

# 基于大间距准则和图像矩阵双向投影的人脸特征提取方法

詹宇斌<sup>1</sup> 殷建平<sup>1</sup> 刘新旺<sup>1</sup>

**摘要** 传统基于降维技术的人脸特征提取需要将图像转换成更高维的向量, 从而加剧维数灾难问题, 对于采用 Fisher 优化准则的特征提取, 这也会使小样本问题更加突出. 基于图像的矩阵表示, 本文提出了一种新的基于大间距准则和矩阵双向投影技术的人脸特征提取方法 (Maximum margin criterion and image matrix bidirectional projection, MMC-MBP). 该方法一方面在计算散度矩阵时引入了能保持数据局部性的 Laplacian 矩阵, 以保持数据的流形结构, 从而提高识别正确率; 另一方面采用了有效且稳定的大间距的优化准则即最大化矩阵迹差, 能克服利用 Fisher 准则所带来的小样本问题; 更重要的, MMC-MBP 方法给出了求解最优双向投影矩阵的迭代计算过程, 该迭代求解过程能保证目标函数的单调递增性、收敛性以及投影矩阵的收敛性, 从而成功解决了传统基于张量 (矩阵) 投影技术的特征提取方法特征维数过高或者无收敛解的问题. 最后广泛而系统的人脸识别实验表明, MMC-MBP 的迭代求解过程能很快收敛, 且相比 Eigenfaces, Fisherfaces, Laplacianfaces 等脸识别方法, 具有更高的识别正确率, 是一种有效的人脸特征提取方法.

**关键词** 特征提取, 维数约简, 大间距, 人脸识别

**DOI** 10.3724/SP.J.1004.2010.01645

## Face Feature Extraction Based on Maximum Margin Criterion and Image Matrix Bidirectional Projection

ZHAN Yu-Bin<sup>1</sup> YIN Jian-Ping<sup>1</sup> LIU Xin-Wang<sup>1</sup>

**Abstract** Conventional appearance-based face recognition needs to convert the image matrix into a higher dimensional vector. This will aggravate “the curse of dimensionality” problem and small sample size problem for those who adopt Fisher discriminant criterion. Moreover, existing feature extractions based on tensor (matrix) representation of image encounter the problem that dimensionality of extracted feature is still too high or there is no convergent solution. Based on matrix representation of image and matrix bidirectional projection, a novel feature extraction method called MMC-MBP (maximum margin criterion and image matrix bidirectional projection) for face recognition is proposed. On the one hand, since the Laplacian matrix, which can preserve the locality of the data set, has been introduced into the scatter matrix, MMC-MBP can preserve the intrinsic manifold structure of face set. On the other hand, MMC-MBP adopts an efficient and stable maximum margin criterion, so it can conquer the small sample size problem brought by Fisher discriminant criterion. What is more, an iterative procedure, which is proven to be convergent and can monotonously increase the objective value, is devised to compute the optimal projection matrices. Thus, the proposed MMC-MBP approach can conquer the problem of existing tensor (matrix) projection based feature extraction approaches. Extensive and systematic experiments on CMU PIE face database and Yale face database verify the high convergence speed of the iterative solution procedure and demonstrate that the proposed MMC-MBP outperforms the state-of-the-art methods in face feature extraction such as Eigenfaces, Fisherfaces, and Laplacianfaces.

**Key words** Feature extraction, dimensionality reduction, maximum margin, face recognition

自动人脸识别一直是计算机视觉和模式识别领域极具吸引力和挑战性的研究课题. 由于人脸数据的高维性, 人脸特征提取对自动人脸识别至关重要. 经过十几年的研究, 人们提出了许多基于降维技术的人脸特征提取方法, 如主成分分析 (Principal component analysis, PCA)<sup>[1]</sup>, 线性判别分析

(Linear discriminant analysis, LDA)<sup>[2]</sup> 等. PCA 是一种无监督的降维技术, 而 LDA 则利用样本的标号信息指导降维, 因此具有更好的分类性能. 由于 LDA 采用 Fisher 判别准则, 需要计算类内散度矩阵的逆, 而较小的样本数将导致奇异的类内散度矩阵, 因此 LDA 不能直接应用于小样本问题 (Small sample size problem, SSS)<sup>[3]</sup>. 人们提出了许多方法来克服 LDA 的小样本问题, 如 PCA+LDA<sup>[4]</sup>、正则化判别分析 (Regularized discriminant analysis, RDA)<sup>[5]</sup> 等. 最大间距准则 (Maximum margin criterion, MMC)<sup>[6]</sup> 是另一种有效且稳定的线性降维技术, 由于采用类间散度和类内散度之差也即间距

收稿日期 2010-01-17 录用日期 2010-05-05  
Manuscript received January 17, 2010; accepted May 5, 2010  
国家自然科学基金 (60970034, 60603015), 高等学校全国优秀博士学位论文作者资助项目 (2007B4) 资助  
Supported by National Natural Science Foundation of China (60970034, 60603015), and the Foundation for Author of National Excellent Doctoral Dissertation (2007B4)

1. 中国人民解放军国防科学技术大学计算机学院 长沙 410073  
1. School of Computer, National University of Defense Technology, Changsha 410073

作为判别准则,避免了矩阵求逆,能有效克服 LDA 存在的小样本问题.几何上,MMC 通过最大化不同类样本的平均间距来实现降维.

最近有关研究表明<sup>[7-9]</sup>,人脸图像可能位于高维图像空间的低维流形之上.基于流形假设,人们相继提出了多种流形学习方法,如等度映射(Isomap)<sup>[7]</sup>、局部线性嵌入(Locally linear embedding, LLE)<sup>[8]</sup>及 Laplacian 特征映射<sup>[9]</sup>等.由于这些流形学习算法没有显式地处理未见样本的机制,无法在人脸识别中展开应用.人们又提出了这些流形学习方法的线性版本,如局部保持投影(Locality preserving projection, LPP)<sup>[10]</sup>、邻域保持嵌入(Neighborhood preserving embedding, NPE)<sup>[11]</sup>等,其中 He 等提出的 LPP 算法通过保持数据的局部性来实现降维,随后他又提出了利用 LPP 进行人脸特征提取的 Laplacianfaces 方法<sup>[12]</sup>,取得了不错的识别效果.

在以上的特征提取方法中,都需要将大小为  $m \times n$  的人脸图像转换为  $m \times n$  维的向量.这一方面会加剧维数灾难问题,使得采用 Fisher 判别准则的方法如 LDA 所面临的小样本问题更加突出;另一方面会丢失图像中有用的结构化信息.因此人们开始考虑直接利用矩阵或张量的人脸表示,来实现特征提取,如 2DPCA<sup>[13]</sup>、2D-LDA<sup>[14]</sup>及文颖等提出的共同向量结合 2DPCA 的方法<sup>[15]</sup>等.实践也证明,这些方法能获得较 1D 方法更好的性能.但是当采用单向的矩阵投影技术时,所获得的特征维数仍然比较高.鉴于此,Zou 等提出了利用矩阵行投影和列投影的双向主成分分析(Bidirectional PCA, BDPCA)<sup>[16]</sup>,来进一步降低特征维数.但由于分别求解最优行投影和列投影,该方法并不能获得最优解.

在张量表示方面,He 等提出的张量子空间分析(Tensor subspace analysis, TSA)采用 2 阶张量来表示图像<sup>[17]</sup>,通过局部保持来学习图像所在的低维张量子空间,并通过迭代的广义特征分解来寻求最优解,但是由于采用的是 Fisher 准则,该方法并不能保证得到收敛的最优解.针对一般张量分析迭代求解过程无收敛解的问题,Wang 等在每步迭代时通过将优化目标矩阵迹的比等价地转换为矩阵迹的差,给出了一个收敛求解过程,但本质上,其采用的仍是 Fisher 优化准则<sup>[18]</sup>.

针对现有的基于张量(矩阵)投影技术的人脸特征提取特征维数过高或无收敛解的问题,本文提出了一种新的基于大间距准则和矩阵双向投影的人脸特征提取方法(Maximum margin criterion and image matrix bidirectional projection, MMC-MBP).该方法直接基于图像矩阵的双向投影技术,采用有效且稳定的大间距判别准则,并通过在计算散度矩

阵时引入保持数据局部性的 Laplacian 矩阵,来实现特征提取的同时保持数据集的内在非线性结构.更重要的,本文还提出了一种迭代求解最优投影矩阵的方法,并从理论上证明了该迭代求解过程的收敛性以及目标函数值的单调递增性和收敛性,从而克服了现有基于张量(矩阵)投影的特征提取特征维数过高(如 2DPCA 和 2D-LDA)或者无收敛解的问题(如 TSA).最后在标准的人脸测试库上的广泛实验表明,本文提出的迭代求解过程能很快收敛,同时相比其他的人脸特征提取算法,本文方法能获得更高的识别正确率.

## 1 MMC-MBP 方法

### 1.1 Laplacian 散度矩阵

令矩阵  $x$  表示大小为  $m \times n$  的图像,传统的 1D 特征提取方法需要将  $x$  转化成  $m \times n$  的一维向量.MMC-MBP 直接采用矩阵来表示图像,基于矩阵双向投影实现特征提取,形式化表示如下:

$$y = U^T x V \quad (1)$$

其中,  $y$  为提取的特征矩阵,  $U$  和  $V$  分别是大小为  $m \times m'$  ( $m' \leq m$ ) 和  $n \times n'$  ( $n' \leq n$ ) 的列投影和行投影矩阵.

给定属于  $c$  个不同类的  $N$  个训练图像  $X = [x_1^{(1)}, \dots, x_{N_1}^{(1)}, \dots, x_j^{(i)}, \dots, x_{N_c}^{(c)}] = [X_1, X_2, \dots, X_c] = [x_1, x_2, \dots, x_N]$ , 其中,第  $i$  类包含  $N_i$  ( $\sum_{i=1}^c N_i = N$ ) 个图像,  $X_i = [x_1^{(i)}, x_2^{(i)}, \dots, x_{N_i}^{(i)}]$  为由属于第  $i$  类图像矩阵构成的矩阵.记  $y_i$  为利用式 (1) 进行特征提取后图像  $x_i$  的特征矩阵,相应地记  $Y = [Y_1, Y_2, \dots, Y_c] = [y_1, y_2, \dots, y_N]$ , 其中矩阵  $Y_i = [y_1^{(i)}, y_2^{(i)}, \dots, y_{N_i}^{(i)}]$  表示由属于第  $i$  类的特征矩阵构成的矩阵.

每个特征矩阵  $y_i$  可以看成由  $m'$  个行向量构成,则在特征空间中,图像集  $X$  的行间总体散度矩阵可以表示为

$$\begin{aligned} S_t^{\text{row}} &= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{m'} (\mathbf{y}_{ij} - \mathbf{y}^{(j)})^T (\mathbf{y}_{ij} - \mathbf{y}^{(j)}) = \\ &= \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^T (y_i - \bar{y}) = \\ &= \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N (y_i - y_j)^T (y_i - y_j) \quad (2) \end{aligned}$$

其中,  $\bar{y}$ ,  $\mathbf{y}_{ij}$ ,  $\mathbf{y}^{(j)}$  分别为特征矩阵的均值矩阵,矩阵  $y_i$  的第  $j$  行向量及特征矩阵第  $j$  行向量的均值向

量. 类似地, 行间类内散度矩阵可以表示为

$$S_w^{\text{row}} = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} (y_j^{(i)} - m_i)^T (y_j^{(i)} - m_i) = \frac{1}{N} \sum_{i=1}^c \frac{1}{2N_i} \sum_{j,k=1}^{N_i} (y_j^{(i)} - y_k^{(i)})^T (y_j^{(i)} - y_k^{(i)}) \quad (3)$$

其中,  $m_i$  为第  $i$  类图像的均值特征矩阵.

最近很多研究都表明人脸图像具有非线性的流形结构<sup>[7-8]</sup>, 保持人脸图像的这种内在非线性流形结构有助于提高人脸识别率<sup>[12, 19]</sup>. 根据图嵌入理论<sup>[20]</sup>, 可以用无向加权图  $G(X, W)$  来刻画数据集  $X$  的内在非线性流形结构.  $G$  中顶点表示图像矩阵, 实对称矩阵  $W$  则刻画顶点之间的相似性, 可以采取多种方式来构造矩阵  $W$ , 如 Laplacian 特征映射中的高斯相似性<sup>[9]</sup>, LLE 中的局部邻域关系<sup>[8]</sup> 等. 本文中采用高斯相似性:

$$w_{ij} = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4)$$

为了在特征提取时能尽量保持数据集的内在非线性流形结构, 在散度矩阵中引入图  $G$  的 Laplacian 矩阵, 定义行间总体 Laplacian 散度矩阵如下:

$$LS_t^{\text{row}} = \frac{1}{2N^2} \sum_{i,j=1}^N w_{ij} (y_i - y_j)^T (y_i - y_j) = \frac{1}{N^2} \sum_{i,j=1}^N (w_{ij} y_i^T y_i - w_{ij} y_i^T y_j) = \frac{1}{N^2} Y'^T (L \otimes I_{m'}) Y' \quad (5)$$

其中,  $Y' = [y_1^T, y_2^T, \dots, y_N^T]^T$ ,  $D$  为对角矩阵, 对角元素  $d_{ii} = \sum_{j=1}^N w_{ij}$ ,  $L = D - W$  为图  $G$  的 Laplacian 矩阵.  $I_{m'}$  为  $m' \times m'$  的单位矩阵, 而  $\otimes$  则表示矩阵的 Kronecher 积.

记  $w_{jk}^{(i)} = \exp(-\|x_j^{(i)} - x_k^{(i)}\|^2 / (2\sigma^2))$  表示第  $i$  类中第  $j$  和  $k$  个样本的类内相似性, 则类似地, 行间类内的 Laplacian 散度矩阵可以表示如下:

$$LS_w^{\text{row}} = \frac{1}{N} \sum_{i=1}^c \frac{1}{2N_i} \sum_{j,k=1}^{N_i} w_{jk}^{(i)} (y_j^{(i)} - y_k^{(i)})^T (y_j^{(i)} - y_k^{(i)}) = \frac{1}{N} \sum_{i=1}^c \frac{1}{N_i} Y_i'^T (L_w^{(i)} \otimes I_{m'}) Y_i' = \frac{1}{N} Y'^T (L_w \otimes I_{m'}) Y' \quad (6)$$

其中,  $Y_i' = [y_1^{(i)T}, y_2^{(i)T}, \dots, y_{N_i}^{(i)T}]^T$ ,  $L_w^{(i)}$  则是  $i$  类样本的 Laplacian 矩阵,  $L_w = \text{diag}\{L_w^{(1)}/N_1, \dots, L_w^{(c)}/N_c\}$  为块对角矩阵. 这样行间类间 Laplacian 散度矩阵可以如下计算:

$$LS_b^{\text{row}} = LS_t^{\text{row}} - LS_w^{\text{row}} \quad (7)$$

同样地, 每个特征矩阵  $y_i$  也可以看成是由  $n'$  个列向量构成, 则类似于行方向的散度矩阵定义, 在特征空间中, 列间的散度矩阵定义如下:

$$S_t^{\text{col}} = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})(y_i - \bar{y})^T = \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N (y_i - y_j)(y_i - y_j)^T \quad (8)$$

$$S_w^{\text{col}} = \frac{1}{N} \sum_{i=1}^c \frac{1}{2N_i} \sum_{j,k=1}^{N_i} (y_j^{(i)} - y_k^{(i)})(y_j^{(i)} - y_k^{(i)})^T \quad (9)$$

因此, 可以获得如下形式表达的列间 Laplacian 散度矩阵:

$$LS_t^{\text{col}} = \frac{1}{N^2} Y (L \otimes I_{n'}) Y^T \quad (10)$$

$$LS_w^{\text{col}} = \frac{1}{N} Y (L_w \otimes I_{n'}) Y^T \quad (11)$$

$$LS_b^{\text{col}} = LS_t^{\text{col}} - LS_w^{\text{col}} \quad (12)$$

## 1.2 优化目标

在 MMC<sup>[6]</sup> 中, 类间平均间距用下式描述:

$$J = \text{tr}(S_b - S_w) = \text{tr}(S_t - 2S_w) \quad (13)$$

MMC 通过最大化特征空间中该类间平均间距实现特征提取. 由于采用的矩阵表示图像, MMC-MBP 所提取的特征为特征矩阵, 因此可以分别定义行间及列间 Laplacian 类间距如下:

$$J_{\text{row}} = \text{tr}(LS_b^{\text{row}} - LS_w^{\text{row}}) \quad (14)$$

$$J_{\text{col}} = \text{tr}(LS_b^{\text{col}} - LS_w^{\text{col}}) \quad (15)$$

在特征空间中, 上述间距越大, 来自不同类的样本就会间隔越远, 而同类的样本则会越接近. 因此特征提取的目标将是寻找最优投影矩阵  $U$  和  $V$  来最大化行间 Laplacian 类间距与列间 Laplacian 类间距之和, 以最大化不同类样本的可分性. 而按照特征提取式 (1), 每个图像的特征矩阵同时依赖于行投影矩阵  $U$  和列投影矩阵  $V$ , 因此在特征空间中, 相

应的 Laplacian 散度矩阵也同时依赖于投影矩阵  $U$  和  $V$ .

利用特征提取式 (1), 将特征空间中散度矩阵利用原始图像矩阵及投影矩阵  $U$  和  $V$  表示如下:

$$LS_t^{\text{row}} = \frac{1}{N^2} V^T X'^T (\mathbf{e} \otimes U) (L \otimes I_{m'}) (\mathbf{e} \otimes U)^T X' V \quad (16)$$

$$LS_w^{\text{row}} = \frac{1}{N} V^T X'^T (\mathbf{e} \otimes U) (L_w \otimes I_{m'}) (\mathbf{e} \otimes U)^T X' V \quad (17)$$

$$LS_t^{\text{col}} = \frac{1}{N^2} U^T X (\mathbf{e} \otimes V) (L \otimes I_{n'}) (\mathbf{e} \otimes V)^T X^T U \quad (18)$$

$$LS_w^{\text{col}} = \frac{1}{N} U^T X (\mathbf{e} \otimes V) (L_w \otimes I_{n'}) (\mathbf{e} \otimes V)^T X^T U \quad (19)$$

其中,  $\mathbf{e}$  为合适维数的全 1 列向量.

因此特征空间中, 行间和列间 Laplacian 间距分别为

$$J_{\text{row}}(U, V) = \text{tr}(LS_b^{\text{row}} - LS_w^{\text{row}}) = \text{tr}(V^T J_U V) \quad (20)$$

$$J_{\text{col}}(U, V) = \text{tr}(LS_b^{\text{col}} - LS_w^{\text{col}}) = \text{tr}(U^T J_V U) \quad (21)$$

其中

$$J_U = \frac{1}{N^2} X'^T (\mathbf{e} \otimes U) (L \otimes I_{m'}) (\mathbf{e} \otimes U)^T X' - \frac{2}{N} X'^T (\mathbf{e} \otimes U) (L_w \otimes I_{m'}) (\mathbf{e} \otimes U)^T X' \quad (22)$$

仅仅依赖于  $U$ , 而

$$J_V = \frac{1}{N^2} X (\mathbf{e} \otimes V) (L \otimes I_{n'}) (\mathbf{e} \otimes V)^T X^T - \frac{2}{N} X (\mathbf{e} \otimes V) (L_w \otimes I_{n'}) (\mathbf{e} \otimes V)^T X^T \quad (23)$$

依赖于  $V$ . 因此 MMC-MBP 优化目标为

$$\max: J(U, V) = J_{\text{row}}(U, V) + J_{\text{col}}(U, V) \quad (24)$$

对投影矩阵  $U$  和  $V$  强加列正交性约束, 则最优  $U$  和  $V$  可以通过求解如下的优化问题获得:

$$\begin{aligned} \max: & J(U, V) = J_{\text{row}}(U, V) + J_{\text{col}}(U, V) \\ \text{s.t.} & U^T U = I_{m'}, V^T V = I_{n'} \end{aligned} \quad (25)$$

### 1.3 优化问题求解

对于优化问题 (25), 不存在显式的解析解, 但是根据 Laplacian 散度矩阵的定义式 (16)~(19), 很容易验证得到:

$$J_{\text{row}}(U, V) = J_{\text{col}}(U, V) \quad (26)$$

根据式 (26), 优化目标可以进一步写成:

$$J(U, V) = 2J_{\text{row}}(U, V) = 2J_{\text{col}}(U, V) \quad (27)$$

从而可以通过分别最大化  $J_{\text{row}}(U, V)$  和  $J_{\text{col}}(U, V)$  来最大化  $J(U, V)$ . 这启发我们采用迭代方法计算最优的投影矩阵  $U$  和  $V$ . 对于给定的  $U$ , 最优  $V$  可以很容易地通过最大化  $J_{\text{row}}(U, V)$  得到; 反过来, 固定  $V$ , 最优  $U$  也可以很容易地通过最大化  $J_{\text{col}}(U, V)$  得到, 而这两个过程都能够在固定一个变量条件下, 最大化目标函数  $J(U, V)$ . 下面给出求解最优  $U$  和  $V$  详细迭代过程.

#### 算法 1. 迭代求解过程

1) 初始化: 用任意  $m \times m'$  列正交矩阵初始化  $U_0$ .

2) 迭代优化:

For  $k = 1, 2, \dots, k_{\max}$ , do

a) 按照式 (22) 计算  $J_{U_{k-1}}$

b) 求解如下的特征分解问题:

$$J_{U_{k-1}} \mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1, 2, \dots, n' \quad (28)$$

其中,  $\mathbf{v}_i$  为第  $i$  大特征值  $\lambda_i$  所对应的特征向量.

c) 令  $\bar{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n'}]$ .

d) 为保持解的正交变换不变性, 令  $S = \bar{V} \bar{V}^T X'^T X' \bar{V} \bar{V}^T$ , 求解如下的特征分解问题:

$$S \mathbf{v}_i = \gamma_i \mathbf{v}_i \quad (29)$$

并令  $V_k = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n'}]$ , 其中,  $\mathbf{v}_i$  为第  $i$  大特征值  $\gamma_i$  所对应的特征向量.

e) 按照式 (31) 计算  $J_{V_k}$ .

f) 求解如下的特征分解问题:

$$J_{V_k} \mathbf{u}_j = \lambda'_j \mathbf{u}_j, \quad j = 1, 2, \dots, m' \quad (30)$$

其中,  $\mathbf{u}_j$  为第  $j$  大特征值  $\lambda'_j$  所对应的特征向量.

g) 令  $\bar{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{m'}]$ .

h) 为保持解的正交变换不变性, 令  $S = \bar{U} \bar{U}^T X X^T \bar{U} \bar{U}^T$ , 求解如下的特征分解问题:

$$S \mathbf{u}_i = \gamma_i \mathbf{u}_i \quad (31)$$

并令  $U_k = [\mathbf{u}_1, \dots, \mathbf{u}_{m'}]$ , 其中,  $\mathbf{u}_i$  为第  $i$  大特征值  $\gamma_i$  所对应的特征向量.

if  $\|U_k - U_{k-1}\| < \sqrt{m'}\varepsilon$  且  $\|V_k - V_{k-1}\| < \sqrt{n'}\varepsilon$  (实验中, 取  $\varepsilon = 0.01$ ), 则返回.

3) 输出投影矩阵:  $U = U_k, V = V_k$ .

### 1.4 解的收敛性

尽管算法 1 采用迭代优化过程对  $U$  和  $V$  进行联合优化, 但是定理 1 表明, 该迭代过程能单调增加目标函数值. 这一性质从理论上保证了本文方法优于现有基于矩阵投影和张量表示的特征提取方法.

**定理 1.** 目标函数值满足如下不等式:

$$J(U_k, V_k) \leq J(U_k, V_{k+1}) \leq J(U_{k+1}, V_{k+1}) \quad (32)$$

**证明.** 根据算法 1 的求解过程,  $U_k$  和  $V_k$  满足:

$$\text{tr}(V_k^T J_{U_{k-1}} V_k) = \max_{V^T V = I_{n'}} \text{tr}(V^T J_{U_{k-1}} V) \quad (33)$$

$$\text{tr}(U_k^T J_{V_k} U_k) = \max_{U^T U = I_m} \text{tr}(U^T J_{V_k} U) \quad (34)$$

利用式 (26), 可以得到:

$$\begin{aligned} J_{\text{col}}(U_k, V_k) &= J_{\text{row}}(U_k, V_k) \leq J_{\text{row}}(U_k, V_{k+1}) = \\ &= J_{\text{col}}(U_k, V_{k+1}) \leq J_{\text{col}}(U_{k+1}, V_{k+1}) \end{aligned} \quad (35)$$

因此按照式 (27), 目标函数满足:

$$J(U_k, V_k) \leq J(U_k, V_{k+1}) \leq J(U_{k+1}, V_{k+1})$$

这就是所要证明的结论.  $\square$

定理 1 表明, 目标函数值随迭代次数  $k$  单调递增. 另外一个需要注意的事实是目标函数在优化问题 (25) 的约束条件下具有上界. 因此算法 1 的迭代过程能保证目标函数单调收敛. 进一步, 可以得到如下关于投影矩阵收敛性的定理:

**定理 2.** 当目标函数收敛时, 由算法 1 获得的投影矩阵序列  $\{(U_k, V_k) (k \geq 1)\}$  也收敛.

**证明.** 根据定理 1, 假设经过  $K$  次迭代, 目标函数值收敛, 于是有:

$$J(U_K, V_K) = J(U_K, V_{K+1}) = J(U_{K+1}, V_{K+1}) \quad (36)$$

将式 (27) 带入到式 (36), 同时利用式 (20) 和 (21), 得到:

$$\text{tr}(V_K^T J_{U_K} V_K) = \text{tr}(V_{K+1}^T J_{U_K} V_{K+1}) \quad (37)$$

于是  $\exists$  正交矩阵  $Q$  满足:  $V_{K+1} = V_K Q$ .

因此,  $V_{K+1}$  和  $V_K$  具有相同的列空间. 同时算法 1 的步骤 2d), 已经限定  $V_K$  和  $V_{K+1}$  为样本在该列空间中投影后的主特征方向, 因此可以得到:

$$V_{K+1} = V_K \quad (38)$$

类似地, 可以得到  $U_{K+1} = U_K$ . 这意味着当目标函数收敛时, 投影矩阵序列也收敛.  $\square$

### 1.5 讨论

与本文类似的工作是 Yang 等提出的 LBMMC (Laplacian bidirectional maximum margin criterion) 方法<sup>[21]</sup>. 尽管本文方法在思想上与 LBMMC 方法类似, 即都是采用大间距的判别准则, 通过引入 Laplacian 矩阵来显式地考虑人脸的流形结构. 但是在导出的优化目标以及优化问题求解上, 本文的 MMC-BLP 方法与 LBMMC 有着本质的区别. 在 LBMMC 中, 在定义行间 Laplacian 间距时, 只考虑了行矩阵的作用而忽略了列投影矩阵, 而在定义列间 Laplacian 间距时, 又仅仅只考虑了列投影矩阵而忽略了行投影矩阵. 而事实上, 根据特征提取式 (1), 在特征空间中, 无论是行间 Laplacian 间距, 还是列间 Laplacian 间距, 都同时依赖于行投影矩阵  $V$  和列投影矩阵  $U$ . LBMMC 的这种考虑其实是为了简化问题求解而放松了优化目标. 这种放松使所获得的投影矩阵并不是严格意义上的最优解. 这也是大多数基于张量 (矩阵) 投影技术的特征提取方法所面临的共同问题<sup>[19]</sup>. 而本文的 MMC-BLP 方法则是在特征空间中严格最大化行间 Laplacian 间距与列间 Laplacian 间距之和, 并采用迭代方式对行列投影矩阵进行联合优化, 并且从理论上证明了优化目标在迭代过程中的单调递增性和收敛性以及行列投影矩阵的收敛性. 另外, 虽然本文采用迭代方式进行求解, 但实验结果表明迭代过程能很快收敛, 算法的时间开销并没有显著增加.

## 2 实验结果及分析

本节首先通过实验研究迭代求解过程的收敛性, 然后通过人脸识别实验比较 MMC-MBP 与其他方法的性能. 参与比较的方法包括: Baseline (直接在原始高维空间利用最近邻分类器进行识别), Eigenfaces<sup>[22]</sup>, Fisherfaces<sup>[4]</sup>, MMC<sup>[6]</sup>, Laplacianfaces<sup>[12]</sup>, 2DPCA<sup>[14]</sup>, 2DLDA<sup>[23]</sup>, 2DLaplacianfaces<sup>[24]</sup> 和 LBMMC<sup>[21]</sup>. 比较时, 首先利用各种特征提取方法通过训练样本集学习得到投影矩阵 (向量), 然后利用学习得到的投影矩阵 (向量) 进行特征提取, 最后在特征空间利用基于余弦距离的最近邻分类器进行人脸识别. 对于 Fisherfaces 和 Laplacianfaces 在利用 PCA 预处理时, 保持 98% 的信息. 为了公平比较, 对于需要利用图模型来刻画数据流形结构的方法, 如 Laplacianfaces, LBMMC 和本文方法, 实验中都采用相同的加权图, 并利用文献 [25] 中方法来确定式 (7) 中参数  $\sigma$  值.

### 2.1 实验数据

实验所用的人脸数据为人脸识别领域广泛使

<sup>1</sup><http://www.zjucadcg.cn/dengcai/Data/FaceData.html>

用的卡内基梅隆大学的 PIE 人脸库<sup>[26]</sup>、Yale 人脸库<sup>[27]</sup>, 均从何的主页获取<sup>1</sup>. 每个人脸图像均已裁剪成大小为 32 像素  $\times$  32 像素, 并且已按照眼部位置对齐, 对于传统的基于向量的方法, 如 Eigenfaces, Fisherfeaces 等, 需将人脸图像转化为 1024 维的向量, 而对于 2D 方法和本文 MMC-MBP 方法, 直接利用 32 像素  $\times$  32 像素的矩阵来表示人脸图像.

CMU 的 PIE 人脸库: 该人脸库共有 41 368 幅图像, 包含了在 13 种不同姿态、43 种不同光照条件下以及 4 种不同表情下的 68 个人的人脸图像. 实验中我们选取包含所有表情和所有不同光照条件下的 5 个近似正面姿态的人脸图像构成实验数据 (C05, C07, C09, C27, C29), 这样实验数据中每个人有 170 幅图像. 实验中对每个人随机选取  $l$  ( $l = 5, 10, 20$ ) 幅图像作为训练集, 余下的作为测试集. 对每个  $l$ , 做 10 次随机划分实验.

Yale 人脸库: 该库含有在不同光照和表情下的 165 张灰度人脸图像 (15 人, 每人 11 张). 对于该数据库, 我们为每个人随机选择  $l$  ( $l = 2, 4, 6$ ) 幅图像进行训练, 余下的作为测试集. 对于每个不同的  $l$ , 做 10 次随机划分实验.

## 2.2 收敛性实验

实验中, 通过观测投影矩阵  $U$  和  $V$  相邻两次迭代之间的步差来研究投影矩阵的收敛性. 图 1 分别展示了投影矩阵的迭代步差, 以及目标函数值随迭代次数之间的变化关系. 从图中可以看到, 对于数据集 CMU PIE, 大致经过 10 次迭代, 目标函数值收敛于某一固定值; 对于 Yale 数据, 目标函数值经过 5 次左右迭代就收敛. 而当目标函数值收敛时, 投影矩阵也收敛. 因此可以得到结论: 本文提出的迭代求解过程不仅收敛, 而且能很快收敛. 高收敛速度能保证 MMC-MBP 方法快速得到优化的投影矩阵.

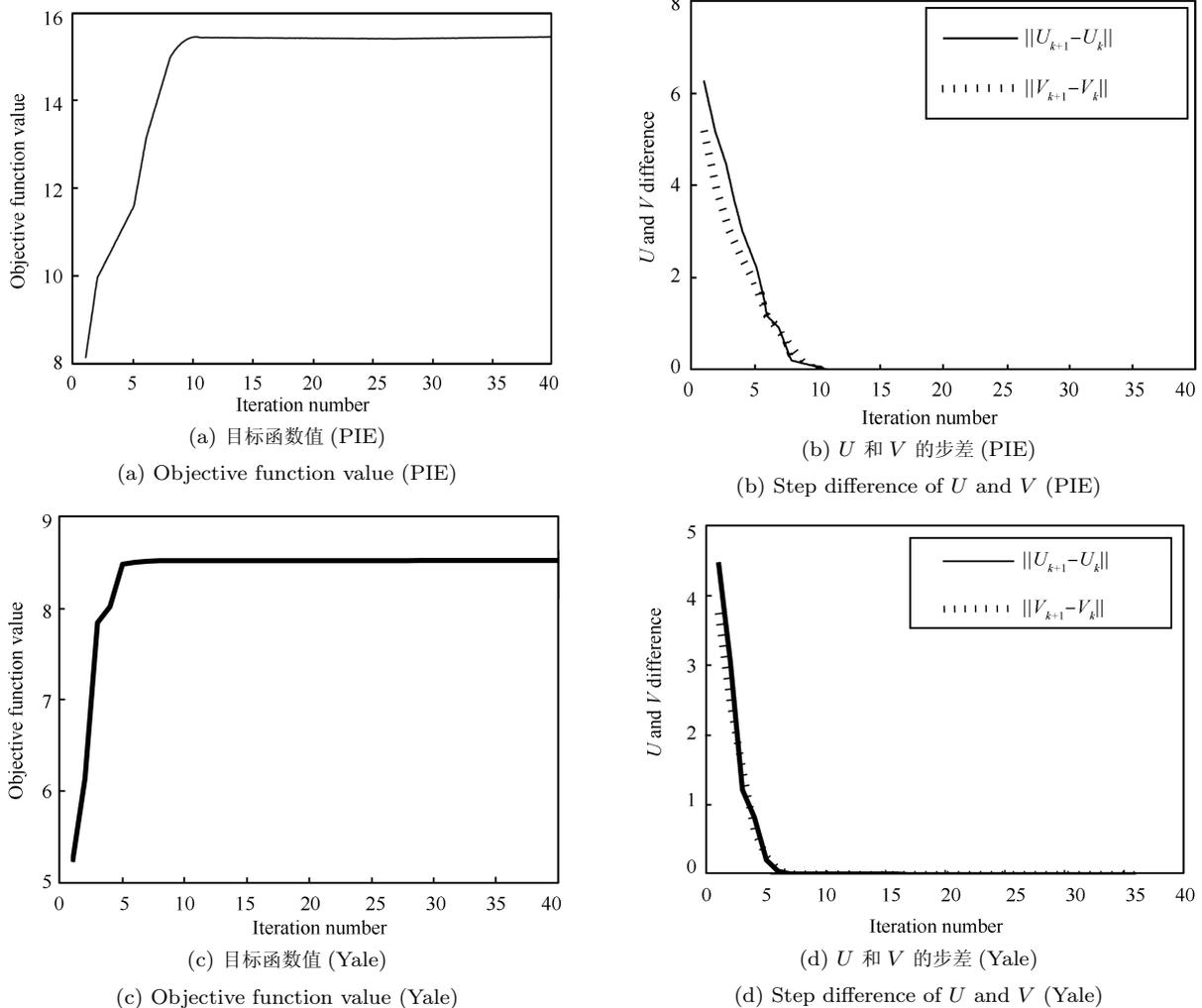


图 1 迭代过程的收敛性

Fig. 1 Convergence of the iterative procedure

### 2.3 性能比较

对每次随机划分实验, 我们都获得不同特征维数条件下的识别正确率, 然后计算不同特征维数条件下的平均识别正确率. 图 3 展示了在不同训练样本数条件下, 10 次随机划分实验所获得的平均识别率与维数的变化关系曲线. 需要注意的是, 对于 Fisherfaces<sup>[4]</sup> 方法, 由于类间散度矩阵秩最多为  $c - 1$  ( $c$  为类别数), 因此所获得的投影方向不超过  $c - 1$ . 对于 2DLDA<sup>[23]</sup>, LBMMC<sup>[21]</sup> 和本文方法 MMC-MBP, 由于采用双向矩阵投影, 我们只展示其特征维数为  $d \times d$  ( $d = 1, 2, \dots$ ), 即  $(1, 4, 9, \dots)$  的结果. 而对于 2DPCA<sup>[14]</sup> 和 2DLaplacianfaces<sup>[24]</sup>, 由于采用单向矩阵投影, 我们展示其维数为  $32 \times d$

( $d = 1, 2, \dots$ ), 即  $(32, 64, \dots)$  的结果. 而对于 MMC, 展示的是其特征维数为  $30 \times d$  ( $d = 1, 2, \dots$ ), 即  $(30, 60, 90, \dots)$  的结果.

表 1 比较了各种方法所获得的最高平均识别率及其相应的特征维数. 为了全面比较, 每次随机划分实验, 各种方法都可以获得一个最高识别率, 我们比较这个最高识别率的 10 次随机划分实验的均值和标准差. 表 2 比较了各种方法最高识别率的均值及其标准差. 在表 1 和表 2 中, 粗体表示最好的结果.

通过对图 2, 表 1 和表 2 的观察、比较和分析, 可以得到如下结论:

1) 在各种情况下, 本文的 MMC-MBP 方法都能获得一致较高的识别正确率, 表明 MMC-MBP 方法具有更好的人脸识别性能.

表 1 最高平均识别率 (%) (相应的特征维数) 比较

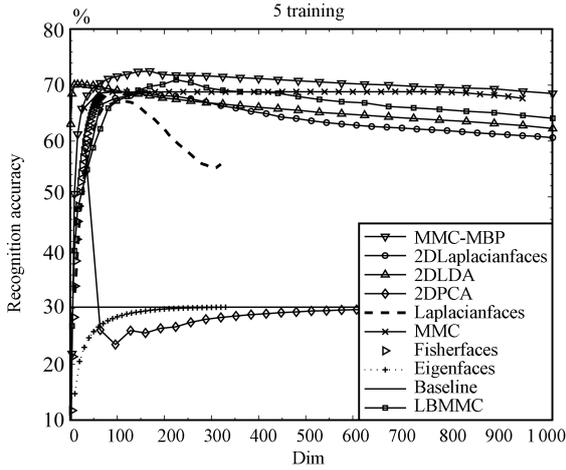
Table 1 Comparison of top average recognition accuracy (%) (its corresponding dimensionality) of different methods

方法 \ $l$ 值	PIE			Yale		
	5 train	10 train	20 train	2 train	4 train	6 train
Baseline	30.3	44.5	61.7	42.3	54.0	59.1
Eigenfaces	30.1 (330)	44.5 (650)	61.7 (1010)	42.3 (28)	54.8 (40)	60.0 (30)
Fisherfaces	68.0 (66)	73.2 (66)	78.6 (66)	44.5 (10)	72.7 (12)	79.9 (14)
MMC	68.8 (690)	84.6 (510)	91.9 (930)	49.2 (150)	68.7 (360)	75.7 (300)
Laplacianfaces	67.1 (110)	75.9 (144)	85.8 (168)	50.0 (14)	73.9 (22)	80.9 (20)
2DPCA	58.8 ( $32 \times 1$ )	71.6 ( $32 \times 1$ )	82.0 ( $32 \times 1$ )	47.6 ( $32 \times 4$ )	58.2 ( $32 \times 4$ )	65.9 ( $32 \times 5$ )
2DLDA	70.2 ( $5^2$ )	83.0 ( $7^2$ )	90.0 ( $7^2$ )	46.8 ( $3^2$ )	69.0 ( $3^2$ )	75.1 ( $3^2$ )
2DLaplacianfaces	68.6 ( $32 \times 5$ )	76.6 ( $32 \times 5$ )	85.5 ( $32 \times 8$ )	56.4 ( $32 \times 3$ )	70.8 ( $32 \times 4$ )	74.4 ( $32 \times 4$ )
LBMMC	71.0 ( $15 \times 15$ )	85.6 ( $15 \times 15$ )	91.7 ( $15 \times 15$ )	56.1 ( $7 \times 7$ )	71.5 ( $8 \times 8$ )	77.1 ( $9 \times 9$ )
MMC-MBP	<b>74.2</b> ( $9^2$ )	<b>86.2</b> ( $17^2$ )	<b>92.3</b> ( $20^2$ )	<b>57.8</b> ( $9^2$ )	<b>74.5</b> ( $8^2$ )	<b>81.4</b> ( $10^2$ )

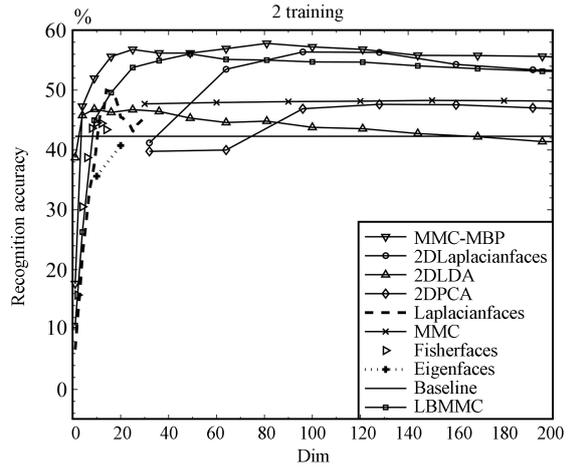
表 2 最高识别率的均值 (标准差比较) (%)

Table 2 Comparison of mean (standard derivations) of top recognition accuracy of different methods

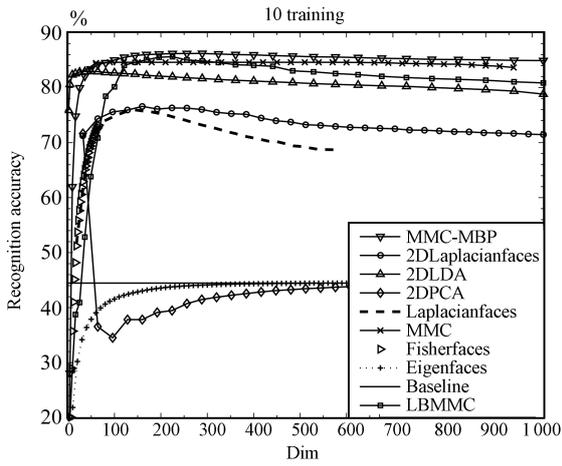
方法 \ $l$ 值	PIE			Yale		
	5 train	10 train	20 train	2 train	4 train	6 train
Baseline	30.1 (0.68)	44.5 (1.0)	61.7 (0.64)	42.3 (4.3)	54.0 (3.7)	59.1 (4.7)
Eigenfaces	30.2 (0.67)	44.5 (1.0)	61.7 (0.64)	42.3 (4.4)	55.1 (3.6)	61.6 (6.4)
Fisherfaces	68.0 (1.0)	73.2 (0.77)	78.6 (0.51)	46.5 (4.6)	73.6 (5.2)	80.4 (3.7)
MMC	68.9 (1.2)	84.7 (1.1)	92.0 (0.55)	49.4 (6.6)	69.1 (5.3)	76.3 (4.3)
Laplacianfaces	67.3 (0.84)	76.0 (0.68)	86.0 (0.45)	51.5 (5.5)	75.1 (5.2)	82.7 (2.8)
2DPCA	58.8 (0.96)	71.6 (0.97)	82.0 (0.56)	49.2 (3.8)	59.2 (3.8)	67.3 (5.2)
2DLDA	70.4 (1.1)	83.0 (0.75)	90.1 (0.71)	49.0 (6.8)	70.5 (4.5)	76.1 (2.7)
2DLaplacianfaces	68.8 (0.98)	76.8 (0.69)	85.7 (0.39)	57.3 (6.4)	72.1 (5.3)	76.0 (4.8)
LBMMC	71.2 (1.1)	85.9 (1.1)	92.0 (0.65)	56.6 (6.4)	71.9 (4.3)	77.2 (4.5)
MMC-MBP	<b>74.6</b> (1.0)	<b>86.3</b> (0.81)	<b>93.3</b> (0.33)	<b>59.1</b> (4.8)	<b>76.5</b> (3.9)	<b>83.4</b> (4.1)



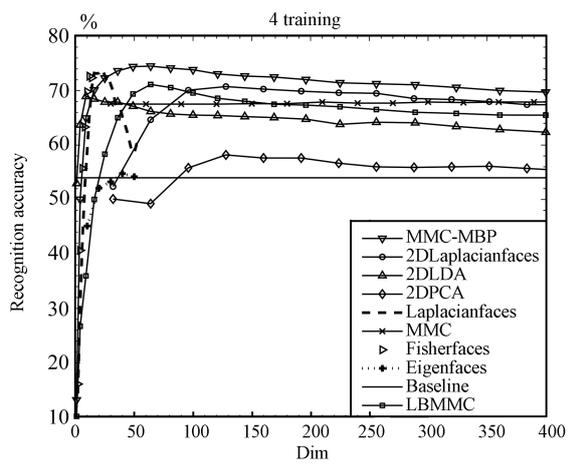
(a) 每个人 5 个训练样本 (PIE)  
 (a) 5 trainings per person (PIE)



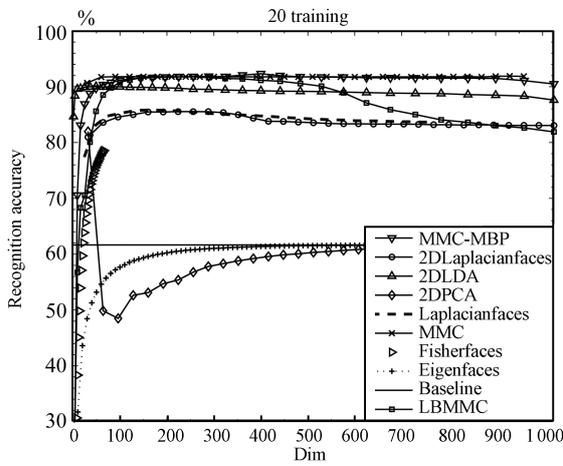
(b) 每个人 2 个训练样本 (Yale)  
 (b) 2 trainings per person (Yale)



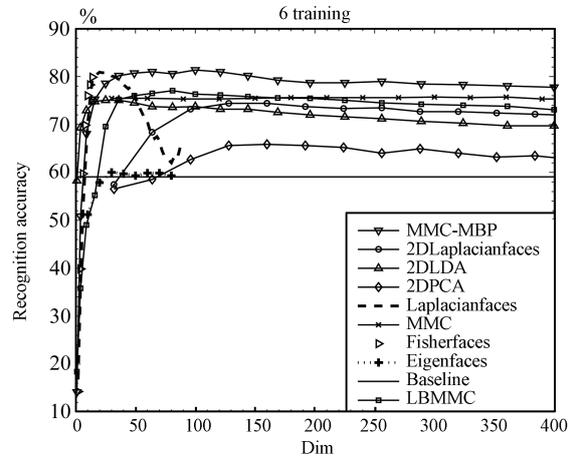
(c) 每个人 10 个训练样本 (PIE)  
 (c) 10 trainings per person (PIE)



(d) 每个人 4 个训练样本 (Yale)  
 (d) 4 trainings per person (Yale)



(e) 每个人 20 个训练样本 (PIE)  
 (e) 20 trainings per person (PIE)



(f) 每个人 6 个训练样本 (Yale)  
 (f) 6 trainings per person (Yale)

图 2 10 次随机划分实验的平均识别率 vs. 维数

Fig. 2 Average recognition accuracy vs. feature dimensionality over 10 random split experiments

2) 各种方法的正确识别率是都随特征维数的变化而变化, 但是对于大间距准则 MMC 方法, 其识别正确率随维数的变化较为平稳, 起伏不大, 这表明大间距准则是一种较为稳定的特征提取方法. 本文 MMC-MBP 方法和 LBMMC 方法都是采用大间距准则, 继承了 MMC 的稳定性, 相比采用 Fisher 准则的其他方法, 识别正确率随特征维数的变化较为平稳, 性能比较稳定.

3) 利用数据的流形结构信息, 能提高识别正确率. 特别是当训练样本数较小时, 这种提高相比较大的训练样本数目更为显著. 这可能是因为较小的训练样本数不足以刻画全局数据集的特点, 而通过引入流形结构信息, 能获取更多关于全局数据的信息.

### 3 结论

基于图像的矩阵表示, 本文提出了一种基于大间距和矩阵双向投影的人脸特征提取方法 (MMC-MBP). 传统的基于矩阵双向投影技术的人脸特征提取方法存在次优解<sup>[17, 21]</sup> 或者无收敛解的问题<sup>[18]</sup>. 本文提出的 MMC-MBP 方法由于采用大间距准则, 具有迭代的收敛解. 在 CMU 的 PIE 人脸库和 Yale 人脸库上系统的实验表明了该方法的有效性. 相比以前的特征提取方法, MMC-MBP 方法有以下 4 个主要特点:

1) 尽管 MMC-MBP 采用迭代方式求解, 但是迭代过程能保证目标函数的单调递增性、收敛性以及投影矩阵的收敛性, 而且实验表明, 算法经过很少次数的迭代就能收敛, 收敛速度快.

2) MMC-MBP 采用矩阵表示图像, 计算开销小. 在每次迭代过程中, 只需计算大小为  $m \times m$  和  $n \times n$  矩阵的广义特征分解问题, 而传统的基于向量的方法, 需计算大小为  $mn \times mn$  矩阵的特征分解问题. 同时基于矩阵的图像表示能保持被基于向量方法忽略的图像中的结构信息.

3) 在 MMC-MBP 中, 由于在计算散度矩阵时, 引入了刻画数据集局部性的 Laplacian 矩阵, 显式地考虑了人脸数据的内在非线性流形结构, 提高了识别正确率.

4) MMC-MBP 是一种较为稳定的特征提取方法. 由于采用了稳定的最大间距准则, MMC-MBP 继承了 MMC 的稳定性, 性能随特征维数的变化较为平稳.

### References

- Jolliffe I T. *Principal Component Analysis*. Berlin: Springer-Verlag, 2002
- Duda R O, Hart P E, Stork D G. *Pattern Classification*. New York: John Wiley and Sons, 2001
- Raudys S J, Jain A K. Small sample size effects in statistical pattern recognition: recommendations for practitioners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, **13**(3): 252–264
- Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(7): 711–720
- Friedman J H. Regularized discriminant analysis. *Journal of the American Statistical Association*, 1989, **84**(405): 165–175
- Li H F, Jiang T, Zhang K S. Efficient and robust feature extraction by maximum margin criterion. *IEEE Transactions on Neural Networks*, 2006, **17**(1): 157–165
- Tenenbaum J B, de Silva V, Langford J C. A global geometric framework for nonlinear dimensionality reduction. *Science*, 2000, **290**(5500): 2319–2323
- Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, **290**(5500): 2323–2326
- Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 2003, **15**(6): 1373–1396
- He X F, Niyogi P. Locality preserving projections. In: *Proceedings of the 16th Conference on Neural Information Processing Systems*. Cambridge, USA: MIT Press, 2003. 153–160
- He X F, Cai D, Yan S C, Zhang H J. Neighborhood preserving embedding. In: *Proceedings of the 10th IEEE Conference on Computer Vision*. Beijing, China: IEEE, 2005. 1208–1213
- He X F, Yan S C, Hu Y X, Niyogi P, Zhang H J. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(3): 328–340
- Yang J, Zhang D, Frangi A F, Yang J Y. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(1): 131–137
- Li M, Yuan B Z. 2D-LDA: a statistical linear discriminant analysis for image matrix. *Pattern Recognition Letters*, 2005, **26**(5): 527–532
- Wen Ying, Shi Peng-Fei. An approach to face recognition based on common vector and 2DPCA. *Acta Automatica Sinica*, 2009, **35**(2): 202–205  
(文颖, 施鹏飞. 一种基于共同向量结合 2DPCA 的人脸识别方法. *自动化学报*, 2009, **35**(2): 202–205)
- Zuo W M, Zhang D D, Wang K. Bidirectional PCA with assembled matrix distance metric for image recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2006, **36**(4): 863–872
- He X F, Cai D, Niyogi P. Tensor subspace analysis. In: *Proceedings of the 19th Conference on Neural Information Processing Systems*. Cambridge, USA: MIT Press, 2006. 499–506
- Wang H, Yan S C, Huang T, Tang X O. A convergent solution to tensor subspace learning. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence*. Hyderabad, India: Morgan Kaufmann Publishers, 2007. 629–634
- Zhang T H, Yang J, Zhao D L, Ge X L. Linear local tangent space alignment and application to face recognition. *Neurocomputing*, 2007, **70**(7–9): 1547–1553

- 20 Yan S C, Xu D, Zhang B Y, Zhang H J, Yang Q, Lin S. Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(1): 40–51
- 21 Yang W K, Wang J G, Ren M W, Yang J Y, Zhang L, Liu G H. Feature extraction based on Laplacian bidirectional maximum margin criterion. *Pattern Recognition*, 2009, **42**(11): 2327–2334
- 22 Turk M A, Pentland A P. Face recognition using eigenfaces. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Maui, USA: IEEE, 1991. 586–591
- 23 Yang J, Zhang D, Yong X, Yang J Y. Two-dimensional discriminant transform for face recognition. *Pattern Recognition*, 2005, **38**(7): 1125–1129
- 24 Niu B, Yang Q, Shiu S C K, Pal S K. Two-dimensional Laplacianfaces method for face recognition. *Pattern Recognition*, 2008, **41**(10): 3237–3243
- 25 Zelnik-Manor L, Perona P. Self-tuning spectral clustering. In: Proceedings of the 17th Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press, 2004. 1601–1608
- 26 Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression (PIE) database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(12): 1615–1618
- 27 Georghiades A S, Belhumeur P N, Kriegman D J. From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(6): 643–660



**詹宇斌** 中国人民解放军国防科学技术大学计算机学院博士研究生. 2004 年获得武汉大学数学与统计学院学士学位, 2006 年获得中国人民解放军国防科学技术大学计算机学院硕士学位. 主要研究方向为机器学习, 数据降维技术, 模式识别. 本文通信作者.

E-mail: yubinzh@nudt.edu.cn

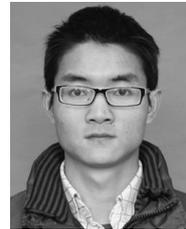
(**ZHAN Yu-Bin** Ph. D. candidate at the School of Computer, National University of Defense Technology. He received his bachelor degree from the School of Mathematics and Statistics, Wuhan University and master degree from the School of Computer, National University of Defense Technology in 2004 and 2006, respectively. His research interest covers machine learning, dimensionality reduction, and pattern recognition. Corresponding author of this paper.)



**殷建平** 中国人民解放军国防科学技术大学计算机学院教授. 主要研究方向为人工智能, 模式识别, 算法设计与信息安全.

E-mail: jpyin@nudt.edu.cn

(**YIN Jian-Ping** Professor at the School of Computer, National University of Defense Technology. His research interest covers artificial intelligence, pattern recognition, algorithm design, and information security.)



**刘新旺** 中国人民解放军国防科学技术大学计算机学院博士研究生. 2008 年获得中国人民解放军国防科学技术大学计算机学院硕士学位. 主要研究方向为机器学习, 核方法和模式识别.

E-mail: xwliu@nudt.edu.cn

(**LIU Xin-Wang** Ph. D. candidate at the School of Computer, National University of Defense Technology. He received his master degree from the School of Computer, National University of Defense Technology in 2008. His research interest covers machine learning, kernel methods, and pattern recognition.)