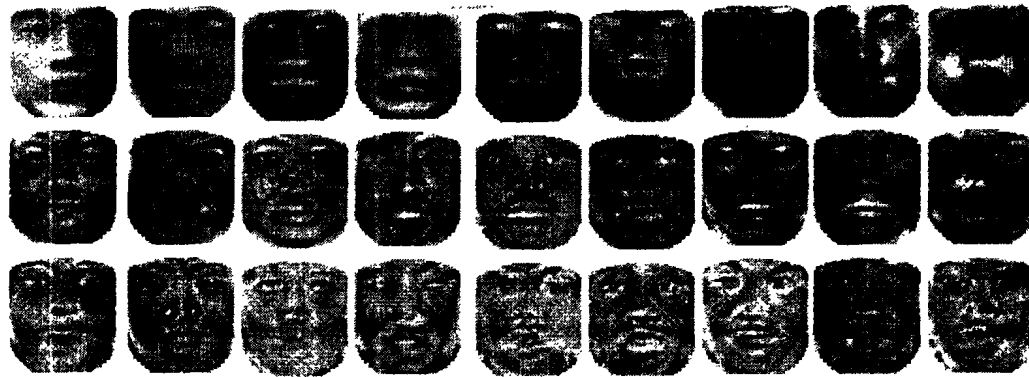


图 9 由 ICT-YC 库人脸图像得到的 64 维 Eigenfeatures 对应的“特征脸”(部分)

**结论** 本文介绍了通过对人的面像进行预处理,使用 Eigenface 技术得到的 Eigenfeatures 数据表示人脸,将人脸识别转化为两类问题,应用支持向量机识别人脸的方法。给出了在 ICT-YC 人脸库、ORL 人脸库上的实验结果,从结果来看, SVM 算法对于人脸识别是可行的。

**致谢** 作者感谢四川银展网讯有限公司的赞助。

### 参 考 文 献

- Burges C J C. A tutorial on support vector machines for pattern recognition. *Datamining and knowledge discovery*, 1998, 2: 121~167
- Vapnik V. *The nature of statistical learning theory*. Springer, New York, 1995
- Turk M, Pentland A. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 1991, 3(1): 71~86
- Kirby M, Sirovich L. Application of the Karhunen-Loève Procedure for the characterization of human faces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990, 12(1): 103~108
- Ries F, Nagy B S Z. *Function Analysis (Vol 2) 泛函 分析讲义 (第二卷)*. 科学出版社, 1980. 110~116
- Collobert R, Bengio S. *Support Vector Machines for Large-Scale Regression Problems*, IDIAP-RR-00-17, 2000
- Snika A, Scholkopf V. A tutorial on support vector regression. *NeuroColt 2*: [TR 1998-03]. 1998
- Platt J C. Fast training of support vector machines using sequential minimal optimization. *Advances in Kernel Methods*, MIT press, 1999. 271~284
- Shevade S K, et al. Improvements to SMO Algorithm for SVM Regression. [Technical Report CD-99-16]
- Weston J, Watkins C. Multi-class Support Vector Machines. [Technical Report CSD-TR-9804]
- ftp://ftp.uk.research.att.com/pub/data/att-faces.tar.z
- Training Support Vector Machines: An Application to Face Detection
- SVM Application List. <http://clopinet.com/isabelle/Projects/SVM>
- 袁亚湘, 孙文瑜. *最优化理论和方法*. 科学出版社, 1999. 422~431
- Platt J C, et al. Large Margin DAGs for Multiclass Classification
- 边肇祺, 张学工, 等. *模式识别 (第二版)*. 清华大学出版社, 2000. 225~226
- Frieb J-T, et al. The Kernel-Adatron Algorithm: A Fast and Simple Learning Procedure for Support Vector Machines
- 阎平凡, 张长水. *神经网络与模拟进化计算*. 清华大学出版社, 2000. 149~151
- Osuna E, Freund R, Girosi F. Training Support Vector Machines: An Application to Face Detection. *Proc. Computer Vision and Pattern Recognition*, 1997. 130~136
- Joachims T. Text Categorization with Support Vector Machines. In: *European Conf. on Machine Learning (ECML)*, 1998
- Papageorgiou C, Oren M, Poggio T. A General Framework for Object Detection. In: *Intl. Conf. on Computer Vision*, 1998
- Vapnik V. *Estimation of Dependencies Based on Empirical Data*. Springer-Verlag, Berlin, 1982
- Scholkopf B, et al. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *IEEE Transactions on Signal Processing*, 1997, 45(11): 2758~2765
- Campbell C, Cristianini N. Simple training algorithms for support vector machines, 1999
- Veropoulos K, Campbell C, Cristianini N. Controlling the sensitivity of support vector machines. In: *Proc. of the Sixteenth Int. Joint Conf. on Artificial Intelligence (IJCA99)*, Stockholm, Sweden. MYGD: <http://www.mips.biochem.mpg.de/>
- Joachims T. Making large-scale SVM learning practical, chapter 11. MIT Press, Cambridge, MA, 1998
- Phillips P J. Support vector machines applied to face recognition. In: M. I. Jordan, M. J. Kearns, S. A. Solla, eds. *Advances in Neural Information Processing System 11*, 1998
- Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: a survey. *Proceedings of IEEE*, 1995, 83(5): 705~741
- Castleman K, Castleman R. *Digital Image Processing*. 清华大学出版社(影印版), 1998. 294~298
- Chapelle O, et al. Choosing Multiple Parameters for Support Vector Machines, submitted to *Machine Learning*, 2000
- Lawrence S, et al. Face Recognition: A Convolutional Neural Network Approach. *IEEE Transactions on Neural Network*, 1997, 8(1): 98~113
- 容观澳. *计算机图像处理*. 清华大学出版社, 2000. 102~105