

基于类别信息的邻域保持嵌入算法

包 兴 张 莉 赵梦梦 杨季文

(苏州大学计算机科学与技术学院 苏州 215006)

摘 要 邻域保持嵌入通常被广泛用于发现高维数据的固有内在维数。为了充分利用样本的类别信息,构建了一个具有判别信息的邻接矩阵,其可以使同类样本点更加紧凑而异类样本点更加疏远。在此基础上,提出了基于类别信息的邻域保持嵌入算法。基于类别信息的邻域保持嵌入算法在不破坏原始高维数据局部几何结构的同时,可以使处于不同子流形上的样本点尽量分开。在 UCI 数据集和 ORL 人脸数据集上的实验结果表明,基于类别信息的邻域保持嵌入算法具有较高的识别率。

关键词 降维,邻接矩阵,类别信息,人脸识别

中图法分类号 TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.5.019

Label Information-based Neighborhood Preserving Embedding

BAO Xing ZHANG Li ZHAO Meng-meng YANG Ji-wen

(School of Computer Science and Technology, Soochow University, Suzhou 215006, China)

Abstract Neighborhood preserving embedding (NPE) is widely used for finding the intrinsic dimensionality of the data with high dimension. In order to make full use of the classification information of samples to get optimal features, we constructed an adjacent matrix which can separate different sub-manifolds as far as possible without destroying local geometry structure of the original data. By introducing the adjacent matrix, this paper proposed label information-based neighborhood preserving embedding (LINPE). Experiments on UCI data and ORL face databases were performed to test and evaluate LINPE. Experimental results demonstrate the effectiveness of LINPE.

Keywords Dimension reduction, Adjacent matrix, Label information, Face recognition

1 引言

随着数据采集技术的飞速发展,高维数据的数量正在迅速增加,尤其在模式识别领域内^[1-3]。这给数据的分析带来了许多挑战。降维是数据处理的必要步骤,它能够避免维数灾难^[4-6]。降维的目的是在数据维数的约简过程中,尽量保持高维数据的某些特征。降维不仅能够提高分类器的性能,还能够降低数据计算的复杂度。

在过去的几十年中,大量的降维方法被提出,其中大部分已经被成功地应用于许多领域,如文本处理、人脸识别等^[7-9]。这些方法大致可以分为两类:无监督学习方法和监督学习方法。到现在为止,主成分分析(Principal Component Analysis, PCA)^[10,11]是最流行的无监督方法,它旨在寻找在最小均方意义下最能代表原始数据的投影方向。线性判别分析(Linear Discriminant Analysis, LDA)^[12]是一种有监督的降维算法,它通过最大化类间散度和类内散度的比值寻求最优的投影方向。也就是说,PCA 寻找的是有效表示数据的方向,而 LDA 寻找的是期望有效分类数据的方向。值得注意的是,PCA 和

LDA 都有一个共同的特点,即它们只考虑到数据的全局线性结构,并没有考虑数据的局部几何结构。其结果是,当数据分布在一个非线性的流形上时,PCA 和 LDA 都不能发现数据潜在的内部结构信息。

非线性降维方法的典型代表是流形学习。通常,流形学习首先要构建一个邻接图来刻画数据的分布或几何结构,然后寻求能够有效保持该结构的最优映射或投影方向。一些流形学习降维算法相继被提出,如拉普拉斯特征映射(Laplacian Eigenmaps, LE)^[13]、局部线性嵌入(Locally Linear Embedding, LLE)^[14,15]、局部保持投影(Locality Preserving Projections, LPP)^[16]、邻域保持嵌入(Neighborhood Preserving Embedding, NPE)^[17]和判别邻域嵌入(Discriminant Neighborhood Embedding, DNE)^[19]。

局部保持投影(LPP)是拉普拉斯特征映射(LE)的线性化算法,它能够较好地保持非线性子流形中局部数据特征的线性流形,其目标是保持数据之间的相似关系,即原始数据空间上相邻的数据点在投影后的空间上也能保持相应的相邻关系。邻域保持嵌入(NPE)是局部线性嵌入(LLE)的一种线性

到稿日期:2014-02-18 返修日期:2014-04-15 本文受国家自然科学基金(61373093, 61033013, 61271301),江苏省自然科学基金(BK2011284, BK201222725),江苏省高校自然科学基金项目(13KJA520001),江苏省青蓝工程资助。

包 兴(1989-),男,硕士生,主要研究方向为机器学习与模式识别, E-mail: 20124227034@suda.edu.cn; 张 莉(1975-),女,教授,硕士生导师, CCF 高级会员,主要研究方向为机器学习、模式识别、图像处理; 赵梦梦(1991-),女,硕士生,主要研究方向为机器学习; 杨季文(1963-),男,教授,硕士生导师,主要研究方向为中文信息处理、嵌入式系统及应用、智能化信息处理技术。

近似,具有保持数据流形上局部邻域结构信息的能力。邻域保持嵌入已经在人脸识别问题中受到了广泛的关注。但是,在人脸识别问题中,NPE是作为一种无监督的降维方法引入的,它没有利用已知数据的类别信息。为了能利用已知数据的类别信息,文献[18]提出了一种新的子空间学习方法,该方法称为邻域保持判别嵌入(Neighborhood Preserving Discriminant Embedding, NPDE)。NPDE在保持流形数据局部邻域结构信息的同时,强调数据集的判别信息,在保证类内局部邻域重构误差最小和权重固定的情况下,使得嵌入映射得到的样本点具有最小类内距离和最大类间距离。判别邻域嵌入^[19]也是一种有监督的降维方法。由于同类的数据会分布在一个较小的子流形上,不同类的数据会分布在不同的子流形上。为了充分利用样本数据的类别信息,DNE构建一个具有判别信息的邻接矩阵,目的是在降维时能够使同类样本点更加紧凑而异类样本点更加疏远。

本文结合DNE和NPE两种方法的优点,提出了基于类别信息的邻域保持嵌入算法。基于类别信息的邻域保持嵌入算法在不破坏原始高维数据局部几何结构的同时,可以使处于不同子流形上的样本点尽量分开,从而进一步提高识别性能。

2 相关工作

本节回顾与本文工作紧密相关的研究工作,包括邻域保持嵌入和判别邻域嵌入。

2.1 邻域保持嵌入(NPE)

令训练样本矩阵表示集为 $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$,其中 $\mathbf{x}_i \in R^m$, n 为训练样本的个数, m 为训练样本的维数。NPE的目的是在降维的同时保持数据集固有的局部邻域流形结构不变。它寻找一个最优的映射变换矩阵 $\mathbf{A}=[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$,其中 $\mathbf{a}_i \in R^m$, d 为训练样本映射变换后的维数,将 R^m 空间的数据嵌入映射到一个相对低维的特征空间 R^d ($d \in m$)中。数据点 \mathbf{x}_i 在 R^d 中表示为 \mathbf{y}_i ,且 $\mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$ 。

NPE从本质上说是局部线性嵌入(LLE)的线性逼近,它采用了与LLE相同的方法来构造训练样本上的近邻图。NPE假定每个局部近邻都是线性的,因此每个数据点 \mathbf{x}_i 都可以通过它的个近邻点线性重构。令 $knn(\mathbf{x}_i)$ 表示 \mathbf{x}_i 的近邻集合,近邻图重构权重系数矩阵 \mathbf{W} 满足下面的要求:

$$\begin{cases} W_{ij} = 0, & \text{if } \mathbf{x}_j \notin knn(\mathbf{x}_i) \\ W_{ij} \neq 0, & \text{if } \mathbf{x}_j \in knn(\mathbf{x}_i) \end{cases} \quad (1)$$

对于非零的权重系数,要通过下面的重构损失函数最小化来求得:

$$\min \varphi(\mathbf{W}) = \sum_{i=1}^n \left\| \mathbf{x}_i - \sum_{j=1}^k W_{ij} \mathbf{x}_j \right\|^2 \quad (2)$$

约束条件为 $\sum_{j=1}^k W_{ij} = 1, i=1, 2, \dots, n$ 。

NPE认为,如果 \mathbf{W} 能在 R^m 空间中重构数据点 \mathbf{x}_i ,则它也可以在 R^d 空间中重构对应的点 \mathbf{y}_i 。因此,嵌入映射矩阵 \mathbf{A} 可以通过求解下面最小化问题得到:

$$\min J_1(\mathbf{A}) = \sum_{i=1}^n \left\| \mathbf{y}_i - \sum_{j=1}^k W_{ij} \mathbf{y}_j \right\|^2 \quad (3)$$

2.2 判别邻域嵌入(DNE)

文献[19]提出了判别邻域嵌入(DNE)方法。DNE结合邻域和原始样本的类别信息,有效地提高了降维后的分类效果。

已知训练样本集 $\{\mathbf{x}_i, c_i\}_{i=1}^n$,其中 $\mathbf{x}_i \in R^m$, c_i 是 \mathbf{x}_i 的类别

标签, $c_i \in \{1, 2, \dots, C\}$, C 是类别数。引入线性变换矩阵 \mathbf{A} 进行投影嵌入 $\mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$, $\mathbf{y}_i \in R^d$ 。为了使分布在不同子流形上的样本点尽可能的分开,DNE构建一个具有判别信息的邻接矩阵 \mathbf{H} 。邻接矩阵 \mathbf{H} 的构造如下:

$$H_{ij} = \begin{cases} -1, & (\mathbf{x}_i \in knn(\mathbf{x}_j) \vee \mathbf{x}_j \in knn(\mathbf{x}_i)) \wedge (c_i \neq c_j) \\ +1, & (\mathbf{x}_i \in knn(\mathbf{x}_j) \vee \mathbf{x}_j \in knn(\mathbf{x}_i)) \wedge (c_i = c_j) \end{cases} \quad (4)$$

即对不同类数据,如果 \mathbf{x}_j 是 \mathbf{x}_i 的 k 个最近邻之一或者 \mathbf{x}_i 是 \mathbf{x}_j 的 k 个最近邻之一,那么 $H_{ij} = 1$;对同类数据,如果 \mathbf{x}_j 是 \mathbf{x}_i 的 k 个最近邻之一或者 \mathbf{x}_i 是 \mathbf{x}_j 的 k 个最近邻之一,那么 $H_{ij} = -1$ 。其中, \mathbf{H} 是对称矩阵。其目标函数可以简化为:

$$\max J_2(\mathbf{A}) = \sum_{i,j} \left\| \mathbf{A}^T \mathbf{x}_i - \mathbf{A}^T \mathbf{x}_j \right\|^2 H_{ij} \quad (5)$$

3 基于类别信息的邻域保持嵌入(LINPE)

本节提出了一种有监督的NPE方法,该方法在不破坏原始高维数据局部几何结构的同时,通过构建一个具有判别信息的邻接矩阵,使高维数据在降维时能够让处于不同子流形上的样本点尽可能的分开。

已知训练样本集 $\{\mathbf{x}_i, c_i\}_{i=1}^n$,其中 $\mathbf{x}_i \in R^m$, c_i 是 \mathbf{x}_i 的类别标签, $c_i \in \{1, 2, \dots, C\}$, C 是类别数。引入线性变换矩阵 \mathbf{A} 进行投影嵌入 $\mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$ 。

利用式(3)来保持样本原有的局部几何结构;利用式(5)使分布在不同子流形上的类内样本更加紧凑而类间样本彼此远离。结合式(3)和式(5),LINPE的目标函数可以简化为:

$$\max_{\mathbf{A}} \frac{\sum_{i,j} \left\| \mathbf{A}^T \mathbf{x}_i - \mathbf{A}^T \mathbf{x}_j \right\|^2 H_{ij}}{\sum_{i=1}^n \left\| \mathbf{y}_i - \sum_{j=1}^k W_{ij} \mathbf{y}_j \right\|^2} \quad (6)$$

最大化式(6)等价于最大化式(5)的同时最小化式(3)。

式(6)的分子可以转化为:

$$\begin{aligned} \sum_{i,j} \left\| \mathbf{A}^T \mathbf{x}_i - \mathbf{A}^T \mathbf{x}_j \right\|^2 H_{ij} &= 2tr(\mathbf{A}^T \mathbf{X}(\mathbf{D} - \mathbf{H})\mathbf{X}^T \mathbf{A}) \\ &= 2tr(\mathbf{A}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{A}) \end{aligned} \quad (7)$$

其中, tr 表示矩阵的迹, $\mathbf{L} = \mathbf{D} - \mathbf{H}$, $D_{ii} = \sum_j H_{ij}$ 。式(6)的分母可以转化为:

$$\begin{aligned} \sum_{i=1}^n \left\| \mathbf{y}_i - \sum_{j=1}^k W_{ij} \mathbf{y}_j \right\|^2 &= 2tr(\mathbf{A}^T \mathbf{X}(\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W})\mathbf{X}^T \mathbf{A}) \\ &= 2tr(\mathbf{A}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{A}) \end{aligned} \quad (8)$$

其中, \mathbf{I} 是单位矩阵, $\mathbf{M} = (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W})$ 是对称半正定矩阵。把式(7)和式(8)代入到式(6)中,可以把优化问题(6)转化为下面的广义特征值求解问题:

$$\mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{a} = \lambda \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{a} \quad (9)$$

假设 $\mathbf{A}=[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m]$ 是式(9)特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 所对应的特征向量,则嵌入映射变换矩阵是由前 d 个特征向量组成的,即 $\mathbf{A}=[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$ 。因此,给定 $\mathbf{x}_i \in R^m$,可以得到其相应嵌入映射 $\mathbf{y}_i \in R^d$: $\mathbf{x}_i \rightarrow \mathbf{y}_i = \mathbf{A}^T \mathbf{x}_i$ 。LINPE算法步骤可描述如下:

Step 1 计算重构权重矩阵 \mathbf{W} ,以及 $\mathbf{M} = (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W})$ 。

Step 2 根据训练样本的类别信息,通过式(4)计算邻接矩阵 \mathbf{H} 。

Step 3 计算矩阵 $\mathbf{X} \mathbf{L} \mathbf{X}^T$ 和 $\mathbf{X} \mathbf{M} \mathbf{X}^T$ 。

Step 4 通过求解式(9)的广义特征值问题求得嵌入映射矩阵 \mathbf{A} 。其中 $\mathbf{A}=[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$ 为式(9)求得的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ 所对应的特征向量。

Step 5 利用 1-NN 分类器进行分类。

4 实验

为了验证本文算法的有效性,除了在人工数据上进行实验外,我们还选择两个 UCI 数据集^[20]以及 ORL^[21]人脸数据集作为实验数据对象。实验中,本文算法与 PCA、LPP、NPE、DNE、NPDE 方法进行了对比。采用最近邻分类器对降维后的数据进行分类。本文所有实验都是在 MATLAB 平台上实现的。

4.1 人工数据

为了分析、对比本文方法与相关方法 DNE 及 NPDE 的分类效果,设计了一组人工数据。该人工数据是 3 类三维的高斯分布的随机数据。这 3 类数据的协方差矩阵均为单位矩阵,其中第一类中心在 $[0, 0, 0]^T$,第二类中心在 $[2, 2, 2]^T$,第三类中心在 $[4, 4, 4]^T$ 。每次生成数据时,每类训练样本和测试样本各产生 100 个样本点,将其降维到二维平面上。实验中,由于涉及到近邻参数的选择问题,取 k 的值为从 1 到 15 时,对比 DNE、NPDE 和 LINPE 3 种方法在降维后的识别率。从图 1 可以看出近邻参数对降维后的识别率影响不是很大,因此 DNE、NPDE 和 LINPE 3 种方法在计算近邻图重构系数时,近邻参数设置为 $k=5$ 。另外,DNE 和本方法在构建邻接矩阵 H 时,近邻参数也设置为 5。

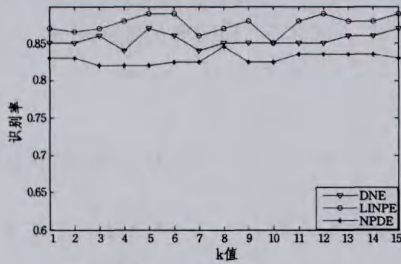


图 1 不同近邻参数下的识别率

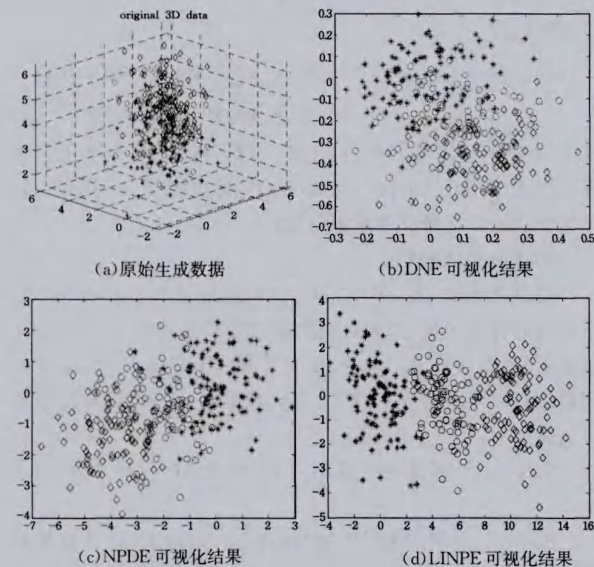


图 2 人工数据及其可视化

图 2(a) 给了一次随机生成的数据,其中符号“+”、“o”和“◇”分别表示不同的类别。图 2(b) — (d) 分别显示通过 DNE、NPDE 和 LINPE 降维后的二维可视化数据分布情况。

由于在原高维空间中数据有重叠,从图 2 可以看到,降维后的数据仍然有重叠。其中 DNE 的重叠现象是最为严重的,这说明该方法并没有很好地融入类别信息。相对而言,LINPE 的重叠现象不如其他两种方法严重。

4.2 基于 UCI 数据集的分类实验

为了验证本文算法的有效性,本实验从 UCI 数据集上选择了 Wpbc 和 Heart 两个子数据集,其信息描述如表 1 所列。

表 1 UCI 数据集

数据集	训练样本个数	测试样本个数	数据维数	类别个数
Wpbc	133	65	33	2
Heart	203	100	13	2

对于 Wpbc 数据集,为了考察各个算法降维后的分类性能,我们保留一定的特征维数,然后利用最近邻分类器进行分类。实验中,构建权值系数时近邻参数设置为 $k=5$,构建邻接矩阵时,近邻参数也设置为 5。不同特征维数下几种算法的识别率如图 3(a) 所示。从图 3(a) 可以看出,本文提出的 LINPE 方法的识别率明显高于 PCA、LPP、NPE、DNE 和 NPDE。然而,当特征维数保留得比较少时,几种方法的识别率均比较低,这是因为降维后原始信息的损失量比较大。当降维后的特征维数大于 15 时,LINPE 的效果明显高于其它几种算法。

对于 Heart 数据集,参数设置同 Wpbc 数据集,实验结果如图 3(b) 所示。从图 3(b) 可以看出,通过 LINPE、DNE 和 NPDE 方法降维后的识别率明显高于 LPP 和 NPE 方法降维后的识别率,这是因为 LINPE、DNE 和 NPDE 考虑了数据的类别信息,是有监督的学习方法;并且当特征维数小于 5 时, NPDE 比 LINPE 的效果好,当降维后的特征维数大于 5 时, LINPE 比 NPDE 的效果好,且识别率增加得比较稳定。

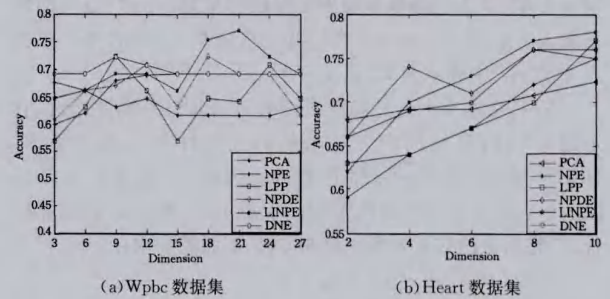


图 3 UCI 数据集上不同特征维数下的识别率

4.3 基于 ORL 人脸数据集的分类实验

ORL 人脸数据库由 40 人、每人 10 幅共 400 幅人脸图像构成。每幅图像灰度级为 256,分辨率为 112×92 。ORL 人脸图像是在不同时间、不同角度、各种表情和不同脸部细节的条件下拍摄的。ORL 人脸数据库的部分人脸图像如图 4 所示。



图 4 ORL 人脸数据库部分图像

实验中,每人分别随机选取 5 幅图像用作训练图像,其余的图像用于测试图像。构建权值系数时近邻参数设置为每类人脸选取的个数,即 $k=5$ 。构建邻接矩阵时,近邻参数也设置为 5。由于初始图像数据的高维性,为了避免发生奇异、降低运算量和消除噪音,首先采用 PCA 方法将训练样本和测试样本分别投影到 100 维的低维线性子空间。对降维后的数据,再采用 LPP、NPE、NPDE、DNE 和 LINPE 方法进行二次降维,分别将训练图像和测试图像的特征维数降到 10, 20, ..., 100, 然后利用最近邻分类器进行识别。

为了消除选择样本的随机独立性,独立重复实验 20 次,最后取平均识别率作为最终识别率。图 5 示出了在 ORL 人脸数据库上,每类人脸训练图像选取 5 幅时,几种不同降维方法在不同特征维数下的识别率。可以看出,随着保留特征维数的增加,各个方法的识别率基本呈上升趋势。DNE、NPDE 和 LINPE 整体上优于 PCA、NPE 和 LPP,这是由于前者包含了类别信息,是有监督的学习方法,同时又提取了图像的非线性特征。在这些方法中,PCA 保持了图像空间的全局结构信息,NPE 和 LPP 保持了图像空间的邻域几何结构,而 LINPE 在保持图像空间邻域局部空间的同时,也考虑了高维数据信息的类别信息。LINPE 也明显地优于另外两种有监督的 NPE,即 DNE 和 NPDE,并且比它们稳定。

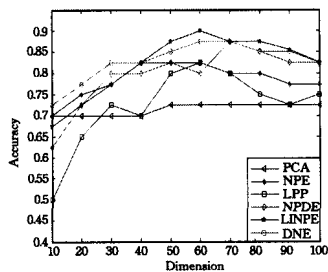


图 5 ORL 数据库上 6 种方法的识别率比较(5 Train)

结束语 本文提出了基于类别信息的邻域保持嵌入算法(LINPE)。LINPE 引入了具有类别信息的邻接矩阵,在降维时,LINPE 既能保持流形的局部几何结构,又能充分利用原始高维数据的类别信息,实现了从无监督到有监督的扩展。通过实验对比可以看出 LINPE 方法在人脸识别上的有效性。在将来的工作中,我们将致力于更好地利用数据的判别信息,与本文提出的方法相结合,以提高算法的鲁棒性的研究。

参考文献

[1] Bellman R. Active control processes: A guided tour [M]. Princeton University Press, 1961

[2] Amador J J. Random projection and orthonormality for lossy image compression [J]. Image and Vision Computing, 2007, 25(5): 754-766

[3] Duda R O, Hart P E, Stork D G. Pattern classification [M]. John Wiley & Sons, 2012

[4] Dzwiniel W, Blasiak J. Method of particles in visual clustering of multi-dimensional and large data sets [J]. Future Generation Computer Systems, 1999, 15(3): 365-379

[5] Novak E, Ritter K. The curse of dimension and a universal method for numerical integration [M]// Multivariate approximation and splines. Birkhäuser Basel, 1997: 177-187

[6] Hyvarinen A. Survey on independent component analysis [J]. Neural Computing Surveys, 1999, 2(4): 94-128

[7] 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

[8] Chang Y L, Han C C, Jou F D, et al. A modular eigen subspace scheme for high-dimensional data classification [J]. Future Generation Computer Systems, 2004, 20(7): 1131-1143

[9] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: A literature survey [J]. Acm Computing Surveys (CSUR), 2003, 35(4): 399-458

[10] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137

[11] Jolliffe I. Principal component analysis [M]. John Wiley & Sons, Ltd, 2005

[12] Belhumeur P N, Hespanha J P, 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

[13] Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation [J]. Neural Computation, 2003, 15(6): 1373-1396

[14] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(5500): 2323-2326

[15] Li B, Zheng C H, Huang D S. Locally linear discriminant embedding: An efficient method for face recognition [J]. Pattern Recognition, 2008, 41(12): 3813-3821

[16] He X F, Partha N. Locality Preserving Projections [C]// Proceedings of the 17th Annual Conference on Neural Information Processing Systems. Vancouver, 2003: 153-160

[17] He X, Cai D, Yan S, et al. Neighborhood preserving embedding [C]// Tenth IEEE International Conference on Computer Vision, 2005 (ICCV 2005). IEEE, 2005, 2: 1208-1213

[18] 杜海顺, 柴秀丽, 汪凤泉, 等. 一种邻域保持判别嵌入人脸识别方法 [J]. 仪器仪表学报, 2010, 31(3): 625-629

[19] Zhang W, Xue X, Lu H, et al. Discriminant neighborhood embedding for classification [J]. Pattern Recognition, 2006, 39(11): 2240-2243

[20] Kifer D, Ben-David S, Gehrke J. Detecting change in data streams [C]// Proceedings of the Thirtieth international conference on Very large data bases - Volume 30. VLDB Endowment, 2004: 180-191

[21] Samaria F S, Harter A C. Parameterisation of a stochastic model for human face identification [C]// Proceedings of the Second IEEE Workshop on Applications of Computer Vision, 1994. IEEE, 1994: 138-142