

一种改进的邻域保持嵌入算法

娄 雪 闫德勤 王博林 王 族

(辽宁师范大学数学学院 辽宁 大连 116029)

摘 要 邻域保持嵌入(NPE)是一种新颖的子空间学习算法,在降维的同时保持了样本集原有的局部邻域流形结构。为了进一步增强 NPE 在人脸识别和语音识别中的识别功能,提出了一种改进的邻域保持嵌入算法(RNPE)。在 NPE 的基础上通过引入类间权值矩阵,使得类间离散度最大,类内离散度最小,增加了样本类间散布约束。最后利用极端学习机(ELM)分类器进行分类,在 Yale 人脸库、Umist 人脸库、Isolet 语音库上的实验结果表明,RNPE 算法的识别率明显高于 NPE 算法、LMMDE 算法以及 RAF-GE 算法。

关键词 邻域保持,邻域嵌入,人脸识别

中图分类号 TP391 文献标识码 A

Improved Neighborhood Preserving Embedding Algorithm

LOU Xue YAN De-qin WANG Bo-lin WANG Zu

(College of Mathematics, Liaoning Normal University, Dalian, Liaoning 116029, China)

Abstract Neighborhood persistence embedding (NPE) is a novel subspace learning algorithm that preserves the original local neighborhood structure of the sample set while maintaining dimensionality. In order to further improve the recognition function of NPE in face recognition and speech recognition, this paper proposed an improved neighborhood preserving embedding algorithm (RNPE). On the basis of NPE, by introducing the interclass weight matrix, the dispersion between classes is the largest, the intra-class dispersion is the smallest, distribution constraint between the classes is increased. The classification experiments are done by the extreme learning machine (ELM) classifier with Yale face database, Umist face database, Isolet speech database. The results show that the recognition rate of RNPE algorithm is significantly higher than NPE algorithm and other traditional algorithms.

Keywords Neighborhood preserving, Neighborhood embedding, Face recognition

1 引言

目前,人脸识别技术和语音识别技术在安检、门禁、反恐、身份鉴别、入口控制等方面具有广阔的应用前景,是计算机识别和模式识别的一个研究热点。但设备采集的面部数据和语音数据往往都具有相关性大、空间维数高的特点,需要将数据从高维空间映射到低维子空间中。在过去的几十年里,子空间分析方法越来越引起人们的注意,成为当前人脸识别和语音识别的主流方法之一,其中最流行的是主成分分析(Principal Component Analysis, PCA)和线性判别分析(Linear Discriminant Analysis, LDA)^[1-3]。

虽然 PCA 和 LDA 在识别功能中有比较成功的应用,但是人脸图像往往会受诸如表情、光照以及皮肤的反射特性等因素的影响,而语音易受场合、情绪、身体状况、语调等因素的影响,它们更可能存在于非线性流形上。最近,研究人员提出了几种流形学习算法可以学习到非线性流形结构^[4-5],如局部线性嵌入算法(Locally Linear Embedding, LLE)、拉普拉斯特征映射算法(Laplacian Eigenmaps, LE)^[6]、等距映射(Isometric

Mapping, Isomap)算法^[7]等。由于降维算法的有效性取决于对人脸图像和语音数据判别信息的有效提取,因此利用这些非线性方法是非常耗时的。同时,对如何构建测试数据的邻接图关系也是未知的,这些算法在应用时无法揭示图像数据的局部结构。

为了克服上述缺点,研究人员提出了邻域保持嵌入(Neighborhood Preserving Embedding, NPE)^[8-9]、局部最大间隔判别嵌入(Local Maximal Margin Discriminant Embedding, LMMDE)、鲁棒的线性降维方法(RAF-Global Embedding, RAF-GE)^[10]等,这些算法能够很好地提取人脸图像数据的判别信息,解决了非线性方法的耗时问题。但其在人脸和语音数据分类问题上仍具有局限性,因为它们在进行线性重构时没有考虑到样本的类间信息。为此,本文提出一种改进的邻域保持嵌入算法(RNPE)。该算法引入类间权值矩阵,并使用极端学习机(ELM)分类器^[11-13]进行分类。基于 Yale 人脸数据库、Umist 人脸数据库和 Isolet 语音数据库的实验结果表明,本文算法是有效的。

本文受国家自然科学基金(61105085)、辽宁省教育厅项目(L2014427)资助。

娄 雪(1993—),女,硕士生,主要研究方向为数据降维、机器学习等;闫德勤(1962—),男,博士,教授,主要研究方向为模式识别、机器学习等;王博林(1993—),女,硕士生,主要研究方向为数据降维、机器学习等;王 族(1992—),女,硕士生,主要研究方向为数据降维、机器学习等。

2 邻域保持嵌入(NPE)

假设高维空间 R^D 中有 N 个样本训练集 $X=(x_1, x_2, \dots, x_N) \in R^D$, 寻求一个最优的映射变换矩阵 A , 将 $x_i \in R^D (i=1, 2, \dots, N)$ 嵌入映射到一个相对低维的空间 $R^d (d \leq D)$ 中。记 $y_i \in R^d$, 将嵌入映射数据集表示为 $Y=(y_1, y_2, \dots, y_N)^T \in R^d$ 。

事实上, NPE 在构建高维空间的数据集的近邻图这一步骤与 LLE 方法相似。在计算权重时, NPE 假设每个局部近邻是线性的, 有 $y_i = A^T x_i$ 。这样, 高维空间 R^D 中的每个点 x_i 都可以通过其 k 个最近邻线性重构, 损失函数为:

$$\varphi(W) = \sum_i \|x_i - \sum_j W_{ij} x_j\|^2 \quad (1)$$

式(1)的限制条件为 $\sum_j W_{ij} = 1, j=1, 2, \dots, N$ 。同样, 在 NPE 中 W_{ij} 也可以在 R^d 中重构与 x_i 相对应的嵌入向量 y_i , 因此有:

$$\varphi(y) = \sum_i \|y_i - \sum_j W_{ij} y_j\|^2 \quad (2)$$

因为在 NPE 中假设变换是线性的, 有 $y_i = A^T x_i$, 则 $\varphi(y)$ 可以化简为:

$$\varphi(y) = A^T X M X^T A \quad (3)$$

其中, $M=(I-W)^T(I-W)$, $I=diag(1, 1, \dots, 1)$, 且限制条件为 $A^T X X^T A = I$, 而且式(3)的最优化问题可以转化成以下广义特征值求解问题, 即 $X M X^T A = \lambda X X^T A$, 它的广义特征值为 $\lambda_0, \lambda_1, \dots, \lambda_{d-1}$ (其中 $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{d-1}$), 相应的广义特征向量为 $\alpha_0, \alpha_1, \dots, \alpha_{d-1}$, 因此 $x_i \rightarrow y_i = A^T x_i$, 其中 $A=(\alpha_0, \alpha_1, \dots, \alpha_{d-1})$ 。

3 改进的邻域保持嵌入(RNPE)

3.1 RNPE 算法的基本思想

假设高维空间 R^D 中有 N 个样本训练集 $X=(x_1, x_2, \dots, x_N) \in R^D$, 嵌入映射数据集表示为 $Y=(y_1, y_2, \dots, y_N)^T \in R^d (d \leq D)$, 我们将训练集 X 分成 c 类, 记 $X=(X_1, X_2, \dots, X_c)$, 因此相应的嵌入映射数据集为 $Y=(Y_1, Y_2, \dots, Y_c)^T$, RNPE 的目标函数为:

$$\min J(W) = \sum_{i=c} \sum_{i=1}^n \|y_i^c - \sum_{j=1}^n W_{ij}^c y_j^c\|^2 - \lambda \sum_{c=1}^c \|u_c - \sum_{k=1}^k S_{ck} u_k\|^2 \quad (4)$$

式(4)的限制条件为 $\sum_{j=1}^n W_{ij} = 1, \sum_{k=1}^k S_{ck} = 1$ 。其中, λ 为权衡参数; $y_i^c (y_j^c)$ 为第 c 类的第 $i (j)$ 个嵌入向量; n_c 为第 c 类的嵌入向量的总数; W_{ij}^c 为 y_i^c 与 y_j^c 之间的权重; $u_c (u_k)$ 为第 $c (k)$ 类的嵌入向量的平均值, $u_c = \frac{1}{n_c} \sum_{i=1}^n y_i^c, u_k = \frac{1}{n_k} \sum_{i=1}^n y_i^k; S_{ck}$ 为第 c 类和第 k 类之间的权重。

接下来, 化简目标函数。由 NPE 思想, 每个局部邻域结构都是线性的, 有 $y_i = A^T x_i$, 从而 $Y^T = A^T X, Y_c^T = A^T X_c$ 。我们定义 $z_i^c = y_i^c - \sum_{j=1}^n W_{ij}^c y_j^c, Y_c = (Y_1^c, Y_2^c, \dots, Y_n^c)^T, Z_c = (z_1^c, z_2^c, \dots, z_{n_c}^c)^T, W_c =$

$$W_c = \begin{bmatrix} W_{11}^c & W_{12}^c & \dots & W_{1n_c}^c \\ W_{21}^c & W_{22}^c & \dots & W_{2n_c}^c \\ \vdots & \vdots & \ddots & \vdots \\ W_{n_c 1}^c & W_{n_c 2}^c & \dots & W_{n_c n_c}^c \end{bmatrix}, \text{因此 } Z_c =$$

$Y_c - W_c Y_c = (I - W_c) Y_c$, 其中 $I = diag(1, 1, \dots, 1)$ 。

被减数可化简为:

$$\begin{aligned} & \sum_{i=c} \sum_{i=1}^n \|y_i^c - \sum_{j=1}^n W_{ij}^c y_j^c\|^2 \\ &= \sum_{c=1}^c \sum_{i=1}^n (z_i^c)^2 = \sum_{c=1}^c (Z_c)^2 \\ &= \sum_{c=1}^c Z_c^T Z_c \\ &= \sum_{c=1}^c Y_c^T (I - W_c)^T (I - W_c) Y_c \\ &= \sum_{c=1}^c Y_c^T H_1 Y_c = \sum_{c=1}^c A^T X_c H_1 X_c^T A = A^T X H_1 X^T A \quad (5) \end{aligned}$$

其中, $H_c = (I - W_c)^T (I - W_c), H_1 = diag(H_1, H_2, \dots, H_c)$ 。

减数可化简为:

$$\begin{aligned} & \sum_{c=1}^c \|u_c - \sum_{k=1}^k S_{ck} u_k\|^2 \\ &= \sum_{c=1}^c \|m_c^T A - \sum_{k=1}^k S_{ck} m_k^T A\|^2 \\ &= \|M^T A - S M^T A\|^2 = \|(I - S) M^T A\|^2 \\ &= A^T M (I - S)^T (I - S) M^T A = A^T M H_2 M^T A \quad (6) \end{aligned}$$

其中, $m_c (m_k)$ 为 R^D 中第 $c (k)$ 类样本数据集的平均值, 即

$$m_c = \frac{1}{n_c} \sum_{i=1}^n x_i^c, m_k = \frac{1}{n_k} \sum_{i=1}^n x_i^k; u_c^T = A^T m_c, u_k^T = A^T m_k; M = diag$$

$$(M_1, M_2, \dots, M_c); S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1k} \\ S_{21} & S_{22} & \dots & S_{2n_c} \\ \vdots & \vdots & \ddots & \vdots \\ S_{c1} & S_{c2} & \dots & S_{ck} \end{bmatrix}; H_2 = (I - S)^T$$

$(I - S)$ 为样本数据集 X 的类间离散度矩阵。

因此, RNPE 的目标函数最终化简为:

$$\begin{aligned} \min J(W) &= \sum_{i=c} \sum_{i=1}^n \|y_i^c - \sum_{j=1}^n W_{ij}^c y_j^c\|^2 - \lambda \sum_{c=1}^c \|u_c - \sum_{k=1}^k S_{ck} u_k\|^2 \\ &= A^T X H_1 X^T A - \lambda A^T M H_2 M^T A \\ &= A^T S_1 A - \lambda A^T S_2 A = A^T (S_1 - \lambda S_2) A \quad (7) \\ A_{opt} &= \arg \min_A [\sum_{i=c} \sum_{i=1}^n \|y_i^c - \sum_{j=1}^n W_{ij}^c y_j^c\|^2 - \lambda \sum_{c=1}^c \|u_c - \sum_{k=1}^k S_{ck} u_k\|^2] \\ &= \arg \min_A \text{tr}(A^T (S_1 - \lambda S_2) A) \quad (8) \end{aligned}$$

其中, $S_1 = X H_1 X^T, S_2 = M H_2 M^T$, 限制条件为 $A^T X X^T A = 1$ 。

目标函数的被减数项反映了类内的变化, 减数项反映了类间的变化, 因此目标函数的极小化就是在权重一定的情况下, 类内局部线性重构误差最小, 类内距离达到最小, 类中心之间局部线性的重构误差最小, 类间距离最大。

由于 λ 只是一个权衡参数, 在理论上我们极小化目标函数时只考虑 $A^T (S_1 - S_2) A$, 且矩阵 $S_1 - S_2$ 是对称的, 采用拉格朗日乘法将目标函数极小化: $L(A) = A^T (S_1 - S_2) A - \lambda (A^T X X^T A - 1)$, 对此式中 A 的各分量求偏导并令其为零, 可确定一个最优的变换矩阵 A 。令 $\frac{\partial L(A)}{\partial (A)} = 2(S_1 - S_2) A - 2\lambda X X^T A = 0$ 。容易看出, 要使 A 达到最优, A 必须满足: $(S_1 - S_2) A = \lambda X X^T A$ 。求得 $(X X^T)^{-1} (S_1 - S_2)$ 的广义特征值为 $\lambda_0, \lambda_1, \dots, \lambda_{d-1}$ (其中 $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{d-1}$), 其相对应的特征向量为 $\alpha_0, \alpha_1, \dots, \alpha_{d-1}, A = (\alpha_0, \alpha_1, \dots, \alpha_{d-1}), A$ 为 $D \times d$ 维的矩阵, 因此, $Y_{d \times N} = A_{d \times D}^T \times X_{D \times N}$ 。

RNPE 算法的过程描述如下。

给定数据集 $X=(x_1, x_2, \dots, x_N) \in R^D$:

步骤 1 根据类标签信息,计算每一类样本的权重矩阵 W_c 。

步骤 2 特征映射。通过广义特征值问题来计算线性变换矩阵 A ,且 A 必须满足 $(S_1 - S_2)A = \lambda XX^T A$,那么从原始样本空间到嵌入特征空间的最终变换矩阵 $A_{opt} = \arg \min_A \text{tr}(A^T(S_1 - \lambda S_2)A)$ 。

步骤 3 线性降维 $Y = A_{opt}^T \times X$ 。

3.2 RNPE 算法的优点与应用

在人脸识别^[14]和语音识别^[15]技术中,设备采集的面部和语音数据往往具有相关性大、空间维数高的特点,正因如此,在这个过程中会出现两个难题:1)训练样本数很少,这种小样本问题有时候会导致类间离散度矩阵 H_2 是奇异的,而 RNPE 采用的是减式而非除式,恰好避开了这一问题;2)NPE 及其他传统的降维算法没有考虑到样本的类间信息,而 RNPE 引入了类间的权值,增加了类间的离散度约束,使投影后的样本具有最佳的分离性,确保了数据的局部流形结构特征。

极端学习机作为单隐层前向神经网络的一种典型学习算法,与传统的方法不同,它可以随机地选择网络中的隐层神经元,并且其网络的输出层权值可以解析获得^[17]。该方法具有许多优良的特性,如学习速度快、泛化能力好等,因此 RNPE 在人脸识别和语音识别的过程中运用 ELM 分类比其他分类器的分类效果更加理想,具体算法步骤如下:

步骤 1 输入高维空间 R^D 中的训练集样本 X 。利用 RNPE 算法提取数据集特征并对其降维,得到该训练集的低维表示 Y 。

步骤 2 随机指定网络输入权值和偏置值 b_j 。

步骤 3 通过激活函数计算隐层节点输出矩阵 H 。

步骤 4 计算出权值 β 。

步骤 5 输入测试集样本进行识别。

4 实验结果及分析

由于极端学习机激活函数的选择具有问题依赖性,对于不同的问题,选择合适的激活函数对提高神经网络的泛化性能是有利的。为了验证 RNPE 算法的有效性,基于 Yale 人脸库、Umist 人脸库以及 Isolet 语音库进行实验,其中高斯 (agauss) 激活函数的性能表现良好,相较于其他激活函数如 sigmoid 函数、sine 函数、hardlim 函数,其训练误差和测试误差都较小。

除激活函数的适当选择外,训练样本的数量对神经网络泛化能力的影响也非常大,经过本文多次实验,在这 3 个数据库中随机在每类样本中选取的训练个数为 $TrainNum = \{3, 4, 5, 6\}$,其余为测试集,其比选取其他训练样本数量的网络泛

化性能更加理想。

ELM 的结果具有较大的随机性,将 RNPE 算法与 NPE 算法、LMMDE 算法以及 RAF-GE 算法进行比较,在每个数据集上重复多次实验并进行统计检验。本文以在每个数据集上实验两次为例,并将对 $TrainNum$ 检验对比。ELM 运用 NPE 算法、LMMDE 算法以及 RAF-GE 算法在不同维数下的识别率曲线如图 1—图 3 所示,其识别率如表 1—表 3 所列。不同的数据库参数设置如表 4 所列,Yale 和 Umist 图像数据如图 4 所示。

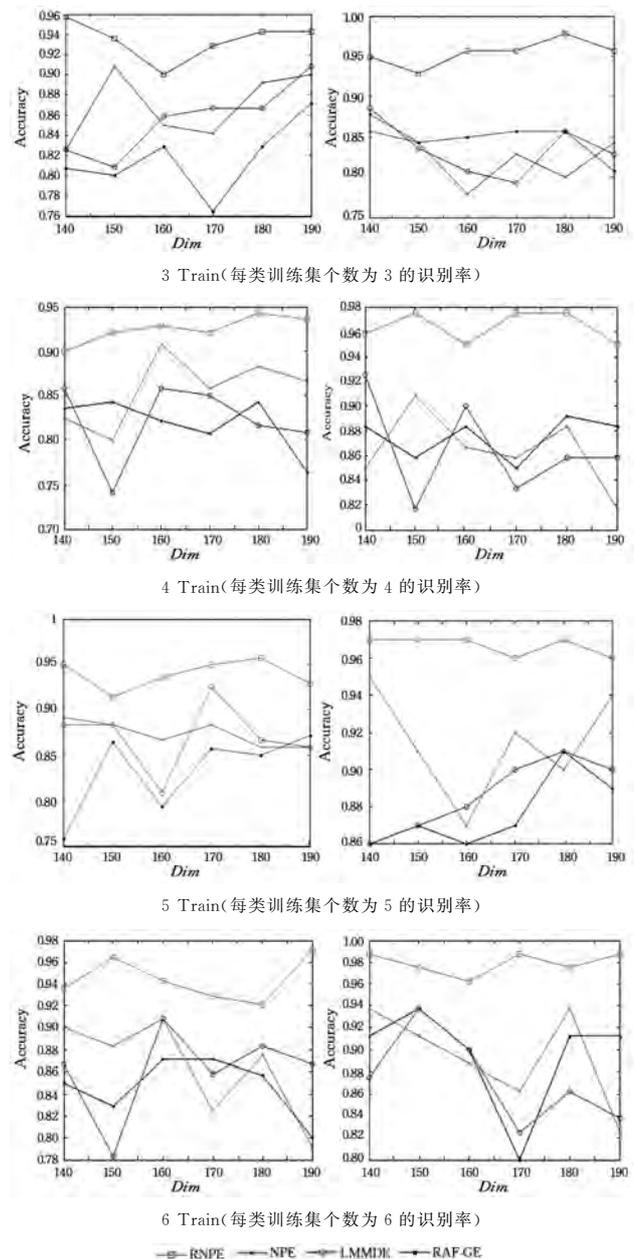


图 1 ELM 运用不同降维算法在 Umist 人脸库上的识别率曲线

表 1 不同降维算法在 Umist 人脸库中的识别率(平均值±相对误差)

Algorithms	Umist							
	3 Train		4 Train		5 Train		6 Train	
	1	2	1	2	1	2	1	2
RNPE+ELM	93.45±1.94	95.48±1.61	92.50±1.48	96.39±1.25	93.93±1.61	96.67±0.52	94.40±1.99	97.92±1.02
NPE+ELM	86.94±3.48	82.50±2.99	85.69±3.92	86.39±3.10	87.36±1.44	91.50±2.88	86.39±4.58	89.37±4.45
LMMDE+ELM	85.56±3.52	83.45±3.36	82.22±4.49	86.53±4.06	87.08±3.83	88.67±1.97	86.11±4.21	87.29±4.14
RAF-GE+ELM	81.67±3.58	84.88±2.37	81.90±3.02	87.50±1.67	83.21±4.62	87.67±1.97	84.64±2.78	89.58±4.85

(单位:%)

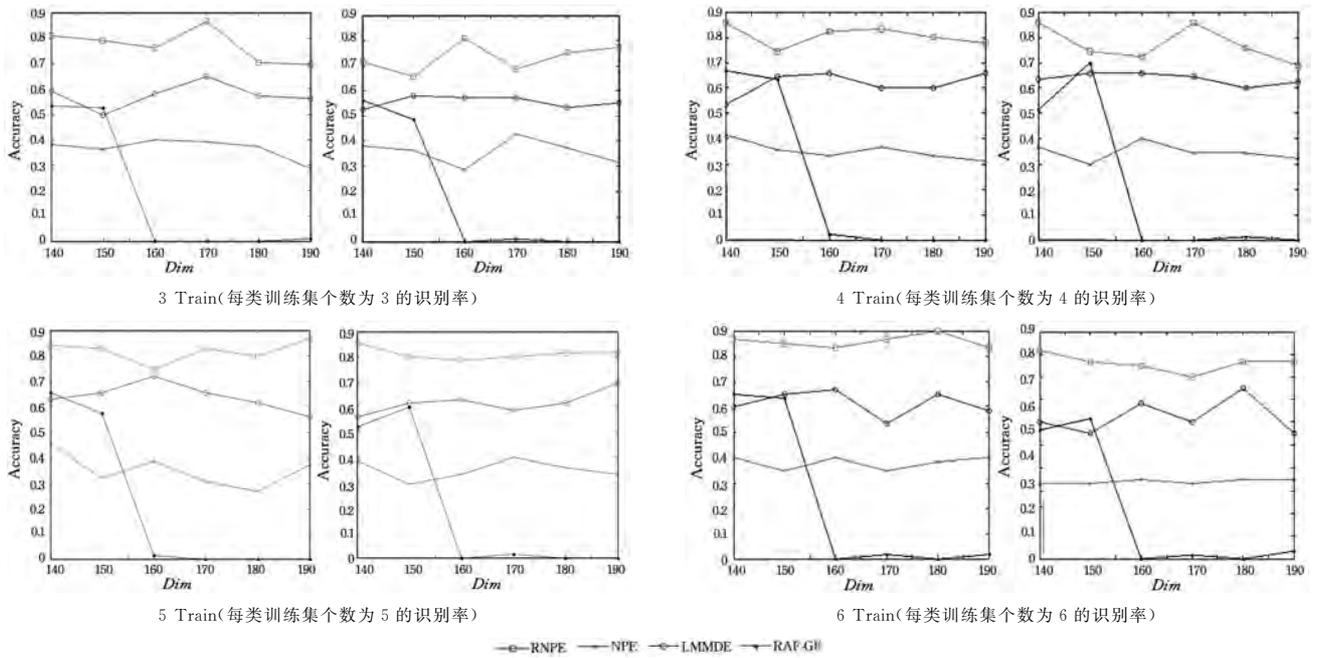


图2 ELM运用不同降维算法在Yale人脸库上的识别率曲线

表2 不同降维算法在Yale人脸库中的识别率(平均值±相对误差)

(单位:%)

Algorithms	Yale							
	3 Train		4 Train		5 Train		6 Train	
	1	2	1	2	1	2	1	2
RNPE+ELM	75.79±6.87	73.17±5.66	80.56±4.02	77.04±6.98	81.78±4.10	81.11±2.30	85.83±2.53	0.8611±0.0375
NPE+ELM	36.51±4.12	35.71±5.07	35.19±3.49	34.63±3.47	35.11±6.67	35.11±3.93	38.06±2.45	0.3417±0.0091
LMMDE+ELM	57.46±4.92	55.56±2.31	61.48±4.75	63.52±2.16	63.78±5.29	61.56±4.50	61.39±5.10	0.6222±0.0793
RAF-GE+ELM	17.78±27.20	17.62±27.10	22.04±33.30	20.37±31.70	24.53±33.70	18.89±28.90	21.95±32.70	0.2056±0.2998

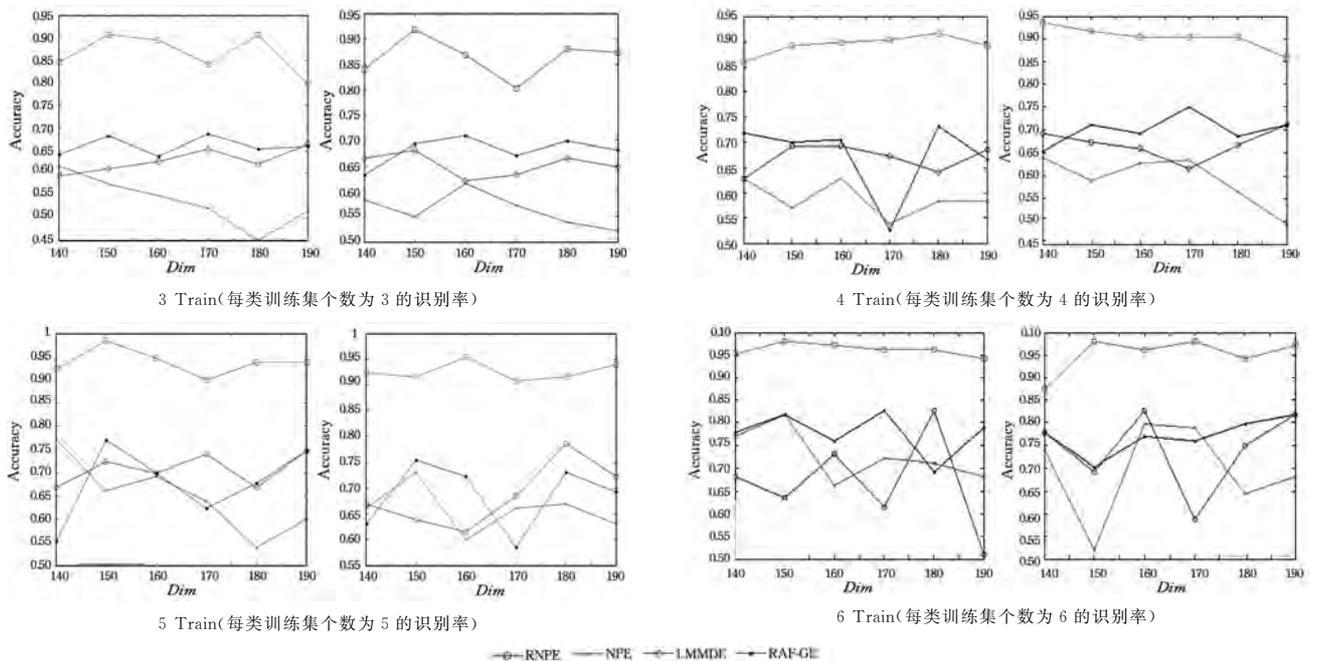


图3 ELM运用不同降维算法在Isolet语音库中的识别率曲线

表3 不同降维算法在Isolet人脸库中的识别率(平均值±相对误差)

(单位:%)

Algorithms	Isolet							
	3 Train		4 Train		5 Train		6 Train	
	1	2	1	2	1	2	1	2
RNPE+ELM	86.54±4.47	86.36±3.89	89.32±1.93	90.38±2.53	93.85±2.79	92.56±1.73	96.15±1.36	95.19±4.03
NPE+ELM	53.85±5.66	56.32±3.36	58.87±3.47	59.19±5.66	65.00±7.90	65.90±4.37	72.76±5.69	69.55±10.50
LMMDE+ELM	62.82±2.68	65.20±2.27	66.88±2.77	66.99±3.26	70.77±3.37	68.59±6.10	66.67±10.80	74.20±9.05
RAF-GE+ELM	66.03±2.01	68.04±2.72	67.41±7.59	70.09±3.21	67.82±7.99	68.59±6.53	77.72±4.84	77.08±3.96

(下转第278页)

的选择。本文的研究成果仍然可以应用于此类问题的后期资源选择中。

参考文献

- [1] 沈玉. MF-TDMA 卫星通信系统网络规划技术研究[J]. 无线电通信技术, 2014, 40(4): 11-14.
- [2] 李学远, 回征, 朱历洪, 等. 卫星数据链战场应用的规划问题研究[J]. 空间电子技术, 2015, 12(5): 51-54.
- [3] 吕智勇, 汪宏武, 张健. 卫星转发器资源的选择[J]. 卫星与网络, 2007(4): 36-41.

- [4] 巴红波, 王佳嘉, 吴建军. 多波束卫星移动通信频率规划方案研究[C]//第9届卫星通信学术年会论文集. 2013: 361-366.
- [5] 郭小海. 多层卫星网络信道资源管理策略应用研究[D]. 南京: 南京邮电大学, 2016.
- [6] 孙艳峰. 宽带卫星通信系统资源管理策略研究[D]. 北京: 北京邮电大学, 2013.
- [7] FIGUEIRA J, GRECO S, EHRGOTT M, et al. Multiple Criteria Decision Analysis: State of the Art Surveys[M]. London: Springer Science, 2005.
- [8] HWANG C L, YOON K. Multiple Attributes Decision Making Methods and Applications[M]. Berlin: Springer, 1981.

(上接第 258 页)



(a) Yale database



(b) Umist database

图 4 不同图像数据

表 4 数据集参数描述

Datasets	Dim	Samples	samples/subject	Classes
Yale	1024	165	11	15
Umist10304	575	18	20	
Isolet	617	1560	60	26

图 1—图 3 分别给出了 ELM 运用不同降维算法在 Yale 人脸库、Umist 人脸库和 Isolet 语音库上的识别率曲线。由图 1—图 3 可以看出本文提出的 RNPE 算法的识别率曲线明显高于 NPE 算法、LMMDE 算法以及 RAF-GE 算法的识别率曲线。这是由于类间权值矩阵的引入增加了样本类间的散布约束, 确保了数据的局部流形结构特征, 使得该算法在针对不同数据时具有很高的稳定性。由图 1—图 3 所示, 可以看出在这 3 组数据库中随机在每类样本中选取的训练个数为 $TrainNum = \{3, 4, 5, 6\}$, 其余为测试集, ELM 运用 RNPE 算法在 140~190 维数下的识别率曲线远远高于 NPE 算法、LMMDE 算法以及 RAF-GE 算法。由图 1、表 1 所示, NPE 的算法识别率达到 90% 以上甚至接近于 100%; 由图 2、图 3, 表 2、表 3 所示, NPE 算法的识别率大多数都达到 80% 以上甚至 90% 以上, 并且随着训练个数的增多, 识别效果会更好。

结束语 本文提出了一种改进的邻域保持嵌入算法 (RNPE), 在 NPE 的基础上, 引入了类间的权值, 增加了类间的离散度约束。该算法使投影后的样本具有最佳的分异性, 确保了数据的局部流形结构特征, 目的是在降维的同时保持样本集原有的局部邻域流形结构不变。通过 Yale 人脸库、Umist 人脸库、Isolet 语音库的实验结果表明, RNPE 算法在识别率上明显高于 NPE 算法和其他传统算法。

参考文献

- [1] 李瑞敏, 陆化普. 基于 WebGIS 的智能交通管理指挥调度系统[J]. 计算机工程, 2007, 33(21): 232-234.

- [2] SCHMITT E J, JULA H. Vehicle Route Guidance Systems: Classification and Comparison[C]// Proc. of IEEE Intelligent Transportation Systems Conference. 2006: 242-247.
- [3] 李旭, 舒薇, 李铭璐, 等. 车载传感器网络中的流量数据处理[C]//第17届计算机通信与网络国际会议. IEEE Press, 2008: 1-5.
- [4] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [5] SEUNG H S, LEE D D. Cognition—the manifold ways of perception[J]. Science, 2000, 290(5500): 2268.
- [6] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural Computation, 2003, 15(6): 1373-1396.
- [7] TENENBAUM J B, DE SILVA V, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290(5500): 2319-2324.
- [8] SONG B, TAN S, SHI H. Process monitoring via enhanced neighborhood preserving embedding[J]. Control Engineering Practice, 2016, 50: 48-56.
- [9] LIANG J Z, CHEN C, YI Y F, et al. Bilateral Two-Dimensional Neighborhood[C]// Preserving Discriminant Embedding for Face Recognition. 2017: 17201-17212.
- [10] 冯林, 刘胜蓝, 张晶, 等. 高维数据中鲁棒激活函数的极端学习机及线性降维[J]. 计算机研究与发展, 2014, 51(6): 1331-1340.
- [11] BENOIT F, VAN HEESWIJK M, MICHE Y, et al. Feature selection for nonlinear models with extreme learning machines[J]. Neurocomputing, 2013, 102(2): 111-124.
- [12] PENG Y, LU B L. Discriminative graph regularized extreme learning machine and its application to face recognition[J]. Neurocomputing, 2015, 149(PA): 340-353.
- [13] HUANG G B. An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels[J]. Cognitive Computation, 2014, 6(3): 376-390.
- [14] 戴礼荣, 张仕良, 黄智颖. 基于深度学习的语音识别技术现状与展望[J]. 数据采集与处理, 2017, 32(2): 221-231.
- [15] TANG J, DENG C, HUANG G B. Extreme learning machine for multilayer perceptron[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 27(4): 809-821.