

## 基于局部子域的最大间距判别分析

皋 军<sup>1,2,3</sup>, 黄丽莉<sup>2</sup>, 王士同<sup>4</sup>

(1. 盐城工学院 信息学院, 江苏 盐城; 224000; 2. 安徽理工大学 电气与信息工程学院, 安徽 淮南 232001; 3. 东南大学 自动化学学院, 南京 210096; 4. 江南大学 数字媒体学院, 江苏 无锡 214122)

**摘 要:** 线性判别分析(LDA)作为一种经典的特征提取方法被广泛地加以研究和运用,然而LDA作为全局判别准则在一定程度上忽视了样本空间的局部结构和局部信息.为此,通过引入局部加权均值(LWM)并结合最大间距判别分析(MMC)提出了具有一定局部学习能力的有监督的特征提取方法——基于局部加权均值的最大间距判别分析(LBMMC).算法结合了QR分解技术,提高了其执行效率,并通过在数据集上的测试结果表明了该算法的有效性.

**关键词:** 线性判别分析; 局部加权均值; QR分解; 最大间距判别分析

**中图分类号:** TP273

**文献标志码:** A

### Local sub-domains based maximum margin criterion

GAO Jun<sup>1,2,3</sup>, HUANG Li-li<sup>2</sup>, WANG Shi-tong<sup>4</sup>

(1. School of Information Engineering, Yancheng Institute of Technology, Yancheng 224000, China; 2. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China; 3. School of Automation, Southeast University, Nanjing 210096, China; 4. Digital Media Institute, Jiangnan University, Wuxi 214122, China. Correspondent: GAO Jun, E-mail: gj0104211@163.com)

**Abstract:** Linear discrimination analysis(LDA) as a classic feature extraction method is widely studied and used. However, LDA as a global criterion is neglected to some extent sample space inner local structure and local information. Therefore, when combined with local weighted mean(LWM) and maximum margin criterion(MMC), a supervised feature extraction method of local learning ability, known as local sub-domains based maximum margin criterion(LBMMC), is proposed. The method is also combined with the QR decomposition technique to improve the efficiency of the algorithm. Finally, the test on the datasets show the effectiveness of the proposed method.

**Key words:** linear discrimination analysis; local weighted mean; QR-decomposition; maximum margin criterion

## 0 引 言

在实际智能识别问题过程中,待处理识别数据一般具有明显的高维特征,故应在尽可能保持原始输入空间判别信息的基础上对其进行适当地降维<sup>[1]</sup>.

主成分分析(PCA)<sup>[2-3]</sup>和线性判别分析(LDA)<sup>[4-5]</sup>是较为经典的两种降维算法. PCA依据提取主成分来得到最佳投影矢量;而LDA依据最大化类间离散度与类内离散度的比率来得到最佳的投影矢量.然而,LDA在处理高维的小样本数据(如人脸数据)时会产生小样本问题(SSS).为了解决这一问题,研究者提出了一系列的方法<sup>[6-10]</sup>,其中基于最大间距的方

法<sup>[11-12]</sup>被广泛地使用.该方法将最大化类间离散度与类内离散度的比率改为最大化类间离散度与类内离散度的差值,这样不仅可以解决小样本问题,同时也降低了时间复杂度<sup>[13]</sup>.

LDA方法是使用总体均值替代方差.而根据统计学习理论<sup>[14]</sup>,总体均值反映的是样本总的分布信息和全局特征,适合表示具有明显高斯分布特征的数据(如有团状结构特点的数据).所以,LDA方法在处理具有明显局部流形的数据(如人脸数据、高维文本数据等)时缺乏适应性.常用的一些经典降维方法无法揭示数据的内在的局部流形结构,因此局部流

收稿日期: 2013-01-21; 修回日期: 2013-03-26.

基金项目: 国家自然科学基金项目(61375001, 61272210); 江苏省自然科学基金项目(BK2011417); 苏州大学江苏省计算机信息处理技术重点实验室开放课题项目(KJS1126); 江苏省新型环保重点实验室开放课题项目(AE201068).

作者简介: 皋军(1971-),男,副教授,博士后,从事数据挖掘、人工智能等研究;王士同(1964-),男,教授,博士生导师,从事模糊系统、知识发现、图像识别等研究.

形学习方法<sup>[15-16]</sup>逐渐发展起来,其中局部保持投影法(LPP)<sup>[17]</sup>被广泛地研究和应用。

本文在LDA方法的基础上,引入局部加权均值(LWM)<sup>[18]</sup>,提出了基于局部加权均值的最大间距判别分析(LBMMC)法。通过分析LBMMC方法,可以发现该方法具有以下4个优势:1) LBMMC方法引入MMC后,解决了小样本问题;2) LBMMC引入QR分解技术,降低了算法的时间复杂度;3) LBMMC方法使用LWM来替代标准均值,反映了样本空间内部不同样本对保持样本内在的局部几何结构的贡献程度不同;4) LBMMC作为一种有监督的方法,在方法构造过程中使用流形学习理论<sup>[15]</sup>对原样本数据进行有效的划分,构造了具有明显几何含义的类内散度和类间散度,提高了方法的泛化能力。

## 1 相关工作

本节简单回顾了LWM、LDA方法的概念。

### 1.1 局部加权均值

为了反映不同样本对保持样本内在的局部几何结构的贡献程度不同,LWM作为一种标准均值概念的泛化形式,具有这一功能。

**定义 1**<sup>[18]</sup> 假设  $X_{1q} = \{x_{1q}^i\}_{i=1}^{n_1}$  表示一个局部子域,  $X_{1q}$  的LWM可写为  $\sum_{i=1}^{n_1} \left( \beta_{qi} x_{1q}^i / \sum_{p=1}^{n_1} \beta_{qp} \right)$ , 其中:  $0 \leq \beta_{qi} \leq 1$  且  $\beta_{qi} = \exp\left(\frac{\|x_{1q} - x_{1q}^{(i)}\|^2}{h}\right)$  是权值参数( $\beta_{qi}$  只与局部子域  $X_{1q}$  中的样本相关),  $h$  是热核函数  $\exp(-d^2/h)$  的热核参数, 本文中令不同权值的热核参数  $h$  相等。

从定义1可以发现,LWM通过给在局部子域中的样本分配不同权值来表示不同样本在保持内在局部几何结构所反映的贡献程度的不同。

### 1.2 线性判别分析

**定义 2**<sup>[9]</sup> 假设在高维欧式空间  $R^d$  中有数据集  $X = \{x_1, \dots, x_n\}$ ,  $\forall x_i \in R^d$ .  $C$  为类别数,  $X_c$  为第  $C$  类、大小为  $n_c$  的一个样本子集。若已知  $\omega$  (最佳投影矢量), 则类内散度矩阵和类间散度矩阵分别为

$$S_W = \sum_{c=1}^C \sum_{x \in D_k} (x - u_c)(x - u_c)^T,$$

$$S_B = \sum_{c=1}^C n_c (u - u_c)(u - u_c)^T.$$

其中: 第  $c$  类均值  $u_c = \frac{1}{n_c} \sum_{x \in X_c} x$ ,  $c = 1, 2, \dots, C$ ; 样本总体均值  $u = \frac{1}{n} \sum_{x \in X} x$ .

根据定义2,可以得到LDA的目标函数为

$$\operatorname{argmax}_{\omega^T \omega = 1} J(\omega) = \operatorname{argmax}_{\omega^T \omega = 1} \frac{\omega^T S_B \omega}{\omega^T S_W \omega}.$$

从LDA的目标函数可以看出,LDA存在小样本问题;同时,由于使用了标准均值表示方差,该方法在一定程度上反映的是样本原始空间的全局结构和全局信息,而缺乏局部学习能力。为此,通过引入LWM到LDA中,提出了LBMMC方法。

## 2 基于局部子域的最大间距判别分析

由以上的分析可知,根据流形学习理论<sup>[18]</sup>,非高斯或流形分布的数据一般都可以分解成若干个可以看成具有局部高斯分布的子域,因此可将具有一定分布的数据集划分成小的局部数据子域,并在各个局部子域中使用LWM来替代标准均值,从而在一定程度上实现局部学习。

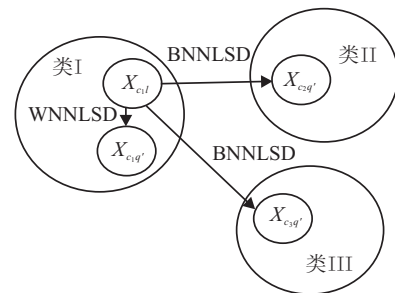


图1 LBMMC基本原理

由图1可以看出LBMMC方法构造的基本原理。假设存在3类(即  $C = 3$ ) 样本,对于任意一类  $X_c$  ( $c = 1, 2, 3$ ),按照如下方法划分局部子域:对于  $X_c$  类中的  $\forall x_{ci}$ ,定义  $x_{ci}$  在该类中的  $k_c$  个最近邻组成的数据子集为  $x_{ci}$  对应的局部子域,记为  $X_{ci}$ 。这样可将样本数为  $n_c$  的数据集  $X_c$  分解成  $n_c$  个数据局部子域。

假设  $X_{ci}$  表示为  $X_c$  中的任一局部子域,可将满足式(1)的  $X_{c1q'}$  称为  $X_{ci}$  在  $X_c$  类中的类内最近邻局部子域(WNNLSD),有

$$\operatorname{dis}(X_{ci}, X_{c1q'}) = \min_{\substack{q=1,2,\dots,n_c \\ q \neq i}} \operatorname{dis}(X_{ci}, X_{cq}) = \min_{\substack{q=1,2,\dots,n_c \\ q \neq i}} \|LWM(X_{ci}) - LWM(X_{cq})\|_F^2. \quad (1)$$

其中:  $\|\cdot\|_F$  表示  $F$  范数,  $LWM(\cdot)$  表示局部加权均值算子。

假设存在两类  $X_{c1}$  和  $X_{c2}$ ,且  $X_{c1i}$  为  $X_{c1}$  类中任意局部子域,则满足下式:

$$\operatorname{dis}(X_{c1i}, X_{c2q'}) = \min_{q=1,2,\dots,n_{c2}} \operatorname{dis}(X_{c1i}, X_{c2q}) = \min_{q=1,2,\dots,n_{c2}} \|LWM(X_{c1i}) - LWM(X_{c2q})\|_F^2. \quad (2)$$

的  $X_{c2q'}$  称为  $X_{c1i}$  在类  $X_{c2}$  中的类间最近邻局部子域(BNNLSD)。由此,可定义LBMMC方法中的类内散度矩阵和类间散度矩阵。

**定义 3** 假设有  $n$  个样本组成的样本集  $X =$

$\{x_1, x_2, \dots, x_n\}, \forall x_i \in R^d$ , 它们分别属于  $C$  个不同的类. 样本子集  $X_c$  属于第  $c$  类, 大小为  $n_c$ , 且  $\forall X_{ci} (i = 1, 2, \dots, n_c)$  为  $X_c$  的任意一个局部子域, 给定分类决策平面的法向量  $\omega$ , 则 LBMMC 方法对应的类内散度和类间散度分别为

$$\alpha_W = \sum_{c=1}^C \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} r_{ij}^{w_c} \|LWM(\omega^T X_{ci}) - LWM(\omega^T X_{cj})\|_F^2, \quad (3)$$

$$\alpha_B = \sum_{c_1=1}^C \sum_{c_2=1}^C \sum_{i=1}^{n_{c_1}} \sum_{j=1}^{n_{c_2}} r_{ij}^{b_{c_1 c_2}} \|LWM(\omega^T X_{c_1 i}) - LWM(\omega^T X_{c_2 j})\|_F^2. \quad (4)$$

其中

$$r_{ij}^{w_c} = \begin{cases} 1, & X_{c_j} \text{ is the WNNLSD of the } X_{c_i} \text{ or} \\ & X_{c_i} \text{ is the WNNLSD of the } X_{c_j}; \\ 0, & \text{otherwise} \end{cases}$$

为类内任意两个局部子域之间的关联系数.

$$r_{ij}^{b_{c_1 c_2}} = \begin{cases} 1, & X_{c_2 j} \text{ is the WNNLSD of the } X_{c_1 i} \text{ or} \\ & X_{c_1 i} \text{ is the WNNLSD of the } X_{c_2 j}; \\ 0, & \text{otherwise} \end{cases}$$

为类间任意两个局部子域之间的关联系数.

LWMMC 方法的目标函数为

$$\operatorname{argmax}_{\omega^T \omega = 1} J(\omega) = (1 - \gamma)\alpha_B - \gamma\alpha_W. \quad (5)$$

为了在一定程度上提高 LWMMC 求解目标函数的效率, 可对式 (3) 和 (4) 进行改写和简化, 得到如下定理.

**定理 1** 根据定义 3, 式 (3) 和 (4) 分别简化为

$$\alpha_W = \operatorname{tr}(\omega^T X L_W X^T \omega), \quad (6)$$

$$\alpha_B = \operatorname{tr}(\omega^T X L_B X^T \omega). \quad (7)$$

其中:  $L_W$  为类内全局关联矩阵,  $L_B$  为类间全局关联矩阵.

**证明** 以式 (6) 的证明为例. 根据定义 1, 等式 (3) 中的  $LWM(\omega^T X_{ci})$ ,  $LWM(\omega^T X_{cj})$  分别可以表示为  $\sum_{m=1}^{k_c} (\beta_{ci}^{(m)} \omega^T x_{ci}^{(m)} / \sum_{p=1}^{k_c} \beta_{ci}^{(p)})$  和  $\sum_{m=1}^{k_c} (c \beta_{cj}^{(m)} \omega^T x_{cj}^{(m)} / \sum_{p=1}^{k_c} \beta_{cj}^{(p)})$ . 其中: 参数  $k_c$  为近邻数,  $x_{ci}^{(m)}$  为局部子域  $X_{ci}$  的第  $m$  个样本,  $x_{cj}^{(m)}$  为局部子域  $X_{cj}$  的第  $m$  个样本. 若将定义在局部子域上的权值扩充到整个样本集, 则上述两个权值可以表示为

$$\beta_{ci} = \left( \underbrace{0, \dots, 0}_{n_1}, \dots, \underbrace{\frac{\beta_{ci}^{(1)}}{\sum_{p=1}^{n_c} \beta_{ci}^{(p)}}, \dots, \frac{\beta_{ci}^{(n_c)}}{\sum_{p=1}^{n_c} \beta_{ci}^{(p)}}}_{n_c}, \dots, \underbrace{0, \dots, 0}_{n_c} \right),$$

$$\beta_{cj} = \left( \underbrace{0, \dots, 0}_{n_1}, \dots, \underbrace{\frac{\beta_{cj}^{(1)}}{\sum_{p=1}^{n_c} \beta_{cj}^{(p)}}, \dots, \frac{\beta_{cj}^{(n_c)}}{\sum_{p=1}^{n_c} \beta_{cj}^{(p)}}}_{n_c}, \dots, \underbrace{0, \dots, 0}_{n_c} \right).$$

从而可得

$$LWM(\omega^T X_{ci}) = \beta_{ci}^T X^T \omega,$$

$$LWM(\omega^T X_{cj}) = \beta_{cj}^T X^T \omega;$$

同时由于  $\|A\|_F^2 = \operatorname{tr}(A^T A)$ , 有

$$\|LWM(\omega^T X_{ci}) - LWM(\omega^T X_{cj})\|_F^2 = \operatorname{tr}(\omega^T X (\beta_{ci} \beta_{ci}^T + \beta_{cj} \beta_{cj}^T - 2\beta_{ci} \beta_{cj}^T) X^T \omega).$$

如果令  $R_{ij}^{w_c} = \operatorname{diag}(r_{ij}^{w_c}, \dots, r_{ij}^{w_c})$  为类内局部子

域关联系数矩阵,  $L_c^{ij} = \beta_{ci} \beta_{ci}^T + \beta_{cj} \beta_{cj}^T - 2\beta_{ci} \beta_{cj}^T$  为类内局部子域分布差异权值矩阵, 则式 (3) 可表示为

$$\begin{aligned} \alpha_W &= \sum_{c=1}^C \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} r_{ij}^{w_c} \|LWM(\omega^T X_{ci}) - LWM(\omega^T X_{cj})\|_F^2 = \\ &= \operatorname{tr}(\omega^T X \left( \sum_{c=1}^C \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} R_{ij}^{w_c} L_c^{ij} \right) X^T \omega) = \\ &= \operatorname{tr}(\omega^T X L_W X^T \omega), \end{aligned} \quad (8)$$

其中  $L_W = \sum_{c=1}^C \sum_{i=1}^{n_c} \sum_{j=1}^{n_c} R_{ij}^{w_c} L_c^{ij}$  为类内总体局部子域关联系数矩阵.

由上可得式 (6) 成立. 同理可以证明式 (7) 成立, 其中

$$\begin{aligned} L_B &= \sum_{c_1=1}^C \sum_{c_2=1}^C \sum_{i=1}^{n_{c_1}} \sum_{j=1}^{n_{c_2}} R_{ij}^{b_{c_1 c_2}} L_{c_1 c_2}^{ij}, \\ L_{c_1 c_2}^{ij} &= \beta_{c_1 i} \beta_{c_1 i}^T + \beta_{c_2 j} \beta_{c_2 j}^T - 2\beta_{c_1 i} \beta_{c_2 j}^T, \\ R_{ij}^{b_{c_1 c_2}} &= \operatorname{diag}(\underbrace{r_{ij}^{b_{c_1 c_2}}, \dots, r_{ij}^{b_{c_1 c_2}}}_n). \end{aligned}$$

由此, 定理得证.  $\square$

如果直接将式 (6) 和 (7) 代入 LBMMC 方法的目标函数 (5), 则当使用 SVD 方法求解时将遇到计算困难, 特别是处理高维数据集时, 尤为明显. 在求解式 (5) 时, 时间复杂度可以达到  $O(d^3)$ . 当处理高维小样本数据时, 有  $d \ll n$ . 故在求解式 (5) 之前首先对数据集  $X$  采用 QR 分解, 对应的时间复杂度将大大降低, 有  $X = QR$ . 其中  $Q \in \mathbb{R}^{d \times r}$  由一组正交列向量组成,  $R \in \mathbb{R}^{r \times n}$  是上三角矩阵,  $r = \operatorname{rank}(X)$ . 求解式 (5) 可转变为求解下式:

$$\operatorname{argmax}_{z^T z = 1} J(z) = (1 - \gamma) z^T R L_B R^T z - \gamma z^T R L_W R^T z. \quad (9)$$

当得到式(9)中的 $z$ 后,有 $\omega = Qz$ , $\omega$ 即为式(5)的解.通过上述定义和定理,可得到本文的LBMMC算法如下.

**Input:** 数据集 $X$ ,并赋初值 $k_c$ 和 $h$ ;

**Output:** 变换矢量 $\omega$ .

**Step 1:** 根据给定的 $k_c$ ,将数据集 $X$ 划分为若干个局部子域;

**Step 2:** 根据定义1,计算局部子域中每一个样本对应的权值 $\beta_{ci}^m$ ;

**Step 3:** 根据式(1)和(2)分别计算每一个局部子域 $X_{c_1i}$ 对应的类内最近邻局部子域 $X_{c_1q'}$ 和类间最近邻局部子域 $X_{c_2q'}$ ;

**Step 4:** 在数据集 $X$ 上进行QR分解,则 $Q \in \mathbb{R}^{d \times r}$ , $R \in \mathbb{R}^{r \times n}$ ;

**Step 5:** 求解式(9),得到满足式(9)最大特征值对应的单位矢量 $z$ ;

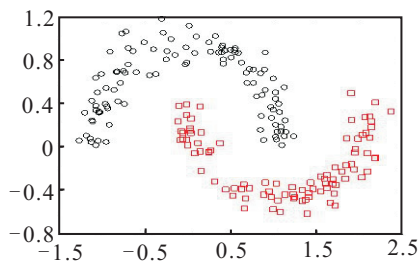
**Step 6:** 计算 $\omega = Qz$ ,并输出 $\omega$ .

### 3 实验分析

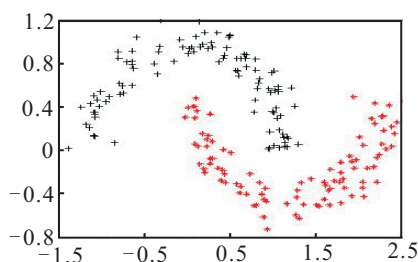
为了说明本文算法LBMMC具有较好的特征提取能力,选取以下几个数据集进行测试.

#### 3.1 测试人造数据集

虽然2moons的维数不高,但图2所示2moons数据集的结构具有明显局部流形,故2moons数据集多用于测试算法的局部学习能力.通过测试2moons数据集,同时与LDA方法相比较,说明了本文LBMMC方法在处理局部流形数据的有效性.



(a) 训练数据



(b) 测试数据

图2 2moons数据集

在测试过程中,设本文的参数 $k_c = [2, 3, 4, 5]$ , $h = [2^{-10}, 2^{-5}, 2^{-3}, 2^{-1}, 2^0, 2^2, 2^4, 2^6, 2^{10}]$ ,使用10-折交叉验证.测试结果见表1.

表1 LDA与LBMMC在2moons数据集测试结果比较

items	LDA	LBMMC
training accuracy	1	1
test accuracy	0.84	0.885( $K=5;h=1;r=0.19$ )

由表1可以看出:LDA和LBMMC在2moons数据集上进行测试,训练精度都可以达到1(最大).而LBMMC方法的测试精度明显高于LDA方法,说明了LBMMC的局部学习能力优于LDA.LBMMC在处理局部流形数据上确实具有优势,由于引入了LWM,算法的局部学习能力得到提高.

#### 3.2 测试UCI数据集

由于UCI数据集被广泛用于验证算法的特征提取能力,本文通过测试UCI数据集的4个数据集(见表2)来说明LBMMC方法的特征提取能力.由于PCA、LDA、MMC、LPP均为经典的降维算法,其在处理不同的数据集时所表现出的特征提取能力均不同.为了更好地说明本文方法具有较强的特征提取能力,将LBMMC方法同以上几种算法进行比较.在测试过程中LBMMC的参数同于3.1节,本文使用10-折交叉验证,测试效果见表3.

表2 Ionosphere、IRIS、Wine、Glass数据集样本性质

items	datasets			
	Ionosphere	IRIS	Wine	Glass
training samples	230	100	118	142
test samples	117	50	60	72
features	33	4	12	9
subjects	2	3	3	6

注:wine数据集是按照6份12个样本,3份11个样本,1份13个样本进行划分,其他是平均划分.

通过分析表3中的数据,有如下结论:

1) 本文的LBMMC方法同其他4种方法相比,其测试精度均大于同数据集不同的训练方法,故LBMMC在处理数据集时表现出较好的特征提取效果和分类精度.这说明LBMMC的局部学习能力相比LDA和PCA有所提高.

2) LBMMC通过引入类内散度和类间散度,在反映样本内全局结构和全局信息的同时保持了样本内在的局部鉴别信息,一定程度上提高了特征提取的能力.

3) 在测试过程中,虽然Ionosphere数据集所对应的类内散度矩阵是非奇异的,但是在LBMMC方法中却没有产生LDA方法中的小样本问题.这充分说明在LBMMC中引入最大间距准则是有效的.

#### 3.3 测试人脸数据集

人脸图像数据集是典型的蕴含着内在的非线性流形结构<sup>[14,16]</sup>的数据.为了说明LBMMC方法不存在小样本问题和其较好的特征提取能力,在ORL人脸数

表3 各算法在 Ionosphere、IRIS、Wine、Glass 数据集测试结果比较

algorithm	Ionosphere		IRIS		Wine		Glass	
	accuracy	dim	accuracy	dim	accuracy	dim	accuracy	dim
PCA	0.931 62	5	0.96	3	0.706 9	1	0.819 44	4
LDA	0.786 32	1	0.94	2	0.689 66	6	0.777 78	4
MMC	0.897 4	5	0.94	4	0.7	1	0.819 44	7
LPP	0.786 32	1	0.96	3	0.666 67	5	0.833 33	5
LBMMC	$(K = 5; h = 2^4)$		$(K = 5; h = 2^{-1})$		$(K = 2; h = 2^{-3})$		$(K = 5; h = 2^{-5})$	
	0.940 17	8	0.98	2	0.766 67	1	0.833 33	6
	$(K = 3; h = 2^4;$ $r = 2e-8)$		$(K = 5; h = 2^{-1};$ $r = 2e-8)$		$(K = 5; h = 2^{-10};$ $r = 2e-8)$		$(K = 3; h = 2^{-3};$ $r = 2e-8)$	

表4 各算法在 ORL、Yale 数据集测试结果比较

algorithm	ORL				Yale			
	4		6		4		6	
	accuracy	dim	accuracy	dim	accuracy	dim	accuracy	dim
PCA	0.854 17	53	0.887 5	47	0.638 1	47	0.666 67	34
PCA+LDA	0.887 5	34	0.918 75	34	0.685 71	13	0.72	16
MMC	0.9	43	0.912 5	34	0.666 67	15	0.786 67	17
PCA+LPP	0.879 17	69	0.931 25	31	0.638 1	18	0.76	80
LBMMC	$(K = 5; h = 2^4)$		$(K = 5; h = 2^5)$		$(K = 5; h = 2^{-6})$		$(K = 5; h = 2^{-4})$	
	0.904 17	74	0.937 53	81	0.685 71	20	0.8	40
	$(K = 5; h = 2^{20};$ $r = 2^{-2})$		$(K = 5; h = 2^{50};$ $r = 2^{-3})$		$(K = 5; h = 2^{20};$ $r = 2^{-4})$		$(K = 5; h = 2^{20};$ $r = 2^{-3})$	

数据集(共40个类, 每类10个样本)、Yale人脸数据集(共15个类, 每类11个样本)上进行测试。

为了更好地反映本文方法LBMMC具有一定程度上的特征提取能力, 将与另外的4种经典算法的测试结果(PCA、PCA+LDA、PCA+LPP、MMC)进行对比。为了更全面地表明算法的有效性, 在ORL数据集和Yale数据集中每类各取2个不同数量的样本作为训练样本(每类各取4个和6个), 分别使用4-折交叉验证和6-折交叉验证。其中LBMMC方法中参数设为

$k_c = [2, 3, 4, 5]$ ,  $h = [2^{10}, 2^{20}, 2^{30}, 2^{40}, 2^{50}]$ , 使用10-折交叉验证。测试结果见表4、图3和图4。

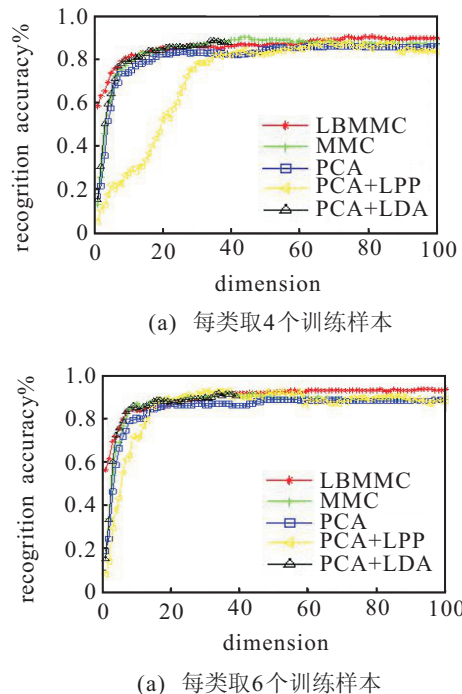


图3 ORL数据特征提取结果(维数-识别精度图)

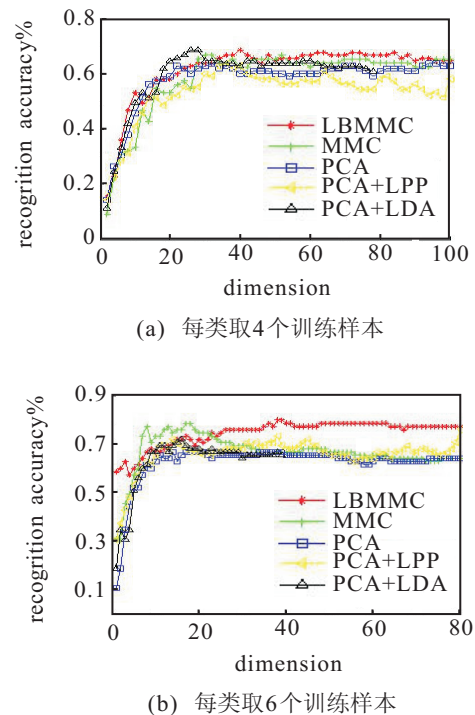


图4 Yale数据特征提取结果(维数-识别精度图)

通过分析上述数据和图表, 发现本文提出的LBMMC方法的测试结果优于其他4种方法的结果, 说明本文方法在处理人脸数据上具有良好的特征提取能力。同时: 1) 由于引入了LWM, LBMMC的局部学习能力加强, 测试精度提高; 2) 由于引入了最大间距判别分析, 减少了产生小样本问题的可能性; 3) 将

QR 分解技术引入到算法中, 大大降低了时间复杂度.

## 4 结 论

本文通过分析 LDA 方法存在的问题, 提出了具有一定局部学习能力的 LBMCM 方法. 该方法首先根据流形学习理论对原样本进行有效划分, 然后将局部加权均值的概念引入每个划分得到的局部子域上. 通过 LWM 表明了不同样本对保持局部子域的内在几何结构和局部信息上的贡献差异是有所不同的. 为了提高算法的有效性, 引进了 QR 分解技术, 并通过测试人造、真实数据集表明了本文方法具有较强的局部学习能力. LBMCM 方法也存在一定的不足, 算法的空间复杂程度和执行效率还有待提高, 这将是进一步研究的内容.

## 参考文献(References)

- [1] Bian Z Q, Zhang X G. Pattern recognition[M]. Beijing: Tsinghua University, 2001.
- [2] Jolliffe I T. Principal component analysis[M]. New York: Springer-Verlag, 1986.
- [3] Todorov, Valentin, Filzmoser, et al. Comparing classical and robust sparse PCA[C]. Advances in Intelligent Systems and Computing. Germany: Springer Verlag, 2013: 1283-291.
- [4] Li Rong-Hua, Liang Shuang, Baci George, et al. Equivalence between LDA/QR and direct LDA[J]. Int J of Cognitive Informatics and Natural Intelligence, 2011, 5(1): 94-112.
- [5] Dhir Chandra Shekhar, Lee Soo-Young. Discriminant independent component analysis[J]. IEEE Trans on Neural Networks, 2011, 22(6): 845-857.
- [6] Deng Weihong, Liu Yebin, Hu Jiani, et al. The small sample size problem of ICA: A comparative study and analysis[J]. Pattern Recognition, 2012, 45(12): 4438-4450.
- [7] Yang Li-Ping, Gu Xiao-Hua, Ye Hong-Wei. Sample locality preserving discriminant analysis for classification[J]. Guangxue Jingmi Gongcheng/Optics and Precision Engineering, 2011, 19(9): 2205-2213.
- [8] Kwak, Nojun. Kernel discriminant analysis for regression problems[J]. Pattern Recognition, 2012, 45(5): 2019-2031.
- [9] Kumar Nitin, Jaiswal Ajay, Agrawal R K. Performance evaluation of subspace methods to tackle small sample size problem in face recognition[C]. ACM Int Conf on Proc Series. United States: Association for Computing Machinery, 2012: 938-944.
- [10] Shu Xin, Gao Yao, Lu Hongtao. Efficient linear discriminant analysis with locality preserving for face recognition[J]. Pattern Recognition, 2012, 45(5): 1892-1898.
- [11] Yang Wankou, Sun Changyin, Du Helen S, et al. Feature extraction using laplacian maximum margin criterion[J]. Neural Processing Letters, 2011, 33(1): 99-110.
- [12] Cui Yan, Fan Liya. Feature extraction using fuzzy maximum margin criterion[J]. Neurocomputing, 2012, 86(1): 52-58.
- [13] Wan Minghua, Lai Zhihui, Jin Zhong. Feature extraction using two-dimensional local graph embedding based on maximum margin criterion[J]. Applied Mathematics and Computation, 2011, 217(23): 9659-9668.
- [14] Vanpanik V. Statistical learning theory[M]. New York: Wiley, 1998.
- [15] Benabdeslem, Khalid, Hindawi, et al. Constrained Laplacian score for semi-supervised feature selection[C]. Lecture Notes in Computer Science. Germany: Springer Verlag, 2011: 204-218.
- [16] Lou Songjiang, Zhang Guoyin, Pan Haiwei, et al. Supervised Laplacian discriminant analysis for small sample size problem with its application to face recognition[J]. Computer Research and Development, 2012, 49(8): 1730-1737.
- [17] Wong W K, Zhao H T. Supervised optimal locality preserving projection[J]. Pattern Recognition, 2012, 45(1): 186-197.
- [18] Atkeson Christopher G, Moore Andrew W, Schaal Stefan. Locally weighted learning[J]. Artificial Intelligence Review, 1997, 11(1/2/3/4/5): 75-113.
- [19] Blake C L, Merz C J. UCI repository of machine learning databases[EB/OL]. 1998. <http://www.ics.uci.edu/mllearn/MLRepository.html>.

(责任编辑: 孙艺红)