

# 基于 QR 分解的鉴别维数压缩及其在人脸识别中的应用

杨静宇, 郑宇杰

(南京理工大学 计算机科学系, 江苏 南京 210094)

**摘 要:** 维数压缩是当前模式识别研究领域中的一个重要研究方向. 但是当前部分维数压缩方法缺乏有效的鉴别信息保留机制, 并且在利用 Fisher 鉴别准则的时候经常会遇到小样本问题. 简单介绍了维数压缩中的鉴别信息保留, 并且提出了一种新的直接线性鉴别分析方法——DLDA/QR 算法. 该方法首先利用矩阵的 QR 分解算法实现目标函数的优化, 再在一个较小的空间内实现有效鉴别信息的提取. 在 ORL 人脸数据库上的实验结果验证了算法的有效性.

**关键词:** 鉴别维数压缩; 模式识别; QR 分解; 直接线性鉴别分析

**中图分类号:** TP391.4 **文献标识码:** A **文章编号:** 1673-4785(2007)06-0048-06

## Discriminant dimensionality reduction based on QR decomposition and its application in face recognition

YANG Jing-yu, ZHENG Yu-jie

(Department of Computer Science, Nanjing University of Science & Technology, Nanjing 210094, China)

**Abstract:** Dimensionality reduction is an important research topic in pattern recognition. At present, partial dimensionality reduction methods lack an effective discriminant information preservation mechanism. And the small sample size problem often occurs when the Fisher discriminant criterion is used. In this paper, discriminant information preservation is briefly presented and a new direct linear discriminant analysis (DLDA) method, the DLDA/QR algorithm, is suggested. With this algorithm, the objective function is optimized through QR decomposition, and then the effective discriminant information is extracted from a smaller space. Experimental results from the ORL face database demonstrate the effectiveness of the proposed method.

**Key words:** discriminant dimensionality reduction; pattern recognition; QR decomposition; direct linear discriminant analysis (DLDA)

随着信息技术和互联技术的发展,使人们能够足不出户而接触到海量的高维信息,如何从这些高维信息中抽取得到有效的信息已成为人们日常生活中的一个必要环节.因此,维数压缩已经成为一个普遍存在的问题.维数压缩的问题可以定义为:在高维的观察样本空间中寻找其隐藏的有意义的低维数据结构,并分析和探索事物的内在规律.例如,当一个人在面前出现时,大脑和视觉系统便能够迅速地捕捉到他的典型特征,并利用这些为数不多的典型特

征很快分辨出他是谁.这个捕捉典型特征的过程实质上就是维数压缩的过程.

总体而言,当前的维数压缩方法可以分为2类:线性维数压缩和非线性维数压缩.对于线性维数压缩,常用的准则有 Bayes 最小错误率准则、相互信息准则、Fisher 鉴别准则等,其中 Fisher 准则<sup>[1-2]</sup>近年来受到普遍的重视和广泛的研究.目前,基于线性维数压缩算法的研究侧重点在于小样本问题,该问题和样本数据的高维性是紧密相关的,因此在维数压缩研究中是不可避免的.线性维数压缩的方法除了经典的主成分分析和 Fisher 鉴别分析之外<sup>[3]</sup>,还有最近兴起的独立成分分析方法<sup>[4-7]</sup>,该方法已经受

收稿日期:2007-04-17.

基金项目:国家自然科学基金重点资助项目(60632050);国家自然科学基金资助项目(60472060).

到了普遍的关注.对于非线性维数压缩,近年来无论理论研究还是应用开发都取得了长足的发展.其中,它的 2 个较大的分支值得注意.其一是基于 kernel 的特征提取技术<sup>[8-9]</sup>,以支持向量机、kernel 主分量分析和 kernel 鉴别分析为代表;其二是以流形学习为主导的维数压缩理论和技术<sup>[10]</sup>.目前典型的方法有主曲线(曲面)分析<sup>[11]</sup>、等距特征映射(isometric feature mapping)<sup>[12]</sup>、局部线性嵌入法(local linear embedding)<sup>[13]</sup>、拉普拉斯特征图(Laplacian eigenmap)<sup>[14]</sup>等.

本文针对维数压缩中的鉴别信息提取,对一种已有的解决小样本问题的直接线性鉴别分析方法(direct linear discriminant analysis, DLDA)<sup>[15-16]</sup>提出了新的解决方案,利用矩阵的 QR 分解实现数据的预处理,并且在低维的空间内实现了特征提取,有望实现算法的实时处理.最后,在 ORL 人脸数据库上的实验结果验证了提出方法的有效性.

## 1 鉴别维数压缩的基本概念和方法

设  $1, 2, \dots, c$  为  $C$  个模式类,模式  $X$  为  $n$  维实向量.那么由该训练样本集构成的类间散布矩阵  $S_b$ 、类内散布矩阵  $S_w$  和总体散布矩阵  $S_t$  分别定义为

$$S_b = \sum_{i=1}^c P(i) (\mu_i - \mu) (\mu_i - \mu)^T, \quad (1)$$

$$S_w = \sum_{i=1}^c P(i) E\{ (X - \mu_i) (X - \mu_i)^T / i \}, \quad (2)$$

$$S_t = S_b + S_w = E\{ (X - \mu) (X - \mu)^T \}. \quad (3)$$

式中:  $P(i)$  为第  $i$  类训练样本的先验概率,  $\mu_i = E\{X / i\}$  为第  $i$  类训练样本的均值,  $\mu = E\{X\} = \sum_{i=1}^c P(i) \mu_i$  为全体训练样本的均值.由式(1)~(3)的定义可知,  $S_w$ 、 $S_b$ 、 $S_t$  均为非负定矩阵.当  $S_w$  可逆时,  $S_w$  与  $S_t$  均为正定矩阵,且满足  $S_t = S_b + S_w$ .

### 1.1 主成分分析

主成分分析的思想来源于 KL 变换<sup>[1-2]</sup>,其目的是通过线性变换找到一组最优的单位正交向量基(也称主成分,主元),用这组向量基的线性组合来重建原样本,并使得重建后的样本和原样本的误差最小.在数学上,主成分分析就是通过求解特征值问题来对角化协方差矩阵  $S_t$ .其对应的广义特征方程为

$$= S_t. \quad (4)$$

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.若将 KL 变换用于人脸识别,则需假设不同人脸具有可分性.由于高维图像空间 KL 变换后可得到一组新的正交基,因此可通过保

留部分正交基,以生成低维人脸空间,而低维空间的基是通过分析人脸训练样本集的统计特性来获得.在新的正交空间中,原始人脸样本  $X$  就可以表示为

$$\tilde{X} = T^T (X - \mu). \quad (5)$$

如将子空间的这些正交向量按照图像阵列排列,则可以看出这些正交向量呈现人脸的形状,因此这些正交向量也被称作特征脸.这种人脸识别方法也叫做特征脸(eigenface)方法.关于正交向量的选择有不同的选择依据,即与较大特征值对应的正交向量(主分量)可用来表达人脸的大体形状,而具体细节需通过与小特征值相对应的特征向量(次分量)来加以描述.通过选择前  $m$  个最大特征值所对应的特征向量,可以实现原始人脸样本数据向这个低维子空间的投影.

KL 变换在 20 世纪 90 年代初受到了很大的重视,实际用于人脸识别也取得了很好的效果.从压缩的角度来看,KL 变换是最优的,它不仅使得从  $n$  维空间降到  $m$  维空间前后的均方误差最小,而且变换后的低维空间有很好的脸表达能力.基于 KL 变换的主成分分析方法虽然实现了高维人脸的低维表示,但是从鉴别的角度而言并不是最优的.传统的主成分分析方法对图像的重构而言是最佳的,但是对于图像的分类却缺少有效的鉴别信息.选择训练样本的散布矩阵作为 KL 变换的生成矩阵,是由于其最大特征向量抓住了该样本几何的主要分布,但这是图像统计的方法,而不是人脸统计方法,它虽然考虑图像之间所有的差别,但由于它不管这样的差异是由照明、发型变更或背景所致,还是属于人脸的内在差异,因此特征脸的识别方法用于人脸识别还存在理论的缺陷,研究表明,特征脸的方法随着光线、角度及人脸的尺寸等因素的引入,识别率急剧下降.因此,主成分分析方法根本局限性在于对于分类而言不是最优的.

### 1.2 线性鉴别分析

鉴别分析的思想最早是由 Fisher 提出的<sup>[17]</sup>,其目的是选择使得 Fisher 准则函数达到极值的向量作为最佳投影方向,从而使得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.在 Fisher 思想的基础上,Wilks 和 Duda 分别提出了鉴别矢量集的概念<sup>[18-19]</sup>,即寻找一组鉴别矢量构成子空间,以原始样本在该空间内的投影矢量作为鉴别特征用于识别.该方法被称为经典的 Fisher 线性鉴别分析方法.近来, Fisher 线性鉴别分析无论在理论上还是在应用上都取得了长足的发展,成为一种广泛使用的,十分有效的特征抽取方法.

Fisher 准则函数可定义为

$$J_f(X) = \frac{S_b}{S_w} \tag{6}$$

推广的 Fisher 准则函数定义为

$$J(X) = \frac{S_b}{S_t} \tag{7}$$

当类内散布矩阵  $S_w$  非奇异时,式(6)、(7)两准则完全等价;当  $S_w$  奇异时,准则(7)是准则(6)的合理推广.在类内散布矩阵非奇异的情况下,由 Fisher 准则函数就可以得到原始样本的最优鉴别矢量集.其求解过程也就是对下列广义特征方程求取相应的特征值和特征矢量的过程:

$$S_b = S_w \tag{8}$$

式中:  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)$  是广义特征方程(8)的前  $d$  个最大的非零特征值  $\lambda_1 > \lambda_2 > \dots > \lambda_d$  所对应的特征向量.因此,上述的特征矢量集同时也是矩阵  $S_w^{-1} S_b$  所对应的特征向量集.但是,当小样本问题存在的时候,就会导致相应的类内散布矩阵是奇异的,此时就不能直接得到  $S_w^{-1}$  的值,因此如何在小样本问题下利用 Fisher 鉴别准则得到最佳鉴别矢量集是有效抽取特征的一个关键.

## 2 直接线性鉴别分析的新算法

### 2.1 直接线性鉴别分析方法

直接线性鉴别分析方法是解决当利用 Fisher 鉴别准则进行特征提取时由于类内散布矩阵的奇异而带来的小样本问题.其基本思想是当类间散布矩阵不为零时,此时的类内散布矩阵零空间中就包含了所有的有效鉴别信息.令  $N_b$  和  $N_w$  分别表示样本的类间散布矩阵  $S_b$  和类内散布矩阵  $S_w$  的零空间,那么他们所对应的非零空间分别为  $N_b = R^n - N_b$  和  $N_w = R^n - N_w$ .直接线性鉴别分析方法的目的是从空间  $N_b - N_w$  抽取得到相应的鉴别信息.

### 2.2 一种新的直接线性鉴别分析算法及理论框架

在直接线性鉴别分析的基础上,提出了一种新的 DLDA 算法:DLDA/QR 算法.该算法的目的是在保留 DLDA 算法思想的基础上,降低算法的复杂度,同时提高算法的有效性和稳定性.

该算法首先是解决如下的优化问题:

$$G = \arg \max_{G^T G = I} \text{trace}(G^T S_b G) \tag{9}$$

从式(9)中,可以得到新算法的第 1 步中只是关注类间散布矩阵的最大.在文中的算法中,首先定义矩阵:

$$H_b = [\sqrt{N}(\mu_1 - \mu) \dots \sqrt{N}(\mu_C - \mu)] \tag{10}$$

并且满足

$$S_b = H_b H_b^T \tag{11}$$

由此,可以得到关于矩阵  $H_b$  的 QR 分解<sup>[20-21]</sup>

$$H_b = (Q_1, Q_2) \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{12}$$

式中:  $Q_1 \in R^{n \times t}$ ,  $Q_2 \in R^{n \times (n-t)}$ ,  $R \in R^{t \times t}$ .

此时,很容易得到  $H_b = Q_1 R$  为  $H_b$  的满秩分解.因此对于任意的正交矩阵  $W \in R^{t \times t}$ ,

$$G = Q_1 W \tag{13}$$

这样就可以解决式(9)的优化问题.特别要注意的是,其中  $t$  的是矩阵  $H_b$  的秩,并且  $t$  的上限是  $C-1$ .在实际使用之中, $C$  类数据通常是线性独立的,在这种情况下,QR 分解之后的维数通常为  $t = C-1 = \text{rank}(S_b)$ .因此,通过 QR 分解也可以实现类似 PCA 的降维过程.

DLDA/QR 算法的第 2 步将关注类内散布矩阵的信息.在这一步骤中,此时的特征提取过程类似经典的 DLDA 方法,要注意的一点是此时在一个低维的空间内进行矩阵操作,因此使得 DLDA/QR 方法变得稳定和有效.当得到矩阵  $Q_1$  之后,那么可以得到在新的降维空间之中的相应散布矩阵为

$$\tilde{S}_t = Q_1^T S_t Q_1 \tag{14}$$

和

$$\tilde{S}_b = Q_1^T S_b Q_1 \tag{15}$$

很容易验证  $\tilde{S}_t$  和  $\tilde{S}_b$  是  $t \times t$  的矩阵,并且其中的  $\tilde{S}_b$  是非奇异的.

在这一步骤中,目的是要找到一个同时对角化  $\tilde{S}_t$  和  $\tilde{S}_b$  的矩阵:

$$V^T \tilde{S}_t V = \Lambda, V^T \tilde{S}_b V = I \tag{16}$$

式中:  $\Lambda$  是一个对角矩阵,它的对角元素按照从小到大的次序排列; $I$  是一个单位化矩阵.

首先对角化对称矩阵  $\tilde{S}_b$ .由于  $\tilde{S}_b$  是一个  $t \times t$  的矩阵,并且在通常情况下  $t \ll n$ ,因此很容易实现这一矩阵的对角化.

现假定存在一个矩阵  $U$  满足

$$U^T \tilde{S}_b U = \Lambda \tag{17}$$

式中:  $U^T U = I$ ,  $\Lambda$  是对角矩阵并且其对角元素按照降序排列.

令

$$I = U \Lambda^{-1/2} \tag{18}$$

那么,可以得到

$$(U \Lambda^{-1/2})^T \tilde{S}_b (U \Lambda^{-1/2}) = I \Rightarrow Z^T \tilde{S}_b Z = I \tag{19}$$

此时就实现了新的类间散布矩阵  $\tilde{S}_b$  的对角化.由此,令

$$S_t = Z^T \tilde{S}_t Z \tag{20}$$

利用相似的方法,可以实现总体散布矩阵  $S_t$  的

对角化. 假定存在一个矩阵  $Y$  满足

$$Y^T S_t Y = \Lambda_t, \quad (21)$$

式中  $\Lambda_t: Y^T Y = I$ ,  $\Lambda_t$  是一个对角矩阵, 并且其对角元素按照升序排列.

通常, 选择前  $s(s \leq t)$  个最小特征值对应的特征向量. 假定由这  $s$  个特征向量构成了矩阵  $P = (y_1, \dots, y_i, \dots, y_s)$ , 其中的  $y_i$  是矩阵  $Y$  的第  $i$  列. 因此, 可以得到

$$P^T S_t P = \Lambda_{t,s}, \quad (22)$$

$$\text{令} \quad V_s = ZP, \quad (23)$$

因此, 可以得到

$$V_s^T \tilde{S}_t V = P^T Z^T \tilde{S}_t ZP = P^T S_t P = \Lambda_{t,s} \quad (24)$$

和

$$V_s^T \tilde{S}_b V = P^T Z^T \tilde{S}_b ZP = P^T I P = I. \quad (25)$$

综上所述, 得到了矩阵  $V_s = ZP$  能同时对角化  $\tilde{S}_b$  和  $\tilde{S}_t$ .

因此, 可以得到如下的最终转换矩阵

$$E = Q_1 ZP_s^{-1/2}, \quad (26)$$

那么, 对于一个检测样本  $x_{\text{test}}$  而言, 其在低维空间上的投影数据可以通过

$$x_{\text{test}} = E^T x_{\text{test}} \quad (27)$$

得到.

上述的同时对角化矩阵过程可以通过这样的定理来说明.

**定理 1** 可以同时对角化 2 个对称矩阵  $\tilde{S}_t$  和  $\tilde{S}_b$ ,  $V^T \tilde{S}_t V = \Lambda_t$ ,  $V^T \tilde{S}_b V = I$ , 其中的  $V$  和  $\Lambda_t$  是矩阵  $\tilde{S}_b^{-1} \tilde{S}_t$  的特征向量和特征值, 并且满足  $\tilde{S}_b^{-1} \tilde{S}_t V = V \Lambda_t$ .

假定对角矩阵  $\Lambda_t$  的对角元素按照升序排列. 相应地, 最后的转换矩阵是  $E = Q_1 V_s \Lambda_{t,s}^{-1/2}$ , 其中的  $V_s$  是由  $V$  的前  $s$  列向量所构成的  $t \times s$  的矩阵,  $\Lambda_{t,s}$  是由矩阵  $\Lambda_t$  所构成的一个  $s \times s$  的矩阵.

基于以上讨论, 得到完整的 DLDA/QR 算法为

1) 在原始样本的基础上得到  $H_b$ ,  $S_b$  和  $S_t$ . 利用 QR 分解得到  $Q_1$ , 此时, 通过  $Q_1$  可以得到新的散布矩阵  $\tilde{S}_t = Q_1^T S_t Q_1$  和  $\tilde{S}_b = Q_1^T S_b Q_1$ .

2) 通过计算得到矩阵  $(\tilde{S}_b)^{-1} \tilde{S}_t$  的特征值和特征向量, 分别用  $\Lambda_t$  和  $V$  表示. 假定矩阵  $V$  的前  $s(s \leq t)$  个特征向量对应着矩阵  $\Lambda_t$  的前  $s$  个最小特征值, 由这  $s$  个特征值和对应的特征向量所构成的矩阵分别为  $\Lambda_{t,s}$  和  $V_s$ . 那么可以得到最终的转换矩阵  $E = Q_1 V_s \Lambda_{t,s}^{-1/2}$ .

3) 利用公式 (27) 将样本投影到低维空间并进行分类.

### 3 实验结果

采用 ORL 人脸数据库作为实验数据. ORL 人脸数据库由 40 个人的脸部图像组成, 每人由不同的 10 幅图像构成, 人脸图像原始维数为  $92 \times 112$ , 其中有些图像拍摄于不同的时期; 人脸脸部表情与脸部细节均有变化, 如笑或不笑、眼睛睁着或闭着、戴或不戴眼镜; 人脸姿态有变化, 深度旋转与平面旋转可达到  $20^\circ$ , 人脸尺度也有 10% 的变化. 该人脸数据库中的部分人脸图像如图 1 所示.

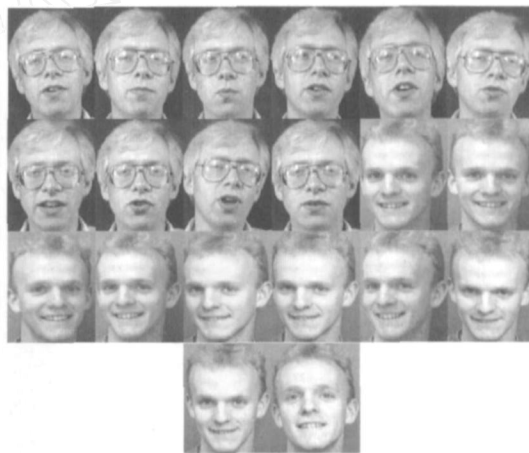


图 1 ORL 人脸数据库中的部分人脸图像

Fig. 1 Some face images from ORL face database

在实验中, 所有的训练样本集和测试样本集均随机产生. 每一次的实验中, 均从样本库的所有人脸数据中随机地从每类中抽取  $n = \{4, 5, 6\}$  个人脸组成训练样本集, 余下的人脸数据构成测试样本集, 不同的训练样本数目下均做 10 次随机的数据选择. 在经过特征提取之后, 最后的分类器采用最近邻分类器来分类, 在实验中, 分别以欧氏距离和余弦距离作为距离的度量值.

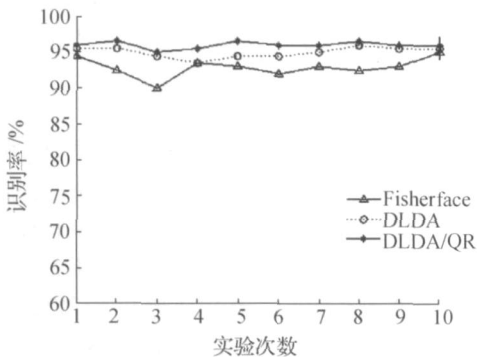
表 1 和表 2 分别显示了当采用不同距离度量标准的时候, DLDA/QR 算法、DLDA 算法和 Fisherface 在不同情形下的识别率比较. 从表中的结果可以看到, 在不同的距离度量下, DLDA/QR 算法总是优于传统的 DLDA 算法和 Fisherface 算法. 尤其是在经过 QR 分解之后, DLDA/QR 算法的第 2 步是在一个低维空间内实现. 此时相应的散布矩阵由于受原始类间散布矩阵秩的限制, 因此最大的可能值为  $C-1$ , 有效地避免了处理高维矩阵, 并且提高了算法的稳定性. 实验结果也验证了理论分析的正确性.

表 1 利用欧氏距离度量在 ORL 人脸数据库上的识别率

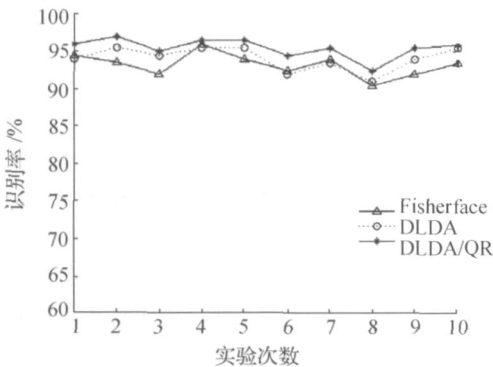
Table 1 Recognition rates on the ORL face database with Euclidean distance %			
识别方法	不同训练样本本数的均值和方差		
	4	5	6
Fisherface	89.85	92.90	93.13
	±1.88	±1.37	±1.02
DLDA	92.58	95.00	95.10
	±1.33	±0.75	±0.78
DLDA/QR	92.71	96.00	96.56
	±1.26	±0.47	±1.11

表 2 利用余弦距离度量在 ORL 人脸数据库上的识别率

Table 1 Recognition rates on the ORL face database with cosine distance %			
识别方法	不同训练样本本数的均值和方差		
	4	5	6
Fisherface	91.62	93.25	93.31
	±0.64	±1.55	±1.56
DLDA	92.6	94.10	94.75
	±21.02	±1.58	±1.54
DLDA/QR	93.05	95.50	96.31
	±0.83	±1.29	±0.62



(a) 欧氏距离下识别率

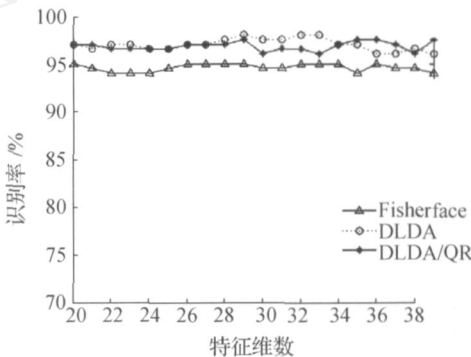


(b) 余弦距离下识别率

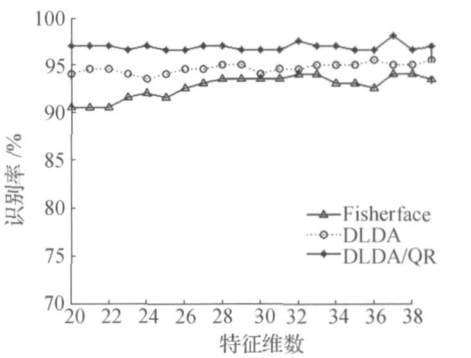
图 2 不同训练样本集下的识别率比较

Fig. 2 Comparison of recognition rates with different sample set

图 2 显示了当固定类别数 (40)、最终特征维数 (39) 和从每类中抽取的训练样本个数 (5) 的情况下, 在 10 次不同的训练集下的识别率比较. 从相应的结果中可以看出, DLDA/QR 算法有相对的稳定性. 图 3 描述了当固定类别数 (40) 和从每类中抽取的训练样本个数 (5) 的情况下, 最终抽取得到的特征维数从 20 变化到 39, 在不同特征维数下的识别率比较. 从以上的实验结果可以得出新算法用矩阵的 QR 分解方法取代了传统的降维预处理过程, 但是这一过程并没有降低算法的识别性能, 并且在不同的实验环境下均表现出良好的性能.



(a) 欧氏距离下识别率



(b) 余弦距离下识别率

图 3 不同特征数目下的识别率比较.

Fig. 3 Comparison of recognition rates with different features numbers

4 结束语

维数压缩很重要的一个目的是为了实现样本分类, 利用 Fisher 鉴别准则能在维数压缩过程之中融入样本的鉴别信息, 但是小样本问题是在利用 Fisher 鉴别准则时经常会遇到的问题, 直接线性鉴别分析方法是解决此类问题的一个有效鉴别维数压缩方法. 本文在直接线性鉴别分析方法中引入了矩阵 QR 分解的思想, 为高维、小样本的有效鉴别信息提取提供了新的思路和新的理论框架. QR 分解的引入, 使得无需处理一个高维的原始样本矩阵. 通过分

析矩阵的 QR 分解过程,可以在一个相对低维的空间中实现目标函数的优化.在第一步实现矩阵的目标函数优化之后,可以在一个较小的空间中实现特征提取过程.此时在新的空间之中,最佳鉴别矢量的计算只需在一个最大为  $C-1$  ( $C$  为样本类别数) 维空间中计算,有效降低了计算复杂度和对硬件存储性能的要求.

## 参考文献:

- [1] FU KUNAGA K. Introduction to statistical pattern recognition[M]. New York: Academic Press, 1990.
- [2] DUDA R O, HART P E, STORK D G. Pattern classification[M]. 2nd ed. New York: John Wiley & Sons, 2001.
- [3] BEL HUMEUR P N, HESPANHA J P, KRIEGMAN D J. Eigenfaces vs fisherfaces: recognition using class specific linear projection[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1997, 19(7): 711 - 720.
- [4] COMON P. Independent component analysis—a new concept[J]. Signal Processing, 1994, 36(3): 287 - 314.
- [5] BARTLETT M S, MOVELLAN J R, SEJNOWSKI T J. Face recognition by independent component analysis[J]. IEEE Transaction on Neural Network, 2002, 13(6): 1450 - 1464.
- [6] HYVARINEN A. Fast and robust fixed-point algorithm for independent component analysis[J]. IEEE Transaction on Neural Network, 1999, 10(3): 626 - 634.
- [7] YANG Jian, ZHANG David, YANG Jingyu. Is ICA significantly better than PCA for face recognition[A]. IC-CV 2005[C]. [S.l.], 2005.
- [8] SCHOL KOPF B, SMOLA A, MULLER K R, et al. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation. 1998, 10(5): 1299 - 1319.
- [9] SCHOL KOPF B, MIKA S, BURGESS C. Input space vs feature space in kernel-based methods[J]. IEEE Trans on Neural Networks, 1999, 10(5): 1000 - 1017.
- [10] SMOLA A J, MIKA S, SCHOL KOPF B, et al. Regularized principal manifolds [J]. Journal of Machine Learning Research, 2001, 1(3): 179 - 209.
- [11] STANFORD D C, RAGTER Y A E. Finding curvilinear features in spatial point patterns: principal curve clustering with noise[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(6): 601 - 609.
- [12] BENGIO Y, PAIEMENT J F, VINCENT P. Out of sample extensions for LLE, isomap, MDS, eigenmaps, and spectral clustering [A]. Neural Information Processing Systems 16 (NIPS 2003) [C]. [S.l.], 2003.
- [13] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290: 2323 - 2326.
- [14] HE X, YAN S, HU Y, et al. Face recognition using Laplacian faces[J]. IEEE Trans Pattern and Machine Intell, 2005, 27(3): 328 - 340.
- [15] YANG J, YU H, KUNZ W. An Efficient LDA algorithm for face recognition[A]. International Conference on Automation, Robotics, and Computer Vision (ICARCV2000) [C]. Singapore, 2000.
- [16] CHEN L F, LIAO Y M, LIN J C, et al. A new LDA-based face recognition system which can solve the small sample size problem[J]. Pattern Recognition, 2000, 33(10): 1713 - 1726.
- [17] FISHER R A. The use of multiple measurements in taxonomic problems[J]. Annals of Eugenics, 1936(7): 178 - 188.
- [18] WILKS S S. Mathematical statistics[M]. New York: Wiley, 1962.
- [19] DUDA R, HART P. Pattern classification and scene analysis[M]. New York: Wiley, 1973.
- [20] YE J P, LI Q. A two-stage linear discriminant analysis via QR-decomposition [J]. IEEE Trans Pattern Anal Machine Intell, 2005, 2(6): 929 - 941.
- [21] GOLUB G H, VAN LOAN C F. Matrix computations [M]. 3rd ed. Baltimore: The Johns Hopkins Univ Press, 1996.

## 作者简介:



杨静宇,男,1941年生,教授,博士生导师,主要研究方向为模式识别理论与应用、智能机器人、图像分析等.1982~1984在美国伊理诺斯大学(UIUC)CSL实验室进行合作研究,1993~1994在美国密苏里大学(UMKC)计算机系担任访问教授,1998年在加拿大康科迪亚大学(Concordia University)模式分析与机器智能中心担任访问教授.发表论文300余篇,论(译)著7部.

E-mail: yangjy@mail.njust.edu.cn.



郑宇杰,男,1977年生,博士,主要研究方向为模式识别、人工智能、机器学习等,被SCI、EI检索论文多篇.

E-mail: yjzheng13@yahoo.com.cn.