

基于子块优化及全局整合的局部判别投影法

郑建炜* 王万良 姚信威

(浙江工业大学计算科学与技术学院 杭州 310023)

摘要: 已有投影算法都直接通过完整的输入训练集求解最佳变换矩阵, 难以进行增量式学习扩展。针对此问题, 该文通过组合优化策略提出局部判别投影方法应用于分类问题。该算法同时包括类间判别信息和类内局部保持特征, 求得的变换矩阵还具有正交性。此外, 利用核函数将算法扩展至非线性应用, 使之可以适应更多的数据类型。在 ORL 人脸库和小样本说话人辨认应用中验证了该算法的有效性。

关键词: 模式识别; 局部判别投影; 组合优化策略; 核函数; 子空间学习

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2011)09-2175-06

DOI: 10.3724/SP.J.1146.2010.01358

Locally Discriminant Projection Algorithm Based on the Block Optimization and Combination Strategy

Zheng Jian-wei Wang Wan-liang Yao Xin-wei

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: It is difficult for the traditional projection algorithms to extend to incremental learning since they use the whole training sets for solving out the projection matrix directly. To tackle this problem, a novel method, named Block optimization and Combination strategy, used for Locally Discriminant Projection (BCLDP) is proposed. This method takes into account both intra-class and interclass geometries; and has the orthogonality property. Furthermore, BCLDP is extended to nonlinear case using kernel function, which makes BCLDP better suits for diverse application. The experiments on ORL face database and speaker identification application demonstrate the effectiveness of the proposed algorithm.

Key words: Pattern Recognition; Locally discriminant projection; Block optimization and combination; Kernel function; Subspace learning

1 引言

子空间投影分类算法已经成功应用于特征提取^[1]、人脸识别^[2]等领域, 引起了人们的广泛关注。传统的方法中, 主成分分析(Principal Component Analysis, PCA)和线性鉴别分析(Linear Discriminant Analysis, LDA)是两种最为常见的技术。PCA和LDA用于描述输入数据中的线性变化, 在揭示数据的欧氏结构时是有效的, 却无法发现嵌入在数据中的低维子流形。邻域保持嵌入(Neighborhood Preserving Embedding, NPE)^[3]和保局部投影(Locality Preserving Projection, LPP)^[4]两种算法则旨在挖掘数据集中的邻域结构, 在高维数据可视化方面效果显著, 但由于缺少全局判别信息, 分类性能偏弱。针对存在的不足, 已经涌现出

众多改进型投影算法, 例如正交型邻域保持嵌入^[5]和有监督保局部投影^[6]方法, 两者都大幅地提升了原算法的分类性能, 但仍然缺少全局鉴别信息。局部敏感判别分析(Locality Sensitive Discriminant Analysis, LSDA)^[7]通过类内邻接图和类间邻接图构建变换矩阵求取函数, 最大化类间边界的同时仍保持类内邻域结构, 更适于数据分类或可视化。除此之外, 偏判别欧氏嵌入^[8], 线性鉴别投影^[9], 最大边界子空间投影^[10]等也都是兼具保局性和判别性的优秀线性降维技术。

线性算法在描述数据非线性结构方面能力不足, 近年来, 人们提出了一些旨在发掘高维图像数据中非线性因素的学习算法, 如等度规映射(Isometric Mapping, ISOMAP)^[11]、局部线性嵌入(Locally Linear Embedding, LLE)^[12]和拉普拉斯特征映射(Laplacian Eigenmap, LE)^[13]等。这些非线性降维技术在理论上可以发现复杂的低维嵌入, 并且在人造测试数据上产生了很好的效果, 但是在很多实际应用中却并没有取得比传统的线性技术更好的

2010-12-09 收到, 2011-06-10 改回

国家自然科学基金(61070043)和浙江省自然科学基金(Y1100611)资助课题

*通信作者: 郑建炜 zjw@zjut.edu.cn

结果。另外, 限制这些非线性流形学习方法在实际应用中得到推广的一个更大的障碍是: 它们只能产生定义在训练数据点集上的维数约简映射, 不能为新的测试点建立映射关系。而基于核函数的非线性技术并不存在该问题, 因此受到众多研究者的青睐。通过核技术, 可以轻易地对已有的线性投影算法进行非线性扩展, 例如核主成分分析(kernel PCA, kPCA), 核化线性鉴别分析(kernel LDA, kLDA), 核化局部敏感判别分析(kernel LSDA, kLSDA)^[7]等等。

上述所有算法都直接通过完整的输入训练集求解最佳投影矩阵, 虽然都拥有出色的应用性能, 但实现过程缺乏灵活性, 算法扩展能力较弱。基于此, 本文采用组合优化(Block optimization and Combination)策略构建局部判别正交投影(Locally Discriminant Projection, LDP)方法, 简称为BCLDP。该策略通过对输入样本逐个构建邻域子分块, 给出子优化函数并加以全局整合, 最终的目标泛函不仅具有类间判别能力还能够保持类内局部特性, 而且算法可以直接进行增量式学习和核化扩展。

2 局部判别投影法

给定训练集 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in R^{m \times n}$, 表示 m 维空间中的 n 个样本数据, BCLDP 的目标就是寻找一个变换矩阵 \mathbf{U} , 通过 $\mathbf{Y} = \mathbf{U}^T \mathbf{X}$ 将训练集合映射为 $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n] \in R^{d \times n}$, 其中 $d \ll m$, 并且映射后的 $\mathbf{y}_i (i = 1, \dots, n)$ 具有最佳的类内紧性和最大的类间散度。变换矩阵 \mathbf{U} 通过组合优化策略^[14]求取, 主要包括两个步骤, 即: 分块最优和全局整合。在分块最优步骤, 先针对每个训练样本构建一个子分块, 并设计相应的优化目标泛函; 全局整合过程即联合所有分块提出全局的投影策略, 并在正交约束条件下生成变换矩阵 \mathbf{U} 。

2.1 分块最优

对于任意的 $\mathbf{x}_i \in \mathbf{X}, i = 1, \dots, n$, 先选取 k_w 个同类近邻样本, 其集合表达为 $N_w(\mathbf{x}_i) = \{\mathbf{x}_{ij}^w | c(\mathbf{x}_{ij}^w) = c(\mathbf{x}_i), 1 \leq j \leq k_w\}$, 其中 $c(\mathbf{x}_i)$ 表示样本 \mathbf{x}_i 的所属类别; 再选择 k_b 个异类近邻样本, 相应集合表达为 $N_b(\mathbf{x}_i) = \{\mathbf{x}_{ij}^b | c(\mathbf{x}_{ij}^b) \neq c(\mathbf{x}_i), 1 \leq j \leq k_b\}$ 。选择过程中, 样本间距离测度选用热核函数, 即

$$w_{i,j} = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2t^2}\right)$$

通过合并近邻集合 $N_w(\mathbf{x}_i)$ 和 $N_b(\mathbf{x}_i)$ 后, 围绕 \mathbf{x}_i 所得的子分块即为

$$\mathbf{X}_i = \{\mathbf{x}_i, \mathbf{x}_{i1}^w, \dots, \mathbf{x}_{ik_w}^w, \mathbf{x}_{i1}^b, \dots, \mathbf{x}_{ik_b}^b\}$$

将 \mathbf{X}_i 通过局部判别投影后的低维数据表达为

$\mathbf{Y}_i = \{\mathbf{y}_i, \mathbf{y}_{i1}^w, \dots, \mathbf{y}_{ik_w}^w, \mathbf{y}_{i1}^b, \dots, \mathbf{y}_{ik_b}^b\}$, 在此子空间中, 同类数据间距离得以缩小, 而异类样本则趋于分散, 具体投影效果如图 1 所示。

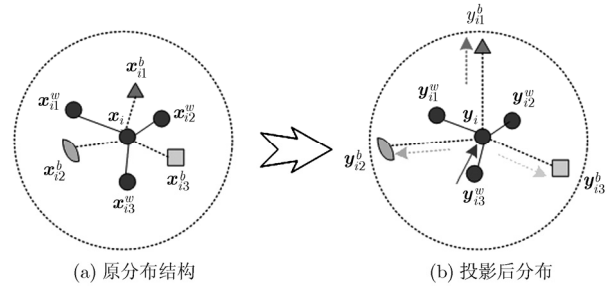


图 1 局部判别投影的子分块投影效果

图 1(a)为原 \mathbf{X}_i 分块的分布结构, 圆形为同类样本, 而方形、三角形、椭圆形数据则为异类样本, 经局部判别投影后数据分布如图 1(b)所示, 其中同类样本都趋向于靠近 \mathbf{y}_i , 而其他数据则尽可能的远离 \mathbf{y}_i , 有利于分类任务的实现。

根据图 1 所示, BCLDP 要求同类样本经投影后趋于紧凑, 其目标函数描述为

$$\arg \min_{\mathbf{y}_i} \sum_{j=1}^{k_w} \|\mathbf{y}_i - \mathbf{y}_{ij}^w\|^2 w_{ij}^w \quad (1)$$

其中

$$w_{ij}^w = \begin{cases} w_{ij}, & \mathbf{x}_j \in N_w \\ 0, & \text{其它} \end{cases}$$

为类内样本距离权值。

引入标记 \mathbf{Y}_i 后, 式(1)可以改写成:

$$\begin{aligned} & \arg \min_{\mathbf{Y}_i} \text{tr}(\mathbf{Y}_i \mathbf{L}_i^w \mathbf{Y}_i^T) \\ & = \arg \min_{\mathbf{Y}_i} \text{tr} \left(\mathbf{Y}_i \begin{bmatrix} D_i^w & -\mathbf{w}_i^{wT} & \mathbf{z}_b^T \\ -\mathbf{w}_i^w & \mathbf{W}_i^w & \mathbf{z}_w \mathbf{z}_b^T \\ \mathbf{z}_b & \mathbf{z}_b \mathbf{z}_w^T & \mathbf{z}_b \mathbf{z}_b^T \end{bmatrix} \mathbf{Y}_i^T \right) \end{aligned} \quad (2)$$

其中 $\mathbf{w}_i^w = \{w_{i1}^w, \dots, w_{ik_w}^w\}^T$ 是类内距离权值向量, $D_i^w = \sum_{j=1}^{k_w} w_{ij}^w$, \mathbf{z}_w 和 \mathbf{z}_b 分别是 k_w, k_b 维零向量, $\mathbf{W}_i^w = \text{diag}(\mathbf{w}_i^w)$ 是元素为 w_{ij}^w 的对角矩阵。

另一方面, BCLDP 希望异类样本远离分块中心 \mathbf{x}_i , 其目标函数可以描述为

$$\arg \max_{\mathbf{y}_i} \sum_{j=1}^{k_b} \|\mathbf{y}_i - \mathbf{y}_{ij}^b\|^2 w_{ij}^b \quad (3)$$

其中

$$w_{ij}^b = \begin{cases} w_{ij}, & \mathbf{x}_j \in N_b \\ 0, & \text{其它} \end{cases}$$

为类间样本距离权值。

同样, 引入 \mathbf{Y}_i 后, 式(3)可以改写成

$$\begin{aligned} & \arg \max_{\mathbf{Y}_i} \text{tr}(\mathbf{Y}_i \mathbf{L}_i^b \mathbf{Y}_i^T) \\ & = \arg \max_{\mathbf{Y}_i} \text{tr} \left(\mathbf{Y}_i \begin{bmatrix} D_i^b & \mathbf{z}_w^T & -\mathbf{w}_i^{bT} \\ \mathbf{z}_w & \mathbf{z}_w \mathbf{z}_w^T & \mathbf{z}_w \mathbf{z}_b^T \\ -\mathbf{w}_i^b & \mathbf{z}_b \mathbf{z}_w^T & \mathbf{W}_i^b \end{bmatrix} \mathbf{Y}_i^T \right) \end{aligned} \quad (4)$$

其中 $\mathbf{w}_i^b = \{w_{i1}^b, \dots, w_{ik_b}^b\}^T$ 为类间权值向量, $D_i^b = \sum_{j=1}^{k_b} w_{ij}^b$, $\mathbf{W}_i^b = \text{diag}(\mathbf{w}_i^b)$ 是元素为 w_{ij}^b 的对角矩阵。

综上所述, 对于任一子分块 \mathbf{X}_i , 联合考虑式(2)和式(4)后, 其最佳投影公式可表示为

$$\begin{aligned} & \arg \max_{\mathbf{Y}_i} \left((1-\beta) \sum_{j=1}^{k_b} \|\mathbf{y}_i - \mathbf{y}_{ij}^b\|^2 w_{ij}^b \right. \\ & \quad \left. - \beta \sum_{j=1}^{k_w} \|\mathbf{y}_i - \mathbf{y}_{ij}^w\|^2 w_{ij}^w \right) \\ & = \arg \max_{\mathbf{Y}_i} \text{tr}(\mathbf{Y}_i ((1-\beta)\mathbf{L}_i^b - \beta\mathbf{L}_i^w) \mathbf{Y}_i^T) \\ & = \arg \max_{\mathbf{Y}_i} \text{tr}(\mathbf{Y}_i \mathbf{L}_i \mathbf{Y}_i^T) \end{aligned} \quad (5)$$

其中 $\beta \in [0, 1]$ 是可调整参数, 用于平衡类内紧性与类间散度, \mathbf{L}_i 中封装了子分块的邻域信息, 其具体表达式为

$$\mathbf{L}_i = \begin{bmatrix} (1-\beta)D_i^b - \beta D_i^w & \beta \mathbf{w}_i^{wT} & -(1-\beta)\mathbf{w}_i^{bT} \\ \beta \mathbf{w}_i^w & -\beta \mathbf{W}_i^w & \mathbf{z}_w \mathbf{z}_b^T \\ -(1-\beta)\mathbf{w}_i^b & \mathbf{z}_b \mathbf{z}_w^T & (1-\beta)\mathbf{W}_i^b \end{bmatrix}$$

2.2 全局整合

设 $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ 为完整的训练样本投影后低维子空间, 根据局部判别投影思想, 任意 \mathbf{X}_i 的投影方法如式(5)所示, 而全局目标泛函则可通过联合所有的子分块获取。先构建选择矩阵 $\mathbf{S}_i \in R^{n \times (k_w + k_b + 1)}$ 如下:

$$(\mathbf{S}_i)_{le} = \begin{cases} 1, & l = (\mathbf{F}_i)_e \\ 0, & \text{其它} \end{cases} \quad (6)$$

其中 $\mathbf{F}_i = \{i, i_1^w, \dots, i_{k_w}^w, i_1^b, \dots, i_{k_b}^b\}$ 表示第 i 个分块的邻域索引集。这样, 子分块 \mathbf{Y}_i 即可由 \mathbf{Y} 计算获取

$$\mathbf{Y}_i = \mathbf{Y} \mathbf{S}_i \quad (7)$$

相应地, 式(5)可改写为

$$\arg \max_{\mathbf{Y}} \text{tr}(\mathbf{Y} \mathbf{S}_i \mathbf{L}_i \mathbf{S}_i^T \mathbf{Y}^T) \quad (8)$$

通过将如式(8)所示的所有子分块进行叠加, 即可得到全局投影目标为

$$\arg \max_{\mathbf{Y}} \sum_{i=1}^n \text{tr}(\mathbf{Y} \mathbf{S}_i \mathbf{L}_i \mathbf{S}_i^T \mathbf{Y}^T) = \arg \max_{\mathbf{Y}} \text{tr}(\mathbf{Y} \mathbf{L} \mathbf{Y}^T) \quad (9)$$

其中 $\mathbf{L} = \sum_{i=1}^n \mathbf{S}_i \mathbf{L}_i \mathbf{S}_i^T \in R^{n \times n}$ 是一个联合矩阵, 可通过迭代过程获取: 初始化 $\mathbf{L}^{t=0} = 0$ 后, $\mathbf{L}^{t+1} = \mathbf{L}^t$

+ $\mathbf{S}_i \mathbf{L}_i \mathbf{S}_i^T$, 其中 t 为迭代次数, $i \in \{1, \dots, n\}$ 是当前子分块索引值, 当所有 \mathbf{L}_i 叠加完成后迭代结束。

为使变换矩阵 \mathbf{U} 具有正交性, 除约束条件 $\mathbf{Y} = \mathbf{U}^T \mathbf{X}$ 外, 添加约束 $\mathbf{U}^T \mathbf{U} = \mathbf{I}_d$, 其中 \mathbf{I}_d 是一个 $d \times d$ 的单位矩阵。相应式(9)变为

$$\left. \begin{aligned} & \arg \max_{\mathbf{Y}} \text{tr}(\mathbf{U}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{U}) \\ & \text{s.t. } \mathbf{U}^T \mathbf{U} = \mathbf{I}_d \end{aligned} \right\} \quad (10)$$

显然, \mathbf{U} 可通过求解如下标准特征值分解问题获取:

$$\mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{u} = \lambda \mathbf{u} \quad (11)$$

即 $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_d\}$, 其中 $\mathbf{u}_i, i \in \{1, \dots, d\}$ 是式(11)特征值 $\lambda_1 > \lambda_2 > \dots > \lambda_d$ 相应的特征向量。

从上述求解过程可以发现, BCLDP 本身已具备增量式学习能力, 由于 \mathbf{L} 矩阵可通过迭代过程求得, 当有新的训练集出现时, 可以在原 \mathbf{L} 的基础上继续叠加子分块, 所需要的工作仅为: 新样本邻域分块的确定和选择矩阵 \mathbf{S} 的计算。

3 核局部判别投影法

当数据流形具有强非线性特征时, 线性投影算法 BCLDP 难以捕捉其本质几何结构, 因此有必要扩展至再生核希尔伯特空间(Reproducing Kernel Hilbert Space, RKHS)进行局部判别投影, 称之为核化的 BCLDP(kernel BCLDP, kBCLDP)。

保持原有训练样本 \mathbf{X} 不变, 通过某种非线性映射, 将问题移至特征空间 F 中求解:

$$\phi: \mathbf{X} \rightarrow F$$

如果 ϕ 选择正确, 则 F 空间中的内积可由输入空间的核函数替代, 即

$$\langle \phi(x), \phi(y) \rangle = k(x, y)$$

在 F 空间中, 联合矩阵 \mathbf{L}^ϕ 的求取与输入空间基本一致, 仅需要调整样本间权值的计算, 方法如下:

$$\begin{aligned} w_{i,j}^\phi &= \exp \left(\frac{-\|\phi(\mathbf{x}_i) - \phi(\mathbf{x}_j)\|^2}{2t^2} \right) \\ &= \exp \left(\frac{-(k(\mathbf{x}_i, \mathbf{x}_i) - 2k(\mathbf{x}_i, \mathbf{x}_j) + k(\mathbf{x}_j, \mathbf{x}_j))}{2t^2} \right) \end{aligned}$$

设 $\Phi = [\phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_n)]$ 是 RKHS 空间的特征样本, 则相应的特征求解问题可以重写为

$$\Phi \mathbf{L}^\phi \Phi^T \mathbf{u} = \lambda \mathbf{u} \quad (12)$$

式(12)所求得特征向量可由 Φ 中的样本线性表达, 即存在系数 $\alpha_i, i = 1, \dots, n$, 使得

$$\mathbf{u} = \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i) = \Phi \boldsymbol{\alpha}$$

其中 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^T \in R^n$, 将上式代入式(12)中, 则

$$\left. \begin{aligned} \Phi L^{\phi} \Phi^T u &= \lambda u \\ \Rightarrow \Phi^T \Phi L^{\phi} \Phi^T \Phi \alpha &= \lambda \Phi^T \Phi \alpha \\ \Rightarrow K L^{\phi} K \alpha &= \lambda K \alpha \end{aligned} \right\} \quad (13)$$

其中 K 称为核矩阵, 元素 $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, 根据式 (13) 求得列向量 $\alpha^1, \dots, \alpha^d$, 对于任意测试样本 \mathbf{x} , 相对于 u^k 的投影结果为

$$(u^k \cdot \phi(\mathbf{x})) = \sum_{i=1}^n \alpha_i^k (\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i)) = \sum_{i=1}^n \alpha_i^k k(\mathbf{x}, \mathbf{x}_i)$$

将上式由不同 $k = \{1, \dots, d\}$ 值情况下所得的结果值串接, 就是最终的投影向量。

4 实验分析

分别采用 ORL 人脸库、自制小样本说话人语料库验证本文所提投影分类算法, 并与经典方法 LPP, LSDA, LDA, PCA 进行结果对比。实验过程分为 3 个步骤: 首先应用不同算法求取各训练样本的最佳变换矩阵, 然后将测试样本进行低维投影, 最后采用最近邻分类器得出辨认结果。实验过程中, BCLDP 算法中的 β 和 γ 都通过网格搜索获得, 其

中 β 取值范围为 (0, 0.5], 选取间隔为 0.05, γ 取值范围为 (0, 0.1], 选取间隔为 0.01, 具体取值根据不同实验各异。

4.1 ORL 人脸库实验

ORL 人脸库 (<http://www.cam-orl.co.uk>) 包括 40 类共 400 张人脸图像, 根据眼睛的位置配准图像, 然后将图像裁剪且缩放为 32×32 。从每类 10 张人脸图像中随机选取 $l = \{3, 5\}$ 张作为训练集, 其余作为测试集, 重复进行 30 次用以获得最终平均实验结果。图 2 是 ORL 人脸库的部分人脸样例。图 3 给出了 5 种算法及其核化版本在不同投影维度下取得的识别结果, 其中图 3(a), 3(b) 为线性算法, 选取平衡因子 $\beta = 0.4$, 热核参数 t 取训练样本间距离的平均值, 类内近邻数设为 $k_w = n_i - 1$, 其中 n_i 为第 i 子分块的同类样本数, 类间近邻数 $k_b = 5$ 。图 3(c), 3(d) 为相应的核化算法, 核函数取应用最为广泛的高斯核, 参数 $\sigma = 2000$ 。表 1 列出了各算法的最佳识别率及其维度, 其中加粗数据是相同实验条件下的最优识别率。从中可见, 较之其它各种算法, BCLDP 具有更高的识别率。



图 2 ORL 人脸库部分样本示例

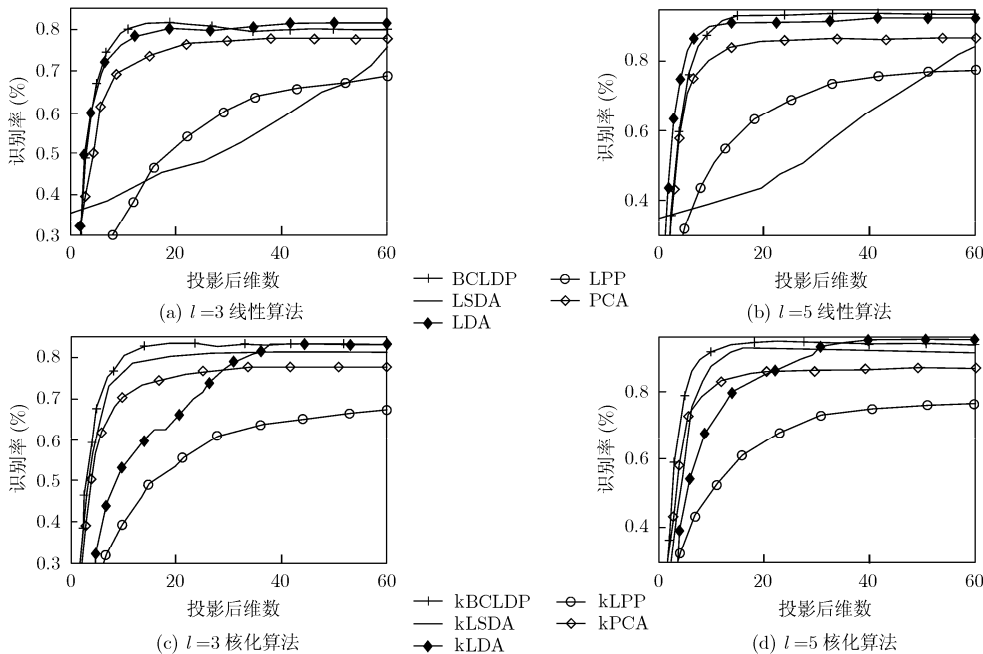


图 3 ORL 人脸库中不同维度时 5 种算法识别率对比

表1 各种线性投影算法及其核化版本在 ORL 人脸库中取得的最佳识别率及相应维度

方法	$l=3$ (线性 投影)	$l=3$ (核化 版本)	$l=5$ (线性 投影)	$l=5$ (核化 版本)
PCA	0.7796(56)	0.7786(36)	0.8715(57)	0.8707(56)
LDA	0.8163(39)	0.8341(39)	0.9255(39)	0.9463(39)
LPP	0.7373(60)	0.6024(65)	0.7065(59)	0.7647(60)
LSDA	0.8045(87)	0.8255(87)	0.8820(74)	0.9237(18)
BCLDP	0.8177(14)	0.8345(19)	0.9357(51)	0.9467(22)

4.2 说话人辨认实验

自制小样本说话人语料库，总人数 10 个，其中男 5 个，女 5 个。数据通过采样频率 8000 Hz，量化位数 16 bit，单声道 A/D 转化获得。每个人的语音信号通过不同时期录制合成。每人混合提取不同时期的语音片段总长度 15 s 作为训练信号，不同时期的 20 个长度 1.5 s 语音片段作为测试信号，即 20 个训练语音，200 个测试语音。语音信号先经高频提升，中心削减等预处理，再通过声音活性检测技术提取其中有效的语音段，以 30 ms 为长度分帧提取 12 维的 MFCC (Mel Frequency Cepstral Coefficient) 特征参数并作一阶差分，共计 24 维特征向量，作为分类参数。表 2 给出了 BCLDP 算法及其核化版本分别在 5 人与 10 人情况下的说话人辨认结果，实验中选取平衡因子 $\beta = 0.3$ ，热核参数 $t = 8$ ，高斯核参数 $\sigma = 28$ ，类内类间近邻数统一选为 $k_w = k_b = 7$ ，LDA 算法受类间散度矩阵秩的影响，子空间维数为 9，其余各算法的子空间维数统一选为 15。

表2 各种线性投影算法及其核化版本在小样本说话人辨认实验的识别率对比

方法	5人(线性 投影)	5人(核化 版本)	10人(线 性投影)	10人(核 化版本)
PCA	0.67	0.90	0.585	0.68
LDA	0.80	0.99	0.77	0.91
LPP	0.72	0.82	0.75	0.83
LSDA	0.68	0.99	0.56	0.865
BCLDP	0.80	0.99	0.745	0.915

4.3 分析讨论

综合上述所有实验结果，值得强调如下几点：

(1)LPP 从局部保持出发，由于缺少类别监督信息，性能较弱；PCA 和 LDA 以全局最优投影为目标，性能较强，但却缺少局部特征保持性质；LSDA 折衷考虑类内近邻保持和类间散度增大，但缺少近邻间距离权值约束，性能仍不突出；BCLDP 综合考虑上述优缺点，整体实验效果优于其它几种算法。

(2)总体来说，基于核函数的投影算法较同类线性算法性能更优，且效果更为稳定，以 LSDA 在 ORL 人脸库中的实验为例，线性 LSDA 在不同维度下识别率变化较大，而 kLSDA 的识别率在很大范围内的维度情况下都具有近似最优的结果。值得一提的是，BCLDP 和 kBCLDP 在 ORL 实验中效果接近，并无多大差异，而在说话人辨认应用时，则是 kBCLDP 远远优于 BCLDP。原因如下：基于核函数的算法出发点是将低维数据映射至高维空间以增加类间间隔，使得在特征空间中可以用常规线性方法求解，而本文算法输入图像数据已经是千维以上向量，本身已具有相当的线性可分离度，因此无法显示核函数的优势，但是语音特征参数仅为 24 维，采用核函数后大大地提升了类间散度，与实验结果相符合。

(3)投影维度对于分类性能影响较大，因此算法往往要根据具体应用对象选择最佳的维数。LDA 的降维范围受到类间散度矩阵秩的影响，最高仅能取 $c-1$ ， c 为类别数；LSDA 的最佳投影维度较难选择，虽然图 3 中仅给出了前 60 维的投影结果，但其最佳维度却仍在此范围之外；相比较而言，BCLDP 的维度选择更为自由，无需 LDA 那样受限制，也不像 LSDA 那样难以确定。

5 结束语

基于组合优化策略，本文提出了正交型局部判别投影法，在每一个子分块中最大化异类近邻样本间隔的同时保持类内近邻结构，通过整体组合后完成目标函数构建，在 ORL 人脸库和小样本语音库上的实验结果表明，BCLDP 优于传统的几种投影法，包括 PCA, LDA, LPP, LSDA 等。另外，采用核函数将 BCLDP 扩展为非线性方法，能够更好地反映数据间的非线性结构，在语音特征参数这类低维样本上具有突出的分类效果，说话人辨认结果也验证了此论点。

参考文献

- [1] Wen Jin-huan, Zheng Tian, She Hong-wei, *et al.*. Feature extraction of hyperspectral images based on preserving neighborhood discriminant embedding. International Conference on Image Analysis and Signal Processing (IASP), Zhejiang, 2010: 257-262.
- [2] Kumar V P, Kishore K V K, and Kumar K H. Hybrid face recognition with locally discriminating projection. International Conference on Signal Acquisition and Processing, Bangalore, India, 2010: 327-331.
- [3] He Xiao-fei, Cai Deng, Yan Shui-cheng, *et al.*. Neighborhood preserving embedding. Tenth IEEE International Conference Computer Vision, Beijing, 2005: 1208-1213.

- [4] He Xiao-fei, Yan Shui-cheng, Hu Yu-xiao, *et al.*. Face recognition using laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(3): 328-340.
- [5] Kokiopoulou E and Saad Y. Orthogonal neighborhood preserving projections: a projection-based dimensionality reduction technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(12): 2143-2156.
- [6] 申中华, 潘永惠, 王士同. 有监督的局部保留投影降维算法. 模式识别与人工智能, 2008, 21(2): 233-239.
Shen Zhong-hua, Pan Yong-hui, and Wang Shi-tong. A supervised locality preserving projection algorithm for dimensionality reduction. *Pattern Recognition and Artificial Intelligence*, 2008, 21(2): 233-239.
- [7] Cai Deng, He Xiao-fei, and Zhou Kun. Locality sensitive discriminant analysis. International Joint Conferences on Artificial Intelligence, Hyderabad, India, 2007: 708-713.
- [8] Wei B and Tao D C. Biased discriminant euclidean embedding for content-based image retrieval. *IEEE Transactions on Image Processing*, 2010, 19(2): 545-554.
- [9] Cai H, Mikolajczyk K, and Matas J. Learning linear discriminant projections for dimensionality reduction of image descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(2): 338-352.
- [10] Chen Yu, Zhang Xin, Zhang Wei-feng, *et al.*. Maximum margin subspace projections for face recognition. International Conference on Innovative Computing and Communication, Macao, 2010: 221-224.
- [11] Tenenbaum J B, Silva V de, and Langford J C. A global geometric framework for nonlinear dimensionality reduction. *Science*, 2000, 290(22): 2319-2323.
- [12] Roweis S T and Saul L K. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, 290(22): 2323-2326.
- [13] Belkin M and Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 2003, 15(6): 1373-1396.
- [14] Zhang T, Tao D, Li X, *et al.*. A unifying framework for spectral analysis based dimensionality reduction. in Proc. International Joint Conference on Neural Network, 2008: 1670-1677.
- 郑建炜: 男, 1982 年生, 讲师, 研究方向为模式识别与智能算法.
王万良: 男, 1957 年生, 教授, 研究方向为智能控制与网络控制.
姚信威: 男, 1986 年生, 博士生, 研究方向为网络流媒体传输.