

# 中心近邻嵌入学习算法的人脸识别研究

孔万增 朱善安

(浙江大学电气工程学院, 杭州 310027)

**摘要** 针对人脸识别问题,提出了一种中心近邻嵌入的学习算法,其与经典的局部线性嵌入和保局映射不同,它是一种有监督的线性降维方法。该方法首先通过计算各类样本中心,并引入中心近邻距离代替两样本点之间的直接距离作为权系数函数的输入;然后再保持中心近邻的几何结构不变的情况下把高维数据嵌入到低维坐标系中。通过中心近邻嵌入学习算法与其他3种人脸识别方法(即主成分分析、线性判别分析及保局映射)在ORL、Yale及UMIST人脸库上进行的比较实验结果表明,它在高维数据低维可视化和人脸识别效果等方面均较其他3种方法取得了更好的效果。

**关键词** 人脸识别 中心近邻嵌入 有监督学习 线性降维

中图法分类号:TP391.41 文献标识码:A 文章编号:1006-8961(2008)04-0691-05

## Learning Algorithm of Center Based Neighborhood Embedding for Face Recognition

KONG Wan-zeng, ZHU Shan-an

(College of Electrical Engineering, Zhejiang University, Hangzhou 310027)

**Abstract** In this paper, a novel learning algorithm called center based neighborhood embedding (CNE) is proposed to deal with face recognition. Unlike the classical methods such as local linear embedding (LLE) and local preserving projection (LPP), CNE is a supervised linear dimensionality reduction method. It first computes centers of all sample classes. The input of the weight function between two samples was replaced by center based neighborhood (CN) distance. Then, the high-dimensional data are embedded into a low-dimensional space with preserving the CN geometric structure. The CNE approach is compared with principle component analysis (PCA), linear discriminant analysis (LDA) and local preserving projection (LPP) on ORL, Yale and UMIST databases. Experiments demonstrate the proposed method is superior to other three methods in terms of both lower-dimensional visualization and recognition accuracy.

**Keywords** face recognition, center based neighborhood embedding, supervised learning, linear dimensionality reduction

## 1 引言

由于人脸识别可应用于公安刑侦、安防监控、人机交互等场合,因此在最近的几年中,已成为计算机视觉和模式识别领域的研究热点。其中研究最广泛的一支是基于表象<sup>[1,2]</sup>(appearance-based)的人脸识别方法,它把 $m \times n$  pixels的人脸图片用一个 $m \times n$ 维的列向量来表示,由于 $m \times n$ 维对人脸

识别来讲,维数太大,影响了人脸识别的快速而稳定的运行,因此需要通过降维来解决此问题。经典的线性降维方法是主成分分析(PCA)<sup>[2]</sup>,线性判别分析(LDA)<sup>[3]</sup>。最近有研究<sup>[4,5]</sup>表明,人脸图片有可能存在于非线性流形子空间,而传统的方法,如PCA和LDA方法只对欧氏空间结构有效,而对非线性流形结构却无能为力。于是一些非线性的流形学习方法,如局部线性嵌入(LLE)<sup>[5,6]</sup>、等规映射(isomap)<sup>[7]</sup>、拉普拉斯本征映射(Laplacian

eigenmap)<sup>[8]</sup> 成为该领域的研究热点。虽然这些方法能对训练样本的非线性流形结构进行很好的学习,但对新测试样本却没有定义。由于这些方法学习泛化能力低,因此仍不适用于人脸识别。于是一种泛化能力较强的保局映射(LPP)方法<sup>[9,10]</sup>被应用于人脸识别,它的识别效果较传统的 PCA、LDA 方法有了很大的改进。这些流形学习方法的一个共同点,就是原始样本可根据  $k$ -近邻或  $\varepsilon$ -邻域构建一幅邻接权系数拓扑图,使得高维数据点映射到低维坐标系时,仍保留了邻接点之间固有的几何结构。可这仍是一种无监督的学习方法。于是本文根据中心最近邻分类器<sup>[11]</sup>思想,提出了一种中心近邻嵌入(CNE)的学习算法,该方法邻接点拓扑图区别于上述几种方法,它首先计算出各类样本中心,同时引入中心近邻距离的概念,即以一样本点到经过该类的样本中心和另一已知样本点的直线的距离;然后采用中心近邻距离代替传统的两样本点之间的直接距离作为拓扑图中权系数函数的输入;最后保持中心近邻的几何结构把高维数据映射到低维坐标系中,用于进行人脸识别分类。

## 2 中心近邻嵌入

考虑一个人脸样本集合  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_n]$ , 有  $C$  个不同的类,共  $n$  个样本,其中  $\mathbf{x}_i \in \mathbf{R}^l$ ,  $l$  是原始人脸样本图片的维数,一般较高。为了降低维数,可选取一传递矩阵  $\mathbf{A}$ ,使得  $\mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$ ,  $\mathbf{y}_i \in \mathbf{R}^d$  ( $d < l$ ), 向量  $\mathbf{y}_i$  是人脸图片在低维空间的表示。从高维空间映射至低维空间时,则采用 LPP 方法<sup>[9,10]</sup>的目标函数,并使之最小化

$$\min \sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 W_{i,j} \quad (1)$$

式中,  $W_{i,j}$  为两样本点之间的权系数。围绕式(1)目标函数,本文 CNE 算法可分构造邻接图、选择权系数、CNE 映射 3 个步骤。

### 2.1 构造邻接图

令  $\mathbf{G}$  为包含  $n$  个结点的连接图,每个结点代表一人脸样本。首先,计算各个类的类中心,即

$$\mathbf{x}_k^c = \frac{1}{N_k} \sum_{\mathbf{x}_i \in \mathbf{c}_k} \mathbf{x}_i \quad (2)$$

式中,  $\mathbf{c}_k$  为第  $k$  类样本集合,  $N_k$  为第  $k$  类的样本个数,  $\mathbf{x}_k^c$  为第  $k$  类的样本中心;然后每个样本结点与该样本所在类的类中心赋一条连接边,即样本与样本之间无直接连接,而是与该类的中心连接(如图 1 所示)。

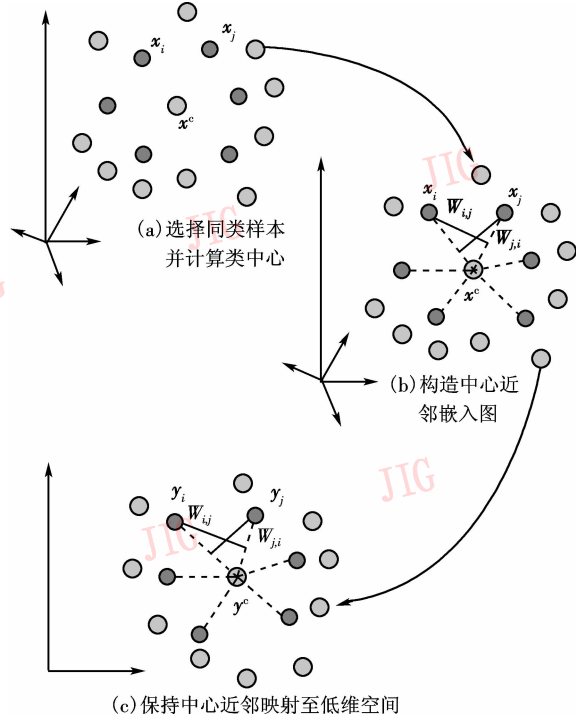


图 1 CNE 流程图

Fig. 1 The procedure in CNE

### 2.2 选择权系数

在中心近邻嵌入中,两样本点  $\mathbf{x}_i, \mathbf{x}_j$  的连接权系数是有向权,即  $\mathbf{x}_i \rightarrow \mathbf{x}_j$ , 对应权  $W_{i,j}$ , 而对于  $\mathbf{x}_j \rightarrow \mathbf{x}_i$ , 则对应权  $W_{j,i}$ 。一般地,  $W_{i,j} \neq W_{j,i}$ 。权系数  $W_{i,j}$  定义为

$$W_{i,j} = \begin{cases} \exp\left(-\frac{d^2(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k^c)}{t}\right) & (\mathbf{x}_i \in \mathbf{c}_k \wedge \mathbf{x}_j \in \mathbf{c}_k) \\ 0 & \text{其他} \end{cases} \quad (3)$$

式中,  $d(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k^c)$  为中心近邻距离,几何示意如图 2 所示。可见 CNE 方法的中心近邻距离区别于 LLE 和 LPP 方法中两样本点之间的直接距离。考虑两个同类的人脸样本,由于姿态、光照及表情等因素影响,使得两样本在高维空间分布较散,从而导致它们之间的直接距离较远,但是由于它们的中心近邻距离却较近,因此以中心近邻距离作为权系数参数得到的嵌入子空间则更为精确。中心近邻距离的数学定义为

$$d(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k^c) = \|\mathbf{x}_i - \mathbf{p}_{i,j}^c\| \quad (4)$$

式中,  $\mathbf{p}_{i,j}^c$  是点  $\mathbf{x}_i$  到直线  $\overrightarrow{\mathbf{x}_j \mathbf{x}_k^c}$  的投影,投影点  $\mathbf{p}_{i,j}^c$  可表示为

$$\mathbf{p}_{i,j}^c = \mathbf{x}_j + \xi(\mathbf{x}_k^c - \mathbf{x}_j) \quad (5)$$

式中,  $\xi \in \mathbf{R}$  称位置因子,可按下式计算:

$$\xi = \frac{(\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_k^c - \mathbf{x}_j)}{(\mathbf{x}_k^c - \mathbf{x}_j)^T (\mathbf{x}_k^c - \mathbf{x}_j)} \quad (6)$$

把式(6)、式(5)、式(4)、式(3)逐级代入,便可求得权系数  $W_{i,j}$ 。

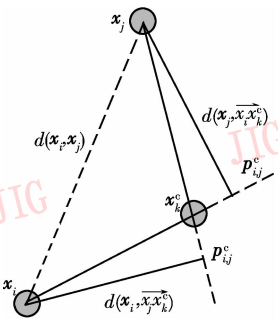


图2 中心近邻距离几何示意图

Fig. 2 Geometric explanation of CN distance

### 2.3 CNE 映射

根据 CNE 方法中权系数的定义,一般  $W_{j,i} \neq W_{i,j}$ ,因此权系数阵  $\mathbf{W}$  为非对称阵,即  $\mathbf{W} \neq \mathbf{W}^T$ 。把降维公式  $\mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$  代入式(1)中,再通过线性代数推导,即可得出与 LPP 不同的结果:

$$\begin{aligned} & \sum_{i,j} \|\mathbf{A}^T \mathbf{x}_i - \mathbf{A}^T \mathbf{x}_j\|^2 W_{i,j} \\ &= \sum_{i,j} (\mathbf{x}_i^T \mathbf{A} \mathbf{A}^T \mathbf{x}_i - \mathbf{x}_i^T \mathbf{A} \mathbf{A}^T \mathbf{x}_j + \mathbf{x}_j^T \mathbf{A} \mathbf{A}^T \mathbf{x}_i - \mathbf{x}_j^T \mathbf{A} \mathbf{A}^T \mathbf{x}_j) W_{i,j} \\ &= \sum_{i,j} \text{tr}(\mathbf{A}^T \mathbf{x}_i \mathbf{x}_i^T \mathbf{A} - \mathbf{A}^T \mathbf{x}_i \mathbf{x}_j^T \mathbf{A} + \mathbf{A}^T \mathbf{x}_j \mathbf{x}_i^T \mathbf{A} - \mathbf{A}^T \mathbf{x}_j \mathbf{x}_j^T \mathbf{A}) W_{i,j} \\ &= \text{tr} \sum_{i,j} (\mathbf{A}^T (\mathbf{x}_i W_{i,j} \mathbf{x}_i^T + \mathbf{x}_j W_{i,j} \mathbf{x}_j^T - \mathbf{x}_i W_{i,j} \mathbf{x}_j^T - \mathbf{x}_j W_{i,j} \mathbf{x}_i^T) \mathbf{A}) \\ &= \text{tr}(\mathbf{A}^T \mathbf{X}(\mathbf{D} + \mathbf{S}) \mathbf{X}^T \mathbf{A} - \mathbf{A}^T \mathbf{X}(2\mathbf{W}) \mathbf{X}^T \mathbf{A}) \\ &= \text{tr}(\mathbf{A}^T \mathbf{X}(\mathbf{D} + \mathbf{S} - 2\mathbf{W}) \mathbf{X}^T \mathbf{A}) \end{aligned}$$

式中,  $\text{tr}(\cdot)$  是矩阵的迹,  $\mathbf{D}$  和  $\mathbf{S}$  是对角阵, 对角线上的元素分别为该元素所在行、列上全部元素之和即  $D_{ii} = \sum_j W_{i,j}$ ,  $S_{jj} = \sum_i W_{i,j}$ , 由于  $\mathbf{W}$  不是对称矩阵, 因此  $\mathbf{D} \neq \mathbf{S}$ 。最优传递矩阵  $\mathbf{A}$  可通过下式计算:

$$\mathbf{A}_{\text{opt}} = \arg \min_{\mathbf{A}} \text{tr}(\mathbf{A}^T \mathbf{X}(\mathbf{D} + \mathbf{S} - 2\mathbf{W}) \mathbf{X}^T \mathbf{A}) \quad (7)$$

从式(7)可知, 矩阵  $\mathbf{A}$  的求解可转化为非对称矩阵  $\mathbf{X}(\mathbf{D} + \mathbf{S} - 2\mathbf{W}) \mathbf{X}^T$  的特征值和特征向量的求解问题, 由于矩阵  $\mathbf{X}(\mathbf{D} + \mathbf{S} - 2\mathbf{W}) \mathbf{X}^T$  是非对称的, 因此特征值分解后可能出现复特征值 (complex eigenvalue) 和复特征向量 (complex eigenvector), 而与图片数据对应的是实特征向量, 那么舍弃那些复特征向量, 选最小的  $d$  个实特征值所对应的特征向量组成

$l \times d$  维矩阵  $\mathbf{A}$ , 即为人脸识别中的特征子空间, 参数  $d$  一般选取样本类个数  $C$  附近的数值。

### 3 几种人脸识别方法比较

为对比几种人脸识别方法的性能, 从学习方式 (是否有监督, 是否基于流形)、目标函数 (是否加权)、分解矩阵 (是否对称) 以及学习泛化能力 (是否具备) 等 5 个方面来评价 CNE 方法与其他几种常用的人脸识别方法的性能。由于 CNE 方法和 LLE、LPP 方法均采用构造邻接图的方法来降维, 属于流形学习, 因此能挖掘人脸潜在的非线性流形结构; 由于这 3 种方法在构造邻接图时都采用了不同的权系数, 所以目标函数有权重; CNE 方法在构造邻接图时利用了各类的中心, 为有监督学习; 由于中心近邻距离的不对称, 从而导致其权系数矩阵的不对称, 也使得其分解矩阵不同于目前的人脸识别方法, 是一个非对称阵。由于 CNE 方法不仅对训练样本有定义, 而且对新的测试样本也有定义, 因此具备泛化能力, 可以应用于人脸识别。综上所述, CNE 方法与其他方法在各个性能指标上的比较如表 1 所示。

表 1 几种识别方法指标比较

Tab. 1 Comparison of several methods in face recognition

识别方法	是否监督学习	是否有权重函数	分解矩阵是否对称	是否有泛化能力	是否流形学习
PCA	否	否	是	是	否
LDA	是	否	是	是	否
LLE	否	是	是	否	是
LPP	否	是	是	是	是
CNE	是	是	否	是	是

### 4 实验结果及分析

为验证本文算法的识别效果, 利用 ORL、Yale 及 UMIST 等人脸库分别进行了低维可视化、人脸识别的实验, 实验结果表明, CNE 方法是有效的。

#### 4.1 高维数据低维可视化

人脸图片空间维数一般都较高, 可是由于人的视觉感知最多只有 3 维, 因此高维向量数据在空间上是不可视的, 也就无法看出这些数据的内在联系。为此本文在 ORL 库进行了数据可视化实验。ORL 人脸库包括 40 个人, 每人 10 张  $112 \times 92$  大小的灰度照片。照片中人脸的表情、位置和尺度均有不同

程度的变化。实验时,首先取其中前 20 个人,每人 10 张共 200 张图片分别用 PCA、LDA、LPP 及 CNE 方法进行训练学习,然后从 10 304 维映射至 2 维空间,映射结果如图 3 所示。由于 CNE 利用中心近邻距离作为权系数,因此不仅能保持样本之间的固有几何关系,而且类中心的应用强化了类属信息,从图 3(a)可见,用 CNE 方法降维后数据的类属关系较为清楚;PCA 方法由于只应用了总体样本中心,使得

样本协方差最大,因此降维后数据的类属关系仍十分模糊。LPP 方法只保持  $k$  近邻或  $\epsilon$  邻域的邻接点关系,因此从高维降至 2 维后,虽然结构上比 PCA 方法要紧一些,但类属关系仍不清楚。由于 LDA 方法使用了类内散布阵最小且类间散布阵最大的原则,因此在 2 维空间有较好的聚类效果,但是由于没有保持各点的邻接点关系,因此空间结构不如 CNE 紧密。

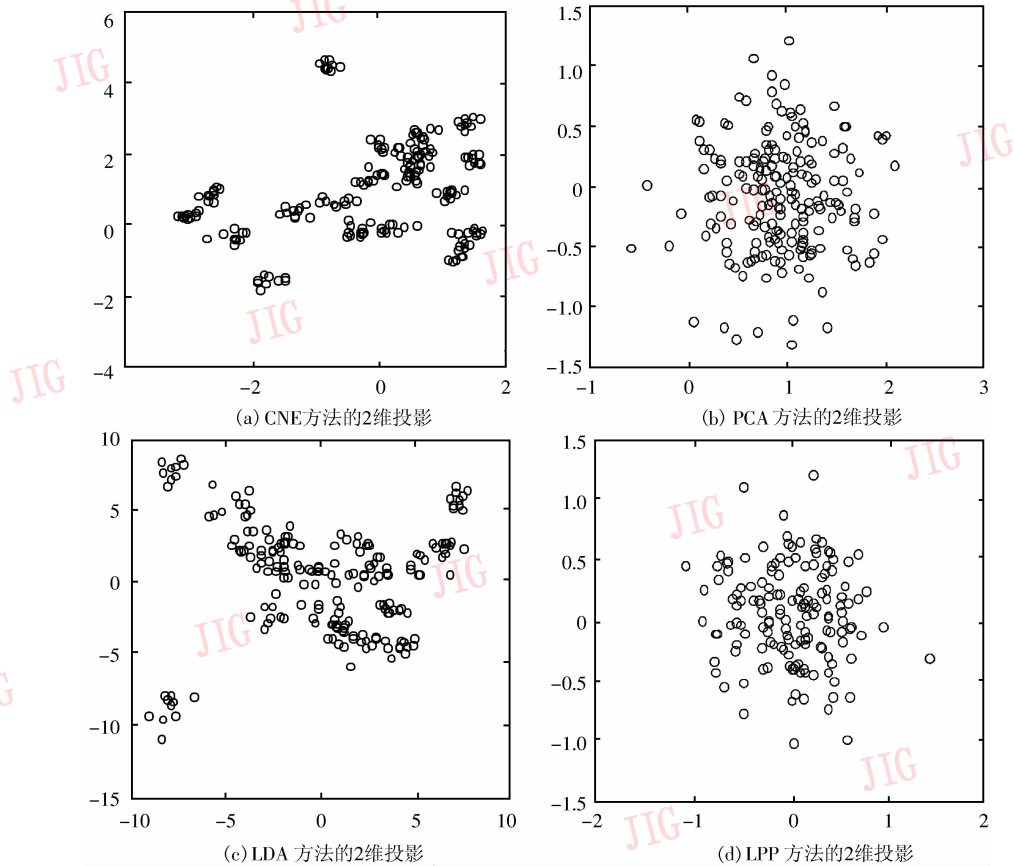


图 3 几种识别方法映射结果的 2 维投影

Fig. 3 Two-dimensional projections achieved

### 4.2 人脸识别

CNE 方法与人脸识别领域中最经典的其他几种算法,如 PCA、LDA、LPP 等方法在 Yale 和 UMIST 库上进行了比较实验。Yale 库包含 15 人,每人 11 张  $100 \times 100$  的灰度照片。照片中人脸有表情、光照变化及是否佩戴眼镜等。UMIST 库包含 20 个不同种族、不同性别共 564 张  $112 \times 92$  大小的灰度图片,每人样本由一系列从侧面至正面的人脸图片组成。实验时,把两个脸库的样本都分作训练和测试两部分:Yale 库每人随机选 6 张作训练样本,共有训练样本 90 张,其余部分用作测试;UMIST 库每人随机选

取 10 张图片作为训练样本,其余 364 张图片作为测试样本。对于 LDA 方法,由于最多有  $C-1$  个非零特征值,因此低维特征空间维数的上界就是  $C-1$ ,其中  $C$  为样本类的个数,而其余 3 种方法的识别效果则受特征空间维数影响,表 2、表 3 中维数均是各种方法在权衡了识别率后的优化维数。由表 2 表 3 可见,即使再增加维数,它们的识别率也不会有明显的改变。识别时是采用最近邻分类器。Yale 库的实验结果显示:CNE、LPP、LDA、PCA 4 种方法各自在 13, 28, 14, 33 维情况下的识别率分别为 93.8%, 89.3%, 80%, 75.5%, 这表明 CNE 方法的特征子空

间维数比其他方法低,但识别率却比其他方法要高。UMIST 库的实验结果说明,4 种算法要达到大致相当的识别率,PCA 方法所需的子空间维数最多,为 116,其次是 LPP 方法需 60,再次是 LDA 方法需 19,最少是 CNE 方法,仅需 15 维就能获得高达 96.11% 的识别率。

表 2 Yale 库上识别性能比较

Tab. 2 Performance comparison on the Yale database

方法	识别率(%)	维数
PCA	75.5	33
LDA	80.0	14
LPP	89.3	28
CNE	93.8	13

表 3 UMIST 库上识别性能比较

Tab. 3 Performance comparison on UMIST database

方法	识别率(%)	维数
PCA	95.6	116
LDA	97.0	19
LPP	96.40	60
CNE	96.11	15

## 5 结 论

随着人脸识别研究的日益深入,基于构造邻接图的学习算法成为该领域的研究热点,本文的方法在构造邻接图时,两样本点的权系数函数的输入是采用中心近邻距离,而不是两点间的直接距离。在高维空间向低维嵌入时是用线性代数的方法推导出与 LPP 方法不同的特征子空间。由于该低维子空间不仅包含原样本的固有几何关系,同时引入了类别信息,因此在 ORL、Yale 及 UMIST 人脸库的低维可视化及人脸识别实验中体现出了较好的性能。此

外,由于中心近邻距离的不对称,因此特征值分解时出现了复特征值和复特征向量,二者的数乘意味着该向量的拉伸和旋转,可见尝试采用复特征子空间来研究人脸的多姿态问题将是下一步的研究目标。

## 参考文献 (References)

- 1 Turk M, Pentland A. Eigenfaces for recognition[J]. *Journal Cognitive Neuroscience*, 1991, 3(1): 71 ~ 86.
- 2 Murase H, Nayar S K. Visual learning and recognition of 3-D objects from appearance[J]. *International Journal of Computer Vision*, 1995, 14(1): 5 ~ 24.
- 3 Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711 ~ 720.
- 4 Chang Y, Hu C, Turk M. Manifold of Facial Expression[A]. In: *Proceeding of IEEE International Workshop on Analysis and Modeling of Faces and Gestures[C]*, Nice, France, 2003: 203 ~ 205.
- 5 Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, 290: 2323 ~ 2326.
- 6 Saul L K, Roweis S T. Think globally, fit locally: unsupervised learning of low dimensional manifolds[J]. *Machine Learning Research*, 2003, 4(6): 119 ~ 155.
- 7 Tenenbaum J B, Silva V De, Langford J C. A global feometric framework for nonlinear dimensionality reduction[J]. *Science*, 2000, 290: 2319 ~ 2322.
- 8 Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. *Neural Computation*, 2003, 15(6): 1373 ~ 1396.
- 9 He X, Yang S, Hu Y, et al. Face recognition using laplacianfaces[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(3): 328 ~ 340.
- 10 He X, Niyogi P. Locality Preserving Projections[A]. In: *Proceedings Conference of Advances in Neural Information Processing Systems[C]*, Vancouver, Canada, 2003: 321 ~ 328.
- 11 Gao Qing-bin, Wang Zheng-zhi. Center-based nearest neighbor classifier[J]. *Pattern Recognition*, 2007, 40(1): 346 ~ 349.