第 12 卷 第 12 期
2007 年 12 月

中国图象图形学报
Journal of Image and Graphics

Vol. 12, No. 12
Dec., 2007

Fisher 大间距线性分类器

陈才扣[1),2)] 杨静宇[1)]

[1)](南京理工大学计算机科学与技术学院, 南京 210094) [2)](扬州大学信息工程学院, 扬州 225009)

摘 要 作为一种著名的特征抽取方法,Fisher 线性鉴别分析的基本思想是选择使得 Fisher 准则函数达到最大值的向量(称为最优鉴别向量)作为最优投影方向,以便使得高维输入空间中的模式样本在该向量投影后,在类间散度达到最大的同时,类内散度最小。大间距线性分类器是寻找一个最优投影矢量(最优分隔超平面的法向量),它可使得投影后的两类样本之间的分类间距(Margin)最大。为了获得更佳的识别效果,结合 Fisher 线性鉴别分析和大间距分类器的优点,提出了一种新的线性投影分类算法——Fisher 大间距线性分类器。该分类器的主要思想就是寻找最优投影矢量 w^{best}(最优超平面的法向量),使得高维输入空间中的样本模式在 w^{best} 上投影后,在使类间间距达到最大的同时,使类内离散度尽可能地小。并从理论上讨论了与其他线性分类器的联系。在 ORL 人脸库和 FERET 人脸数据库上的实验结果表明,该线性投影分类算法的识别率优于其他分类器。

关键词 大间距分类器 支持向量机 Fisher 线性鉴别分析 人脸识别

中图法分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2007)12-2143-05

Fisher Large Margin Linear Classifier

CHEN Cai-kou[1),2)], YANG Jing-yu[1)]

[1)](*College of Computer Science and Technology*, *Nanjing University of Science and Technology*, *Nanjing* 210094)

[2)](*Information Engineering College*, *Yangzhou University*, *Yangzhou* 225009)

Abstract Fisher linear discriminant analysis(LDA), a well-known feature extraction method, searches for the projection axes on which the data samples from different classes are far from each other while requiring data samples of the same class to be close to each other. Large margin classifier(LMC), also referred as linear support vector machine, de finds a project direction onto which two classes of the samples projected reach maximal margin. With combination of advantages of both LDA and LMC, the paper develops a novel linear projection classfication algorithm, called Fisher large margin linear classifier. The underlying idea is that an optimal discrimiant vector w^{best} is found along which the samples of high dimensional input space are projected such that the margin is maximized while within-class scatter is kept as small as possible. In addition, relations to other classifiers are explored in theory in this paper. Finally, the proposed method is tested on ORL face database and FERET face database. The experimental results show that the proposed classifier outperforms other linear classifiers.

Keywords large margin classifier (LMC), support vector machines (SVM), Fisher linear discriminant analysis, face recognition

1 引 言

维度缩减是模式识别中最基本的问题之一。Fisher 线性鉴别分析(linear discriminant analysis, LDA)[1]是维度缩减的最为经典和广泛使用的方法。它的主要思想是选择一个使得与分类相关的准则——Fisher 准则函数达到最大值的矢量作为最优

基金项目:国家自然科学基金项目(60472060, 60632050);江苏省高校自然科学基金项目(05KJB520152);江苏省博士后科研资助计划项目(苏人通[2005]249)

收稿日期:2005-12-14;**改回日期**:2007-02-09

第一作者简介:陈才扣(1967~),男。博士后,副教授,硕士生导师。主要研究方向为模式识别理论与应用、生物特征识别。E-mail: yzcck@126.com

鉴别矢量(投影轴)。其物理意义是,高维输入空间中的模式样本在该鉴别矢量上投影后,可使同一类的模式样本相互集中,不同类的样本相互分离,类间散布程度和类内散布程度之比达到最大。但基于 Fisher 鉴别准则的维度缩减方法面临以下 3 个主要问题:一是类间散布矩阵常常为奇异矩阵;二是模式样本要求假设服从一定的正态分布,否则得到的是次优解;三是不能根据投影结果直接进行分类,因此最后通常需要利用最近邻或最小距离分类器进行分类。

最近,大间距分类器(large margin classifier, LMC)[2-5]正成为机器学习和模式识别领域的一个研究热点。作为大间距分类器中的最著名的算法——支持向量机(support vector machines, SVM),以其良好的性能正受到人们越来越多的关注。大间距分类器的基本思想是寻找一个最优投影矢量(最优分隔超平面的法向量),使得投影后的两类样本之间分类间距最大,它是根据 Vapnik 提出的结构风险最小化(structure risk minimization)原理,其目的是尽量提高学习机的推广能力。但现有的大间距分类器在强调分类间距(margin)最大的同时,没有像 Fisher 鉴别分析那样,同时考虑类内散度尽可能小的问题。为此,结合上述两类算法的优点,本文提出一种新的大间距分类器算法——Fisher 大间距分类器(FLMC),该算法的一个突出优点是借助于现有的支持向量机算法可直接进行计算,而不需要设计新的求解算法。在 ORL 和 FERET 人脸数据库上的测试结果表明,本文的算法都优于现有的大间距分类器和 Fisherfaces 算法。

2 Fisher 大间距分类器

2.1 大间距分类器

设线性可分样本集为$(\boldsymbol{x}_i, y_i), i=1,\cdots,m, \boldsymbol{x}_i \in \mathbf{R}^n, y_i \in \{+1,-1\}$是类别标号,类别为“+1”和“-1”的训练样本的个数分别为 m_1 和 m_2,并且 $m_1 + m_2 = m$。Vapnik 指出,具有最大间隔的分类超平面能够满足结构风险最小化原理。求解最大间隔的最优分类超平面的问题可以表示为如下的约束优化问题,即

$$\begin{aligned} &\min \quad J(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2 = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} \\ &\text{s.t.} \quad y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b) - 1 \geqslant 0, i = 1,2,\cdots,m \end{aligned} \tag{1}$$

定义如下的 Lagrange 函数:

$$L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} - \sum_{i=1}^{m}\alpha_i[y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b) - 1] \tag{2}$$

对上式分别对 $\boldsymbol{w}$ 和 b 求偏导数,并令它们等于 0,就可把原问题转化为如下的对偶问题:

$$\begin{aligned} &\max \quad Q(\boldsymbol{\alpha}) = \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j(\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \\ &\text{s.t.} \quad \sum_{i=1}^{m} y_i\alpha_i = 0, \alpha_i \geqslant 0, i = 1,\cdots,m \end{aligned} \tag{3}$$

若 α_i^{best} 为最优解,则最优分类器为

$$\begin{aligned} f(\boldsymbol{x}) &= \operatorname{sgn}\{(\boldsymbol{w}^{\text{best}} \cdot \boldsymbol{x}) + b^{\text{best}}\} \\ &= \operatorname{sgn}\left\{\sum_{i=1}^{m}\alpha_i^{\text{best}} y_i(\boldsymbol{x}_i \cdot \boldsymbol{x}) + b^{\text{best}}\right\} \end{aligned} \tag{4}$$

对于非线性可分的情况,大间距分类器可表示为如下的约束优化问题:

$$\begin{aligned} &\min \quad J(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{m}\xi_i \\ &\text{s.t.} \quad y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b) - 1 + \xi_i \geqslant 0, i = 1,2,\cdots,m \end{aligned} \tag{5}$$

其中,C 为某个指定的常数,其用于控制错分样本的比例与算法复杂度之间的折衷。

2.2 Fisher 大间距分类器

根据 Fisher 鉴别准则函数的思想,本文对现有的支持向量机的目标函数进行了如下的修正,使得在最优解向量 $\boldsymbol{w}^{\text{best}}$(最优超平面的法向量)上投影后,在保持最大的类间间距的同时,使得类内散度尽可能地小。

可将式(1)修正为如下的形式:

$$\begin{aligned} &\min \quad J_{\mathrm{M}}(\boldsymbol{w}) = \frac{1}{2}(\|\boldsymbol{w}\|^2 + \eta\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{w}) \\ &\text{s.t.} \quad y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b) \geqslant 1 \quad \forall i \end{aligned} \tag{6}$$

其中,$\boldsymbol{S}_{\mathrm{w}} = \sum_{i=1}^{2}\sum_{j=1}^{N_i}(\boldsymbol{x}_j - \boldsymbol{m}_i)(\boldsymbol{x}_j - \boldsymbol{m}_i)^{\mathrm{T}}$ 表示类内散布矩阵,$\boldsymbol{m}_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\boldsymbol{x}_j$ 表示模式样本均值。参数 η 为一正实数,用来平衡最大化类间间距和最小化类内散度两个不同的目标。η 值越大,意味着最小化类内散度越重要(相对于最大化类间散度而言),而当 $\eta=0$ 时,式(6)就是普通的最大间距分类器。该 FLMC(Fisher large margin classifier)分类器选择使得 $J_{\mathrm{M}}(\boldsymbol{w})$ 达到最大值的方向作为最优投影方向。显然,FLMC 分类器所确定的最优投影方向是以另

外一种方式，在使得投影后的类间散度达到最大的同时，类内散度达到最小。

式(5)是一个凸二次优化问题，为了便于计算，可将式(6)等价变换为

$$\begin{aligned}&\min \quad \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}(\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}})\boldsymbol{w}\\&\text{s.t.} \quad y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b)\geqslant 1 \quad \forall i\end{aligned} \tag{7}$$

其中，$\mathbf{I}$ 表示单位矩阵。

由于 $\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}$ 是一个对称矩阵，因此必存在正交矩阵 $\boldsymbol{P}=(\boldsymbol{p}_1,\cdots,\boldsymbol{p}_n)$，使得

$$\boldsymbol{P}^{-1}(\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}})\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}(\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}})\boldsymbol{P}=\boldsymbol{\Lambda} \tag{8}$$

其中，$\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$；$\boldsymbol{p}_1,\cdots,\boldsymbol{p}_n$ 为 $\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}$ 的标准正交的特征向量，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为所对应的特征值，且满足 $\lambda_j>0$，$j=1,\cdots,n$。

$\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}$ 也可写成

$$\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}=\boldsymbol{P}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}\boldsymbol{\Lambda}^{1/2}(\boldsymbol{P}\boldsymbol{\Lambda}^{1/2})^{\mathrm{T}} \tag{9}$$

将式(9)代入式(7)，则有

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}=\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{\Lambda}^{1/2})(\boldsymbol{P}\boldsymbol{\Lambda}^{1/2})^{\mathrm{T}}\boldsymbol{w}=\|\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}\|^2 \tag{10}$$

令 $\boldsymbol{w}_2=\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}$，则式(7)可改写成

$$\begin{aligned}&\min \quad \frac{1}{2}\|\boldsymbol{w}_2\|^2\\&\text{s.t.} \quad y_i(\boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{v}_i+b)\geqslant 1 \quad \forall i\end{aligned} \tag{11}$$

其中，$\boldsymbol{v}_i=(\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}})\boldsymbol{x}_i$。

因此，求解式(11)的最优解可以使用传统的 SVM 算法，下面给出计算 FLMC 的步骤：

(1) 对 $\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}$ 进行特征分解，即 $\mathbf{I}+\eta \boldsymbol{S}_{\mathrm{w}}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{T}}$；

(2) 把所有的训练样本 $\boldsymbol{x}_i$ 变换为 $\boldsymbol{v}_i$，$\boldsymbol{v}_i=(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{P}^{\mathrm{T}})\boldsymbol{x}_i$；

(3) 利用现有的支持向量机算法求解式(10)的二次规划问题来得到解 $\boldsymbol{w}_2$ 和 b；

(4) 计算式(6)的解：$\boldsymbol{w}=(\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{w}_2$，$b=b$。

由于非线性大间距分类器的解与线性大间距分类器几乎完全相同，因此，对于非线性可分问题的 FLMC 也与线性可分情况的 FLMC 几乎完全相同。

3 FLMC 分类器与其他分类器的关系

本节将讨论大间距 Fisher 分类器与传统的支持向量机以及 Fisher 线性鉴别分析之间的关系。

显然，根据式(6)确定的最优投影方向 $\boldsymbol{w}^{\mathrm{best}}$ 与参数 η 的取值有关。下面考察当参数 η 趋向无穷大时以及参数 η 取值为 0 时的两种特殊情况。

3.1 参数 η 趋向无穷大时的 FLMC 分类器

定理 1 若 $\boldsymbol{S}_{\mathrm{w}}$ 为奇异矩阵，则当参数 η 趋向无穷大时，由式(6)确定的最优投影方向等价于由式(12)确定的最优投影方向。

$$\begin{aligned}&\min \quad J_{\mathrm{M}}(\boldsymbol{w})=\frac{1}{2}\|\boldsymbol{w}\|^2\\&\text{s.t.} \quad y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b)\geqslant 1 \quad \forall i\\&\qquad \boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{w}=0\end{aligned} \tag{12}$$

证明 设 λ 为矩阵 $(\mathbf{I}+\eta\cdot \boldsymbol{S}_{\mathrm{w}})$ 的某个特征值，$\boldsymbol{p}$ 为对应的单位特征向量。因 $\boldsymbol{S}_{\mathrm{w}}$ 为奇异矩阵，故存在单位向量 $\boldsymbol{p}_0$，使得 $\boldsymbol{S}_{\mathrm{w}}\boldsymbol{p}_0=\mathbf{0}$。

对于任意的正实数 η，由 λ 和 $\boldsymbol{p}$ 的含义知，

$$(\mathbf{I}+\eta\cdot \boldsymbol{S}_{\mathrm{w}})\boldsymbol{p}=\lambda \boldsymbol{p} \tag{13}$$

可把式(13)改写为

$$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{p}=\frac{1}{\eta}(\lambda-\boldsymbol{p}^{\mathrm{T}}\mathbf{I}\boldsymbol{p})\leqslant \frac{1}{\eta}\lambda \tag{14}$$

另外，由于 $\boldsymbol{S}_{\mathrm{w}}$ 为半正定矩阵，因而有

$$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{p}\geqslant 0 \tag{15}$$

综合式(14)与式(15)不难得出

$$\lim_{\eta\to\infty}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{p}=0 \tag{16}$$

因此，当 η 趋向于无穷大时，式(6)的解 $\boldsymbol{w}$ 一定位于 $\boldsymbol{S}_{\mathrm{w}}$ 的零空间内。故式(12)确定的最优投影轴即为式(6)确定的最优投影轴的极限情况。

定理 2 由式(12)确定的最优投影方向等价于由式(17)确定的最优投影方向。或者说，若 $\boldsymbol{S}_{\mathrm{w}}$ 为奇异矩阵，则当参数 η 趋向无穷大时，由式(6)确定的最优投影方向等价于由式(17)确定的最优投影方向

$$\begin{aligned}&\max \quad \boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{b}}\boldsymbol{w}\\&\text{s.t.} \quad \begin{cases}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{w}=0\\ \|\boldsymbol{w}\|=1\end{cases}\end{aligned} \tag{17}$$

证明 由类内散布矩阵的含义不难知道，当 $\boldsymbol{S}_{\mathrm{w}}$ 为奇异矩阵时，训练样本是完全线性可分的。因 $\boldsymbol{S}_{\mathrm{w}}$ 为奇异矩阵，故存在单位向量 $\boldsymbol{p}$，使得 $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{p}=0$，即

$$\begin{aligned}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{p}&=\boldsymbol{p}^{\mathrm{T}}\sum_{\boldsymbol{x}\in\omega_i}(\boldsymbol{x}-\boldsymbol{m}_i)(\boldsymbol{x}-\boldsymbol{m}_i)^{\mathrm{T}}\boldsymbol{p}\\&=\sum_{\boldsymbol{x}\in\omega_i}(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{p}^{\mathrm{T}}\boldsymbol{m}_i)^2=0,i=1,2\end{aligned} \tag{18}$$

即

$$\forall \boldsymbol{x}\in\omega_i,\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{p}^{\mathrm{T}}\boldsymbol{m}_i,i=1,2 \tag{19}$$

也就是说，在 $\boldsymbol{S}_{\mathrm{w}}$ 的零空间内总可以找到一个投影方向，使得样本沿该方向投影后，同类样本对应着

投影轴上的同一个点,异类样本对应着投影轴上不同的点。此时,投影后两类样本之间的类间间距 $\frac{2}{\|\boldsymbol{w}\|^2}$ 就等于$[\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{m}_1-\boldsymbol{m}_2)]^2$。

而

$$\begin{aligned}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{b}}\boldsymbol{w} &= \boldsymbol{w}^{\mathrm{T}}(\boldsymbol{m}_1-\boldsymbol{m}_2)(\boldsymbol{m}_1-\boldsymbol{m}_2)^{\mathrm{T}}\boldsymbol{w}\\ &=(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_1-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_2)^2 \end{aligned}\tag{20}$$

因此,最小化$\frac{1}{2}\|\boldsymbol{w}\|^2$ 就等价于最大化 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{b}}\boldsymbol{w}$。

3.2 参数 η 取值为 0 时的 FLMC 分类器

当参数 η 取值为 0 时,FLMC 分类器就退化为式(1)的传统大间距分类器。由此可见,传统的支持向量机可以看成 FLMC 分类器的一种特殊形式,或者说,FLMC 是传统支持向量机的推广。

4 实验结果与分析

为验证本文分类器的识别效果,采用 ORL 和 FERET[6] 人脸图像数据库,对不同参数值下的 FLMC 分类器、大间距分类器(LMC,即线性 SVMs)、标准 SVMs 以及 Fisher 鉴别分析(Fisherfaces)的分类效果进行了实验比较。

ORL 人脸图像数据库由 40 人、每人 10 幅 92 × 112 大小的图像所组成。实验时,取每个人脸的前 5 幅图像作为训练样本,后 5 幅图像作为测试样本,因而训练样本和测试样本数均为 200。FERET 人脸图像库由 100 个人,每人 7 幅分辨率为 80 × 80pixels 的灰度图像组成。对于 ORL 人脸图像库,选取每个人的前 5 幅图像作为训练样本,后 5 幅图像作为测试样本,因而训练样本和测试样本数均为 480。在 FERET 人脸图像库上,实验一共进行 10 次,每次随机选取每个人的 3 幅图像作为训练样本,其余 4 幅图像作为测试样本,训练样本和测试样本数分别为 300 和 400。

FLMC 分类器、大间距分类器和标准 SVMs 本质上都是两类分类器,在用于多类分类问题时,需要将原问题分解为若干个两类分类问题。常见的分解方式有以下 3 种:一对一(one-vs-one)、一对多(one-vs-rest)和有向树(directed acyclic graph, DAG)[7]。本实验使用一对一分解方式。除了 Fisherfaces 方法外,其他分类器都是直接基于原始图像进行分类的。

表 1 列出了大间距分类器、标准 SVMs、Fisherfaces 以及不同参数值下的 FLMC 分类器在 ORL 数据集上的正确识别率,其中,Fisherfaces 的识别结果引自文献[6]中的表 1,其他识别结果都是在"一对一"分解方式下获得的。此外,大间距分类器中的参数 C 的取值为 1,标准 SVMs 分类器中是使用高斯核函数,其参数 $\sigma=1.5\times10^{-8}$。

表 1 各个分类器在不同分辨率 ORL 图像下的最优识别率比较

Tab. 1 Optimal recognition rates under different resolutions using several classifiers

分类器		各个分类器在不同分辨率下的识别率(%)				
		112 × 92	56 × 46	28 × 23	14 × 12	7 × 6
大间距分类器		95.5	95.5	95.5	94.5	82.5
标准 SVMs		97.0	97.0	97.0	96.0	85.5
Fisherfaces		86.0	85.5	84.5	82.0	82.0
FLMC 分类器	$\eta=0.05$	96.5	96.5	96.0	95.5	85.0
	$\eta=0.10$	96.5	96.5	96.5	96.0	85.0
	$\eta=1.00$	95.5	95.5	95.5	95.0	83.0
	$\eta=10.0$	95.0	95.0	95.0	94.0	82.0
	$\eta=100$	94.0	94.0	94.0	94.0	81.0

从表 1 中参数 η 在各种取值下的 FLMC 分类器的识别结果不难看出,若 η 取较小的值,则 FLMC 分类器就获得较高的识别结果,当 $\eta=0.1$ 时,识别结果最优,并且优于 Fisher 线性鉴别分析(Fisherfaces)和传统的大间距分类器(线性支持向量机);同时,随着 η 取值的增大,识别性能逐渐下降,其识别精度反而不如传统的大间距分类器。这个实验结果验证了参数 η 在 FLMC 分类器中的作用。但是 FLMC 分类器的识别性能不如标准 SVMs,其原因是由于人脸图像的特征分布通常是非线性的,本文方法以及大间距分类器只能获得线性分类面,而标准 SVMs 则采用了非线性的高斯核函数,因而其最终得到的分类面是非线性的,这正是非线性方法的优势所在。此外,随着图像分辨率的降低(112 × 92、56 × 46、28 × 23),FLMC 的识别性能仅略有下降,但在7 × 6大小的马赛克图像上的正确识别率却突然下降到 85.0%,大间距分类器在同样的图像上的识别率也只有 82.5%。这一方面说明了 FLMC 分类器和大间距分类器对较低分辨率的图像不具有鲁棒性,另一方面表明 FLMC 本质上仍然是大间距分类器(当 η 取很小的值),因此,它俩的实验结果的变化规律具有很大的相似性也就不足为奇。

表 2 列出了 Fisherfaces、大间距分类器、标准 SVMs 及 FLMC 分类器在 FERET 数据集上执行 10 次鉴别分类后的正确识别率。其中,Fisherfaces 方法中采用最近邻分类来对抽取后的特征进行分类。线性 SVM 参数 C 的取值为 1,标准 SVMs 分类器中是使用高斯核函数,其参数 $\sigma = 1.5 \times 10^{-6}$。

表 2　Fisherfaces、LMC、SVMs 以及 FLMC 在 FERET 数据集上的识别率比较

Tab. 2　Recognition rates in the FERET face databse using Fisherfaces、LMC、SVMs and FLMC

实验编号	不同分类器的正确识别率(%)			
	Fisherfaces	LMC	SVMs	FLMC
1	56.10	61.00	62.60	62.20
2	64.75	70.75	74.50	73.75
3	65.25	73.25	77.00	75.65
4	55.50	60.50	63.40	62.85
5	63.25	68.25	69.90	68.55
6	53.90	59.80	62.90	61.10
7	65.85	73.25	75.10	73.95
8	63.65	69.10	73.15	71.25
9	59.35	65.85	67.95	66.30
10	65.00	71.70	75.40	72.45
平均识别率(%)	61.26	67.35	70.19	68.81

表 2 中的数据表明,FLMC 的正确识别率均分别优于 Fisherfaces 和 LMC,但低于 SVMs,这进一步验证了本文方法的有效性。

5　结　论

本文提出的 Fisher 大间距分类器不仅集中了 Fisher 鉴别分析和大间距分类器的优点,而且具有更强的物理意义。在 ORL 和 FERET 人脸数据库上的实验结果表明,FLMC 分类器在识别性能优于大间距分类器,也明显高于 Fisher 线性鉴别分析。由于其固有的线性特性,致使其识别性能略低于非线性支持向量机。

参考文献(References)

1　Fisher R A. The use of multiple measurements in taxonomic problems [J]. Annals of Eugenics, 1936, **7**(part Ⅱ): 179 ~ 188.

2　Vapnik V. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995.

3　Freund Yoav, Schapire R E. Large margin classification using the perceptron algorithm[J]. Machine Learning, 1999, **37**(3): 277 ~ 296.

4　Huang Kai-zhu, Yang Hai-qin, King Irwin. Learning large margin classifiers locally and globally[A]. In: Proceedings of the Twenty-first International Conference on Machine Learning [C], Banff, Alberta, Canada, 2004, **69**: 268 ~ 275.

5　Bian Zhao-qi, Zhang Xue-gong. Pattern Recognition[M]. Beijing: Tsinghua University Press, 2000. [边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社,2000.]

6　Phillips P J, Moon H, Rizvi S A, *et al.* The FERET evaluation methodology for face recognition algorithms[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, **22**(10): 1090 ~ 1104.

7　Hsu C, Lin C. A comparison of methods for multiclass support vector machines [J]. IEEE Transactions on Neural Networks, 2002, **13**(2): 415 ~ 425.