

文章编号: 1001-0920(2007)04-0384-05

基于非线性映射的 Fisher 判别分析

范玉刚, 李平, 宋执环

(浙江大学 a. 工业控制技术国家重点实验室, b. 工业控制技术研究所, 杭州 310027)

摘要: 提出一种非线性分类方法——基于非线性映射的 Fisher 判别分析(NM-FDA). 首先提取基向量; 然后采用 Nystro'm 方法, 以基向量为训练样本, 将形式未知的非线性映射近似表达为已知形式的非线性映射, 这种近似的非线性映射将变量由非线性的输入空间转换到线性的特征子空间; 最后对映射数据进行线性 Fisher 判别分析. 实验采用 7 组标准数据集, 结果显示 NM-FDA 具有较强的分类能力.

关键词: Fisher 判别分析; 核函数; 特征空间

中图分类号: TP181 **文献标识码:** A

Fisher discriminant analysis based on nonlinear mapping

FAN Yu-gang, LI Ping, SONG Zhi-huan

(a. National Lab of Industrial Control Technology, b. Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, China. Correspondent: FAN Yu-gang, E-mail: ygf@iipc.zju.edu.cn)

Abstract: As a nonlinear classification method, Fisher discriminant analysis based on nonlinear mapping(NM-FDA) is presented. Base vectors are selected for formulating the nonlinear mapping. An approximate explicit expression is formed by using the Nystro'm method for nonlinear mapping which is not available originally. The data in the nonlinear input space is mapped into some linear feature subspace and fisher linear discriminant analysis is computed for the mapped data. Seven benchmark datasets are applied in the experiment, and the results show the good performance of NM-FDA.

Key words: Fisher discriminant analysis; Kernel function; Feature space

1 引言

线性 Fisher 判决分析算法(FLDA)是一种旨在降低特征维数的方法^[1,2]. FLDA 算法的目标是建立一个子空间(由所有的投影轴构成),所有样本在该子空间内满足类内散度最小,类间散度最大.样本在这些投影轴上的投影系数可作为样本的特征向量,利用这些特征向量,便可进行样本的分类识别. FLDA 在解决两类模式识别问题中取得了较好的结果,并得到广泛应用.然而作为统计数据分析常用的技术,FLDA 有一定的局限性,常常不能有效地将不同类别的样本区分开^[3].为此,本文提出基于非线性映射的 Fisher 判别分析(NM-FLDA),其基本思想是:首先提取基向量,并以基向量为训练样本,采用 Nystro'm 方法构造非线性映射函数;然后对映射数据进行线性 Fisher 判别分析.与 FLDA 相比,NM-FDA 对输入空间的样本数据进行非线性映射,提高

了分类性能.

文献[4,5]将核函数引入 FLDA,形成基于核函数的非线性特征提取方法——核 Fisher 判决分析算法(KFDA).KFDA 在具体求解过程中可能会出现问题:一是特征空间的维数大于或等于样本数据量 N ,导致特征空间中类内散布矩阵通常是奇异的;二是特征空间的散布矩阵的维数等于训练样本的数目,当样本量很大时,KFDA 所耗费的计算量异常惊人.

与 KFDA 相比,本文方法实现了非线性映射的具体形式,对特征子空间中的映射数据直接进行线性 Fisher 分析,因此其实现原理和过程不同.在本文方法中,选取的基向量只是训练样本的一部分,散布矩阵的维数小于样本数据量 N ,因此可以避免 KFDA 中存在的计算问题.

收稿日期: 2005-12-07; 修回日期: 2006-02-25.

基金项目: 国家 973 计划项目(2002CB312200); 浙江省科技计划项目(2004C31106).

作者简介: 范玉刚(1973—),男,山东文登人,博士生,从事数据挖掘、机器学习等研究;李平(1954—),男,广西北流人,教授,博士生导师,从事工业过程模型化、智能控制等研究.

2 FLDA 和 KFDA 的机理分析

2.1 Fisher 线性判别分析

线性 Fisher 判别分析是把输入空间的样本数据投影到一条直线上,使得在这条直线上样本的投影分开得最好.

假设一个样本集合 X 包含 N 个 d 维样本 $X = \{x_1, \dots, x_N\}$, 其中: N_1 个属于 C_1 类的样本记为 $X_1 = \{x_1^1, \dots, x_{N_1}^1\}$, N_2 个属于 C_2 类的样本记为 $X_2 = \{x_1^2, \dots, x_{N_2}^2\}$. 在进行线性 Fisher 判决分析时,目标是找到线性投影方向,以使训练样本在这些轴上的投影结果为:类内离散度最小,类间离散度最大.

定义类间离散度矩阵 S_b 和类内离散度矩阵 S_w 如下:

$$S_b = (m_1 - m_2)(m_1 - m_2)^T,$$
$$S_w = \sum_{j=1}^2 \sum_{x_j} (x - m_j)(x - m_j)^T.$$

其中: m_j 是各类样本的均值向量,即 $m_j = \frac{1}{n_j} \sum_{x_j} x_j, j = 1, 2$. 那么寻找最好投影方向的问题在数学上就成为寻找最好的变换方向 w^* 的问题,即最大化下式:

$$J(w) = (w^T S_b w) / (w^T S_w w).$$

因此最大化 $J(w)$ 的本质是要找到一个最好方向来最大化投影后的类间离散度(分子),同时最小化该方向上的类内离散度(分母). 设任意一个测试样本 x 到方向为 w^* 的投影为 $G(x) = w^{*T} x$, 那么其决策函数为 $F(x) = \text{sgn}(G(x) - b_0)$, 其中 b_0 为决策阈值.

2.2 核 Fisher 判别分析

线性 Fisher 判别分析是统计分析常用的技术,属于线性方法. 然而,在实际情况下,线性方法非常有限,常常不能有效地解决非线性问题. 为此 KFDA 引入非线性映射 ϕ , 将输入空间的样本数据 $x \in R^d$ 变换到特征空间 F 中,即 $\phi: R^d \rightarrow F, x \rightarrow \phi(x)$, 其中 $x = \phi(x), x \in F$. 这样在特征空间 F 中,可以进行线性 Fisher 判别分析. 此时线性 Fisher 判别分析函数为

$$J(w) = (w^T S_b^\phi w) / (w^T S_w^\phi w). \tag{1}$$

其中: $w \in F, S_b^\phi$ 和 S_w^ϕ 是 F 中相应的类间离散矩阵和类内离散矩阵,即

$$S_b^\phi = (m_1^\phi - m_2^\phi)(m_1^\phi - m_2^\phi)^T,$$
$$S_w^\phi = \sum_{j=1}^2 \sum_{i=1}^{n_j} (\phi(x_i^j) - m_j^\phi)(\phi(x_i^j) - m_j^\phi)^T,$$

其中

$$m_j^\phi = (1/n_j) \sum_{i=1}^{n_j} \phi(x_i^j), j = 1, 2.$$

KFDA 算法首先通过一个非线性映射,将输入数据影射到一个高维的核空间中;然后在该高维核空间中再进行线性 Fisher 判决分析. 然而在 KFDA 中,并没有确定非线性映射 $\phi(x)$ 的具体形式. 为了求解特征空间 F 中的线性 Fisher 判别分析, KFDA 引入核函数,将式(1)转化为只包含特征矢量内积的形式. 由于特征空间中各坐标分量间的计算仅限于内积,用满足 Mercer 条件的核函数代替线性算法中的内积,不需知道非线性映射的具体形式. 因此 KFDA 算法是将核学习方法的思想与 FLDA 算法相结合的产物.

3 基于非线性映射的 Fisher 判别分析

NM-FDA 主要包含 3 个过程:1) 基向量的选取;2) 采用 Nystro'm 方法构造非线性映射函数;3) 对映射数据进行线性 Fisher 判别分析.

3.1 基向量提取

假设样本集 X 有 N 个样本, $X = \{x_i / i = 1, 2, \dots, N\}$, 对应于样本集 X , 有如下定义:

定义 1 设样本集 X 中存在子集 $S = \{x_{s1}, \dots, x_{sL}\}, L \leq N, x_{si}$ 在特征空间 F 中的映射为 $\phi(x_{si})$, 若子空间 F_s 由 $\{\phi(x_{s1}), \phi(x_{s2}), \dots, \phi(x_{sL})\}$ 扩展而成,使得样本 x_i 的映射 $\phi(x_i)$ 在 F_s 中的投影误差小于给定值,则称 x_{si} 为基向量, S 为基向量集, F_s 为特征子空间.

如图 1 所示, S 在 F 中的映射为 $(\phi(x_{s1}), \dots, \phi(x_{sL}))$, 其扩展而成的特征子空间是图 1 的阴影区域,用 $F_s(\phi(x_{s1}), \dots, \phi(x_{sL}))$ 表示. 样本 x_i 的非线性映射 $\phi(x_i)$ 可分解为正交投影 $\hat{\phi}(x_i)$ 和正交投影误差 $\phi(x_i)_{res}$.

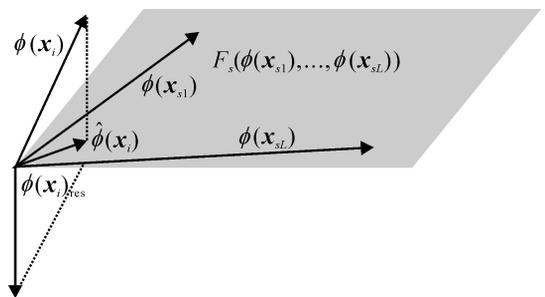


图 1 向量 $\phi(x_i)$ 在子空间 F_s 中的正交投影 $\hat{\phi}(x_i)$ 和正交误差 $\phi(x_i)_{res}$

3.1.1 基向量选取原理

从 X 中选取 L 个基向量,构成基向量集 $S = \{x_{s1}, \dots, x_{sL}\}$, 由 S 的映射扩展而成子空间 F_s , 那么其他样本在子空间 F_s 中的正交投影可以表示为 $\hat{\phi}_i = \sum_{s=1}^L a_i \cdot \phi_{si}$. 其中: $s = (\phi_{s1}, \dots, \phi_{sL}), a_i = (a_{i1}, \dots, a_{iL})^T, a_i$ 是使 $\hat{\phi}_i$ 和 ϕ_i 差异最小的系数向量. 通过计

算 ϕ_i 与 F_s 的夹角求取 a_i , 因此求解 a_i 可归结为

$$a_i = \min_{a_i} \frac{\phi_i - \phi_i^2}{\phi_i},$$

或

$$a_i^2 = \min_{a_i} \frac{\phi_i - \phi_i^2}{\phi_i^2}.$$

求解上式得

$$a_i^2 = 1 - \frac{K_{ii}^T K_{ss}^{-1} K_{si}}{k_{ii}},$$

其中

$K_{ss} = (k_{s_p s_q})_{1 \leq p, q \leq L}$, $k_{s_p s_q} = \Phi^T(x_{s_p}) \Phi(x_{s_q})$, x_{s_p} 和 x_{s_q} 是基向量, $K_{ii} = (k_{s_p i})_{1 \leq p \leq L}$ 是样本 x_i 和基向量集 S 的核函数内积^[5]. 因此所有样本对 F_s 的夹角的正弦值平方和为

$\sum_{i=1}^N \left(1 - \frac{K_{ii}^T K_{ss}^{-1} K_{si}}{k_{ii}}\right)$, 以该式值最小为目标函数选取基向量, 问题可以归结为

$$\min_S \sum_{i=1}^N \left(1 - \frac{K_{ii}^T K_{ss}^{-1} K_{si}}{k_{ii}}\right), \quad (2)$$

即

$$\max_S \sum_{i=1}^N \frac{K_{ii}^T K_{ss}^{-1} K_{si}}{k_{ii}}. \quad (3)$$

为方便表示, 定义

$$J_s = \frac{1}{N} \sum_{i=1}^N J_{si}, \quad (4)$$

其中

$$J_{si} = \frac{K_{ii}^T K_{ss}^{-1} K_{si}}{k_{ii}} = \frac{\phi_i^2}{\phi_i^2}, \quad (5)$$

那么式(2) 等于 $\max_S (J_s)$.

上式得到的基向量的映射 $\Phi(x_{si})$ 扩展生成特征子空间 F_s , F_s 刻画了样本集 X 在 F 中的映射的主要特征. 从 J_s 和 J_{si} 的定义可以看出, 其取值范围为 $(0, 1]$.

3.1.2 基向量选取算法

基向量选取分两步进行: 首先对训练样本集进行聚类, 离各个簇中心最近的样本构成初始基向量集 S ; 然后进行一个循环选取过程, 计算基向量集 S 的代表性, 即计算式(4) J_s 和式(5) J_{si} , 将最小 J_{si} 对应的样本添加到基向量集 S 中, 计算新的基向量集 S 的代表性 J_s , 此过程不断循环, 直到 J_s 达到要求. 下面是基向量选取算法的基本执行步骤:

Step1: 设定停止条件, 即最大代表性指标 $\max \text{Fitness}$ 和最大基向量的容量 M ;

Step2: 利用 FCM 聚类算法将数据样本集聚类成 k 簇, 离各个簇中心最近的 k 个样本构成初始基向量集 $S, L = k$;

Step3: 计算 J_{sj} 和 $J_s, 1 < j < N$;

Step4: 选取样本 $x_j^*, j^* = \arg \min_j J_{sj}$;

Step5: 如果 $L < M$ 并且 $J_s < \max \text{Fitness}$, 则 $L = L + 1, S = S \cup \{x_j^*\}$, 转向 Step3, 否则转向 Step6;

Step6: 退出, S 为选取的基向量集.

得到基向量集 S 后, 以 S 为训练样本, 构造非线性映射.

3.2 构造非线性映射

由 Mercer-Hilbert-Schmidt 定理, 对于任意核函数, 存在一个正交集 $\{\phi_i\}_{i=1}^d$ 和与其对应的非负正数 λ_i , 有

$$k(x, x) = \sum_{i=1}^d \lambda_i \phi_i(x) \phi_i(x) = \Phi(x)^T \Lambda \Phi(x), \quad (6)$$

其中 d 是空间 RKHS(再生核的希尔伯特空间) 的维数. 式(6) 是核函数的分解式, 其中 λ_i 和 ϕ_i 分别是积分等式

$$k(x, x) = \int \phi_i(x) p(x) dx = \lambda_i \phi_i(x) \quad (7)$$

的特征值和特征向量.

非线性映射 $\Phi(x)$ 可以表达为 $d \times 1$ 的向量, 即

$$\Phi(x) = [\sqrt{\lambda_1} \phi_1(x), \dots, \sqrt{\lambda_d} \phi_d(x)]^T. \quad (8)$$

然而 ϕ 的具体形式并没确定, 需要采用 Nyström 方法, 将形式未知的非线性映射近似表达为已知形式的非线性映射^[6].

用基向量集对积分等式(7) 离散化得

$$\frac{1}{L} \sum_{j=1}^L k(x, x_j) \phi_i(x_j) = \lambda_i^{(L)} \phi_i^{(L)}(x). \quad (9)$$

式(9) 可以重新组织为矩阵的特征值求解问题, 即

$$K U_L = U_L \Lambda_L,$$

其中: $K_{ij} = \phi_i(x_j) \phi_i(x_j) = k(x_i, x_j)$; $U_L = [u_1, \dots, u_L]$ $L \times L$ 是由 K 的特征向量组成; Λ_L 是 $L \times L$ 阶对角矩阵, 对角线上的元素是特征值 $\lambda_i^{(L)}$. 对于基向量 $(x_{s1}, x_{s2}, \dots, x_{sL})$, 由式(9) 得到的特征向量 $\phi_i(x_j)$ 和特征值 λ_i 可近似表达为

$$\phi_i(x_j) \approx \sqrt{\lambda_i^{(L)}} u_{ji},$$

$$\lambda_i \approx \frac{1}{L} \lambda_i^{(L)}.$$

对于其他样本有

$$\Phi(x) \approx \frac{\sqrt{\Lambda_L}}{\lambda_i^{(L)}} \sum_{j=1}^L k(x, x_j) u_{ji} = \frac{\sqrt{\Lambda_L}}{\lambda_i^{(L)}} [k(x_1, x), \dots, k(x_L, x)] U_L, \quad (10)$$

其中 u_{ji} 是矩阵 U_L 第 j 行第 i 列元素. 由式(10) 和(8), 样本 x 非线性映射 $\Phi(x)$ 可以近似表达为如下 L

$\times 1$ 的向量:

$$\Phi(x) = \left[\frac{1}{\sqrt{L}} \Lambda_1(x), \dots, \frac{1}{\sqrt{L}} \Lambda_L(x) \right]^T = \left(\frac{w}{\sqrt{L}}, \dots, \frac{w}{\sqrt{L}} \right)^T [k(x_1, x), \dots, k(x_L, x)]^T. \quad (11)$$

将所有样本都按式(11)进行计算,样本在特征空间的映射数据就可近似表达为

$$\begin{bmatrix} \Phi^T(x_1) \\ \Phi^T(x_2) \\ \dots \\ \Phi^T(x_N) \end{bmatrix} \Lambda = \begin{bmatrix} \Phi^T(x_1) \\ \Phi^T(x_2) \\ \dots \\ \Phi^T(x_N) \end{bmatrix} R^{N \times L}.$$

对矩阵 Λ 进行线性 Fisher 判别分析,完成基于非线性映射的 Fisher 判别分析.

3.3 NM-FDA 的实现过程

NM-FDA 训练过程总结如下:

- 1) 选取基向量;
- 2) 以基向量集为训练样本,采用 Nystro'm 方法构造非线性映射;
- 3) 以样本在特征子空间的映射数据为训练样本,进行线性 Fisher 判别分析,得到最好的变换方向 w^* .

对测试样本 x 进行类属判别时,首先将它进行非线性映射得到 $\Phi(x)$,判断类别的决策函数为 $F(x) = \text{sgn}(w^{*T} \Phi(x) - b_0)$,其中 b_0 为决策阈值.

在 KFDA 算法中,所有运算都是通过原空间中定义的核函数来进行的,并未涉及到非线性映射 $\Phi(x)$ 的具体形式. NM-FDA 与 KFDA 相反,采用 Nystro'm 方法构造非线性映射函数,将形式未知的非线性映射,近似表达为已知形式的非线性映射,从而可以直接采用线性 Fisher 方法计算出 S_b^ϕ 和 S_w^ϕ ,简化 KFDA 的计算.

4 实验例子

例 1 采用二维数据作为样本集^[7]. 样本分为两类,每类样本有 100 个样本,且分别满足 $N((-1, 0), \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix})$ 和 $N((1, 0), \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})$ 分布. 数据分布情况见图 2,两类数据分别用 '+' 和 '.' 标识,标识为 'o' 的数据是选取的基向量,本实例共选取 18 个基向量.

用线性 Fisher 判别分析的误分类率为 14.5%,用本文方法进行分类实验的误分类率为 8.5%. 两次实验分别用“0”和“1”标识样本类别,所以决策界面上的点计算值为 0.5,即决策阈值 $b_0 = 0.5$. 第 3 次实验用文献[8,9]提供的 GDA (KFDA) 进行分

析,误分类率为 8.5%. 为了避免混淆, GDA 进行分析时,用“-1”和“1”标识样本类别,所以决策界面上的点计算值为 0,即决策阈值 $b_0 = 0$. 从图 2 可看出, NM-FDA 的决策界面为曲线,表现出非线性分类器的特点,它与 GDA 的误分类率都为 8.5%,低于线性 Fisher 判别分析的误分类率. 由于基向量只训练样本的一小部分(18 个基向量只占全部样本的 9%),因此 NM-FDA 的训练时间和测试时间都小于 KFDA. 例如对 6 400 个样本进行测试时, NM-FDA 训练时间和测试时间分别为 0.320 0 s 和 0.881 0 s; KFDA 用时分别为 1.772 0 s 和 41.650 0 s. 实验采用径向基核函数,参数 $\gamma = 2$.

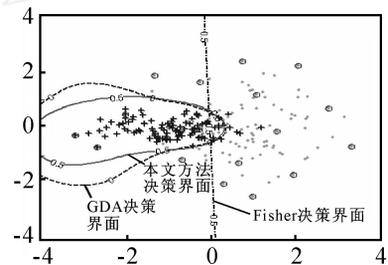


图 2 NM-FDA 和 ELDA 分类界面

例 2 将 NM-FDA 应用于 7 个数据集的分类问题,它们是标准数据集^[10],其基本信息见表 1,详细信息请参见文献[10]. 表 1 中样本量的括号内的数字分别表示第 1 类和第 2 类样本数.

表 1 数据集的基本情况

数据集	样本量	维数	类别
Ion	351(225, 126)	34	2
WDB	569(212, 357)	30	2
Wine	178(107 (1 & 2), 71)	12	2
Can	683(444, 239)	10	2
Iris	100(50, 50)	4	2
Diab	768(500, 268)	8	2
Liv	345(145, 200)	6	2

本文采用文献[11]的方法对 Can Wine 和 Iris 三个样本集进行预处理. Can 数据集含有 699 个样本,其中有 16 个样本数据不完整,实验时将它们删除. Wine 数据集含有 3 类数据,将第 1 类样本和第 2 类样本合并为一类样本,与第 3 类样本组成两类样本的分类问题. 从 Iris 样本集选取第 2 类样本 (versicolor) 和第 3 类样本 (virginica) 组成 Iris 样本集,这两类样本部分重叠,第 1 类样本 (setosa) 与其他差异明显,是完全可分的,所以将第 1 类样本省略.

每次实验都随机抽取数据,将数据平分为两部分:一部分用于训练样本;另一部分用于测试. 实验时对数据进行标准化处理,使得处理后的样本每一

维数据均值为 0, 方差为 1. 表 2 为分类结果, 其中 FLDA1 和 ModLDA 对应的数据直接引用文献[11] 中的结果, FLDA1 为线性 Fisher 判别分析, ModLDA 为本文提出的新方法, 对线性 Fisher 判别分析进行了改进. FLDA2 为本文进行的线性 Fisher 判别分析, 由于采样过程不完全相同, 因此与 FLDA1 的实验结果有所偏差. 由表 2 可以看出, 本文算法在其中的 4 个数据集中分类准确率最高, 表现出优异的分类性能.

表 2 FLDA, ModLDA 和 NM-FDA 分类性能比较

数据集	FLDA1	ModLDA	FLDA2	NM-FDA
Ion	83.86	85.42	80.57	92.57
WDB	95.30	95.81	94.37	95.77
Wine	92.08	92.30	97.75	100
Can	100.0	100.0	95.01	96.48
Iris	93.72	93.72	95.2	95.6
Diab	73.57	73.46	78.13	77.08
Liv	62.80	62.67	67.44	68.02

实验采用径向基核函数, 对应于数据集 Ion, WDB, Wine, Can, Iris, Diab 和 Liv, 所采用的参数分别为 7.1, 5, 3.2, 3.2, 3.2, 4.2 和 2.7.

5 结 语

本文首先提取基向量, 然后采用 Nyström 方法将形式未知的非线性映射近似表达为已知形式的非线性映射, 并将这种近似的非线性映射与线性 Fisher 判别分析相结合, 形成一种非线性分类方法 NM-FDA. 人工数据仿真实验显示, NM-FDA 的分类精度与 KFDA 基本相同, 采用 7 组实际问题的数据进行实验, 结果显示本文方法分类准确率高于其他方法, 在其中的 4 个数据集中分类准确率最高. 与 KFDA 相比, 本文方法思想简练直接, FLDA 的改进形式如 D-LDA^[12-14] 和 R-LDA^[15] 等可以直接对映射数据进行分析, 从而构成相应的非线性算法.

参考文献(References)

- [1] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces versus fisherfaces: Recognition using class specific linear projection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(5): 711-720.
- [2] Martinez A M, Kak A C. PCA versus LDA[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2001, 23(2): 228-233.
- [3] 甘俊英, 张有为. 模式识别中广义核函数 Fisher 最佳鉴别[J]. 模式识别与人工智能, 2002, 15(2): 429-434. (Gan J Y, Zhang Y W. Generalized kernel fisher optimal discriminant in pattern recognition[J]. Pattern Recognition and Artificial Intelligence, 2002, 15(2): 429-434)
- [4] Mika S, Ratsch G, Weston J, et al. Fisher discriminant analysis with kernels[C]. Proc of the 1999 IEEE Signal Processing Society Workshop. Madison, 1999: 41-48.
- [5] Schölkopf B, Mika S, Burges C J C, et al. Input space versus feature space in kernel-based methods[J]. IEEE Trans on Neural Networks, 1999, 10(3): 1000-1017.
- [6] Hoegaerts L, Suykens J A K, Vandewalle J, et al. Subset based least squares subspace regression in RKHS[J]. Neurocomputing, 2005, 63(1): 293-323.
- [7] Baudat G, Anouar F. Feature vector selection and projection using kernels[J]. Neurocomputing, 2003, 55(1-2): 21-38.
- [8] Alex Smola, Bernhard Schölkopf. Generalized discriminant analysis[DB/OL]. (2005-03-15). <http://www.kernel-machines.org/>.
- [9] Baudat G, anouar F. Generalized discriminant analysis using a kernel approach[J]. Neural Computation, 2000, 12(10): 2385-2404.
- [10] UCI machine learning repository[DB/OL]. (2005-03-15). <http://www.ics.uci.edu/~mllearn/MLRepository.html>.
- [11] Chen S C, Li D H. Modified linear discriminant analysis[J]. Pattern Recognition, 2005, 38(3): 441-443.
- [12] Chen L F, Liao H Y M, Ko M T, 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.
- [13] Yu H, Yang J. A direct LDA algorithm for highdimensional data — With application to face recognition[J]. Pattern Recognition, 2001, 34(10): 2067-2070.
- [14] Lu J W, Plataniotis K N, Venetsanopoulos A N. Face recognition using LDA based algorithms[J]. IEEE Trans on Neural Networks, 2003, 14(1): 195-200.
- [15] Lu J W, Plataniotis K N, Venetsanopoulos A N. Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition[J]. Pattern Recognition Letters, 2005, 26(2): 181-191.