

DOI: 10.11992/tis.201808007

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20181223.1553.002.html>

面向局部线性回归分类器的判别分析方法

朱换荣, 郑智超, 孙怀江

(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

摘 要: 局部线性回归分类器 (locality-regularized linear regression classification, LLRC) 在人脸识别上表现出了高识别率以及高效性的特点, 然而原始特征空间并不能保证 LLRC 的效率。为了提高 LLRC 的性能, 提出了一种与 LLRC 相联系的新的降维方法, 即面向局部线性回归分类器的判别分析方法 (locality-regularized linear regression classification based discriminant analysis, LLRC-DA)。LLRC-DA 根据 LLRC 的决策准则设计目标函数, 通过最大化类间局部重构误差并最小化类内局部重构误差来寻找最优的特征子空间。此外, LLRC-DA 通过对投影矩阵添加正交约束来消除冗余信息。为了有效地求解投影矩阵, 利用优化变量之间的关系, 提出了一种新的迹比优化算法。因此 LLRC-DA 非常适用于 LLRC。在 FERET 和 ORL 人脸库上进行了实验, 实验结果表明 LLRC-DA 比现有方法更具有优越性。

关键词: 局部线性回归分类器; 维数约简; 正交投影; 迹比问题; 人脸识别; 特征提取; 判别分析; 线性回归分类器

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2019)05-0959-07

中文引用格式: 朱换荣, 郑智超, 孙怀江. 面向局部线性回归分类器的判别分析方法 [J]. 智能系统学报, 2019, 14(5): 959-965.

英文引用格式: ZHU Huanrong, ZHENG Zhichao, SUN Huaijiang. Locality-regularized linear regression classification-based discriminant analysis[J]. CAAI transactions on intelligent systems, 2019, 14(5): 959-965.

Locality-regularized linear regression classification-based discriminant analysis

ZHU Huanrong, ZHENG Zhichao, SUN Huaijiang

(College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Locality-regularized linear regression classification (LLRC) based face recognition achieves high accuracy and high efficiency. However, the original feature space cannot guarantee the efficiency of LLRC. To improve the performance of LLRC, this study proposes a new dimensionality reduction method called locality-regularized linear regression classification-based discriminant analysis (LLRC-DA), which is directly associated with LLRC. The objective function of LLRC-DA is designed according to the classification rule of LLRC. In LLRC, interclass local reconstruction errors are maximized and simultaneously, intraclass local reconstruction errors are minimized to identify the optimal feature subspace. In addition, LLRC-DA eliminates redundant information using an orthogonal constraint, imposed on the projection matrix. To effectively obtain the solutions of the projection matrix, we exploit the relationship between optimal variables and propose a new trace ratio optimization method. Hence, LLRC-DA suits LLRC well. Extensive experimental results obtained from the FERET and ORL face databases demonstrate the superiority of the proposed method than state-of-the-art methods.

Keywords: locality-regularized linear regression classification; dimensionality reduction; orthogonal projections; trace ratio problem; face recognition; feature extraction; discriminant analysis; linear regression classification

收稿日期: 2018-08-09. 网络出版日期: 2018-12-26.

基金项目: 国家自然科学基金项目 (61772272).

通信作者: 朱换荣. E-mail: zhuhuanrong@foxmail.com.

维数约简是帮助我们理解数据特征结构的有效工具, 被广泛地应用于人脸识别^[1-3]、图像检

索^[4-5]、指纹认证^[6-7]、生物信息学^[8-9]、数据挖掘^[10-11]等。维数约简的目标是减少冗余、不相关的信息和噪声,同时保留数据的潜在结构和本征信息。文献[12]显示,虽然随着样本维数的增加,分类器的性能会逐步上升,但是过高的维数却会使分类器的性能急剧下降,甚至造成“维数灾难”问题。因此,维数约简对于分类具有重要的意义。在过去的几十年中,针对维数约简问题出现了各种不同的解决方案和算法,其中主成分分析^[13](principal component analysis, PCA)和线性判别分析^[14](linear discriminant analysis, LDA)是最具代表性的两个线性降维算法。PCA的思想是最大化投影空间中样本之间的差异。LDA的目标是找到一组投影轴,使得在投影空间中同类样本尽可能靠近,而不同类样本尽可能分开。LDA是一种监督算法,充分利用了样本的类别信息来发现数据的特征结构,因此在维数约简方面通常比PCA更加有效,但是LDA不能很好地处理非线性分类问题。流形学习^[15](manifold learning)是一种非线性的维数约简方法。流形算法的主要思想是学习高维空间中样本的局部邻域结构,并寻找能够保留这种流形结构的子空间,使得样本投影到子空间后有着较好的局部近邻关系。其中局部保持投影^[15](locality preserving projections, LPP)就是最常见的流形学习算法之一。LPP方法得到的投影矩阵,可以使样本在投影之后保留邻域结构。

近年来,基于表示的分类方法引起了人们的广泛关注,稀疏表示分类器^[16](sparse representation based classification, SRC)是最具代表性的基于表示的分类方法之一。SRC中测试样本用训练样本的稀疏线性组合来表示,然后计算测试样本与每类样本之间的重构误差,最终将测试样本分类到重构误差最小的类别。SRC表现出了良好的分类性能,但是由于SRC计算稀疏表示系数需要求解 l_1 范数,所以花费的时间较多。而线性回归分类器^[17](linear regression classification, LRC)则很好地避免了这一问题。LRC假设同一类的样本处于相同的线性子空间中。因此LRC分别计算测试样本在不同类的子空间中的投影,然后通过计算最小重构误差来判别测试样本的类别。LRC利用最小二乘法估算回归系数,所以LRC是一个简单而有效的分类方法。然而,在许多现实应用中,样本通常表现出局部线性,而不是全局线性。于是Brown等^[18]提出了局部线性回归分类器(local-

ity-regularized linear regression classification, LLRC),将流形学习系统嵌入到LRC中,使用流形学习来扩展传统的LRC,从而提高分类精度。

为了防止过高的维数影响LLRC的分类性能并且保留数据的潜在结构,所以在使用原始数据时先进行降维处理。大多数的降维方法独立于分类器的决策准则,因此分类器往往不能有效地利用学习到的特征子空间。为了更好地将维数约简与特定的分类器相联系,根据分类器的决策准则来设计维数约简算法是很有必要的。于是,Huang等^[19]根据LLRC的决策准则提出了局部线性回归判别分析(locality-regularized linear regression discriminant analysis, LLRDA)。LLRDA根据LLRC的决策准则分别定义了类内局部重构误差和类间局部重构误差,然后使用最大特征值分解方法求解一个迹比问题以得到最优解。然而,在LLRDA算法中,线性回归系数在原始空间进行计算并且假定系数恒定,但是实际上回归系数与投影矩阵是相关的,不是独立的,在求解目标函数的过程中投影矩阵的改变会对回归系数产生影响,所以不能假定系数恒定。另外,LLRDA的目标函数是一个迹比问题,但是为了便于处理,LLRDA将迹比问题隐式地转换成了比迹问题,并用最大特征值分解方法进行求解,因而得到的结果会与原来的目标函数存在偏差。

为了克服LLRDA的缺点,本文提出了面向局部线性回归分类器的判别分析方法(locality-regularized linear regression classification based discriminant analysis, LLRC-DA)。LLRC-DA根据LLRC的决策准则设计目标函数,通过最大化类间局部重构误差并最小化类内局部重构误差来寻找最优的特征子空间。在求解目标函数的过程中,本文利用投影矩阵与线性回归系数之间的关系,使用解析解表示线性回归系数,从而消除了回归系数对求解投影矩阵产生的影响。另外,在求解目标函数过程中使用了一种新的迹比优化方法,直接解决迹比问题而不是使用广义特征值分解的方法近似求解。文献[20]的研究显示,直接解决迹比问题是可行的并且比使用广义特征值分解效果更好。在本文提出的方法中,使用了正交投影来消除冗余信息,提高性能。本文在FER-ET、ORL人脸库上进行了实验,实验结果验证了算法的有效性。

1 局部线性回归分类器

假设有 n 个训练样本, 共有 C 类, 其中第 i 类有 n_i 个训练样本, 所以总共有 $\sum_{i=1}^C n_i = n$ 个样本。训练样本矩阵为 $\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_C] \in \mathbf{R}^{N \times n}$, 其中第 i 类训练样本矩阵为 $\mathbf{X}_i = [\mathbf{x}_i^1 \mathbf{x}_i^2 \cdots \mathbf{x}_i^{n_i}] \in \mathbf{R}^{N \times n_i}$, $i = 1, 2, \dots, C$, N 表示样本维数。对于一个测试样本 $\mathbf{y} \in \mathbf{R}^{N \times 1}$, 可以通过 k -最近邻算法计算出其在第 i 类训练样本 \mathbf{X}_i 中的局部线性子集 $\tilde{\mathbf{X}}_i = [\mathbf{x}_i^1 \mathbf{x}_i^2 \cdots \mathbf{x}_i^k] \in \mathbf{R}^{N \times k}$, 在 k -最近邻算法中, 通常使用欧氏距离作为距离度量方法, 则 \mathbf{y} 的线性表示记作:

$$\mathbf{y} = \tilde{\mathbf{X}}_i \tilde{\boldsymbol{\beta}}_i \quad (1)$$

式中: $\tilde{\boldsymbol{\beta}}_i \in \mathbf{R}^{k \times 1}$ 表示第 i 类样本的线性回归系数, $\tilde{\boldsymbol{\beta}}_i$ 可以使用最小二乘估计来计算:

$$\tilde{\boldsymbol{\beta}}_i = (\tilde{\mathbf{X}}_i^T \tilde{\mathbf{X}}_i)^{-1} \tilde{\mathbf{X}}_i^T \mathbf{y} \quad (2)$$

根据线性回归系数 $\tilde{\boldsymbol{\beta}}_i$, 测试样本 \mathbf{y} 在第 i 类样本上的线性重构可以表示为

$$\hat{\mathbf{y}}_i = \tilde{\mathbf{X}}_i (\tilde{\mathbf{X}}_i^T \tilde{\mathbf{X}}_i)^{-1} \tilde{\mathbf{X}}_i^T \mathbf{y} \quad (3)$$

根据 LLRC 的假设, 如果重构样本 $\hat{\mathbf{y}}_i$ 和测试样本 \mathbf{y} 的距离最近, 则 \mathbf{y} 应该属于第 i 类, 所以 LLRC 的决策准则为

$$l(\mathbf{y}) = \min_i \left\| \mathbf{y} - \tilde{\mathbf{X}}_i (\tilde{\mathbf{X}}_i^T \tilde{\mathbf{X}}_i)^{-1} \tilde{\mathbf{X}}_i^T \mathbf{y} \right\|_2^2 \quad (4)$$

2 LLRC-DA 算法

2.1 LLRC-DA 提出

LLRC-DA 的目标是将样本投影到一个特征子空间, 使得类内局部重构误差最小, 同时类间局部重构误差最大。用 $\mathbf{A} \in \mathbf{R}^{N \times d}$ ($d < N$) 表示优化投影矩阵, 原始空间中的训练样本矩阵 $\mathbf{X} \in \mathbf{R}^{N \times n}$ 投影到 d 维子空间为 $\mathbf{Y} = \mathbf{A}^T \mathbf{X} \in \mathbf{R}^{d \times n}$, 即每个训练样本 $\mathbf{x}_i^j \in \mathbf{R}^N$ 投影为 $\mathbf{y}_i^j = \mathbf{A}^T \mathbf{x}_i^j \in \mathbf{R}^d$, 其中 \mathbf{x}_i^j 表示第 i 类的第 j 个训练样本。线性回归系数 $\tilde{\boldsymbol{\beta}}$ 是通过式 (2) 来计算得到的。

从式 (4) 可以看出, 如果 \mathbf{y}_i^j 用第 i 类训练样本的局部线性子集来表示, 可以得到最小的重构误差, 则 \mathbf{y}_i^j 可以被 LLRC 正确分类, 所以本文希望找到一个特征子空间, 在这个子空间中, \mathbf{y}_i^j 用第 i 类训练样本表示时的局部重构误差尽可能地小, 而用其他类训练样本表示时的局部重构误差尽可能地大。为了达到这样的目标, 本文定义了类内局部重构误差和类间局部重构误差。其中, 类内局部重构误差为

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} \left\| \mathbf{y}_i^j - \tilde{\mathbf{Y}}_i^j \tilde{\boldsymbol{\beta}}_i^j \right\|_2^2 = \\ & \text{tr} \left(\frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} (\mathbf{y}_i^j - \tilde{\mathbf{Y}}_i^j \tilde{\boldsymbol{\beta}}_i^j) (\mathbf{y}_i^j - \tilde{\mathbf{Y}}_i^j \tilde{\boldsymbol{\beta}}_i^j)^T \right) = \\ & \text{tr} \left(\frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} (\mathbf{A}^T \mathbf{x}_i^j - \mathbf{A}^T \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j) (\mathbf{A}^T \mathbf{x}_i^j - \mathbf{A}^T \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j)^T \right) = \\ & \text{tr} \left(\frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} \mathbf{A}^T (\mathbf{x}_i^j - \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j) (\mathbf{x}_i^j - \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j)^T \mathbf{A} \right) = \text{tr}(\mathbf{A}^T \mathbf{S}_w \mathbf{A}) \end{aligned} \quad (5)$$

式中: $\mathbf{S}_w = \frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} (\mathbf{x}_i^j - \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j) (\mathbf{x}_i^j - \tilde{\mathbf{X}}_i^j \tilde{\boldsymbol{\beta}}_i^j)^T$ 是类内局部重构散列矩阵; $\tilde{\mathbf{X}}_i^j$ 是通过 k -最近邻算法计算出的 \mathbf{x}_i^j 在第 i 类训练样本 \mathbf{X}_i 中的局部线性子集; $\tilde{\boldsymbol{\beta}}_i^j$ 表示 \mathbf{x}_i^j 用第 i 类样本的局部线性子集 $\tilde{\mathbf{X}}_i^j$ 表示时的线性回归系数。

同样地, 在特征子空间中, 可以计算类间局部重构误差, 即

$$\begin{aligned} & \frac{1}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(\mathbf{x}_i^j)} \left\| \mathbf{y}_i^j - \tilde{\mathbf{Y}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j \right\|_2^2 = \\ & \text{tr} \left(\frac{1}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(\mathbf{x}_i^j)} (\mathbf{y}_i^j - \tilde{\mathbf{Y}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j) (\mathbf{y}_i^j - \tilde{\mathbf{Y}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j)^T \right) = \\ & \text{tr} \left(\frac{1}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(\mathbf{x}_i^j)} \mathbf{A}^T (\mathbf{x}_i^j - \tilde{\mathbf{X}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j) (\mathbf{x}_i^j - \tilde{\mathbf{X}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j)^T \mathbf{A} \right) = \\ & \text{tr}(\mathbf{A}^T \mathbf{S}_b \mathbf{A}) \end{aligned} \quad (6)$$

式中: $\mathbf{S}_b = \frac{1}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(\mathbf{x}_i^j)} (\mathbf{x}_i^j - \tilde{\mathbf{X}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j) (\mathbf{x}_i^j - \tilde{\mathbf{X}}_{im}^j \tilde{\boldsymbol{\beta}}_{im}^j)^T$ 是类间局部重构散列矩阵; $l(\mathbf{x}_i^j)$ 表示与 \mathbf{x}_i^j 欧氏距离最近的 K 个类, 其中不包括第 i 类, 即 $i \notin l(\mathbf{x}_i^j)$, 而 $m \in l(\mathbf{x}_i^j)$ 则表示第 m 类是 \mathbf{x}_i^j 的 K 近邻类; $\tilde{\mathbf{X}}_{im}^j$ 是通过 k -最近邻算法计算出的 \mathbf{x}_i^j 在第 m 类训练样本 \mathbf{X}_m 中的局部线性子集; $\tilde{\boldsymbol{\beta}}_{im}^j$ 表示 \mathbf{x}_i^j 用第 m 类样本的局部线性子集 $\tilde{\mathbf{X}}_{im}^j$ 表示时的线性回归系数。

为了使 LLRC-DA 表现出更好的性能, 本文使用正交投影, 即 $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ 。因为投影矩阵是正交的, 所以可以消除大量的冗余信息, 得到更简洁的特征子空间。在特征子空间最小化类内局部重构误差同时最大化类间局部重构误差, 可以得到以下目标函数:

$$\mathbf{A}^* = \arg \min_{\mathbf{A}} \frac{\text{tr}(\mathbf{A}^T \mathbf{S}_w \mathbf{A})}{\text{tr}(\mathbf{A}^T \mathbf{S}_b \mathbf{A})} \quad \text{s.t.} \quad \mathbf{A}^T \mathbf{A} = \mathbf{I} \quad (7)$$

2.2 LLRC-DA 优化

LLRC-DA 是一个迹比最小化问题, 通常此类

问题可以使用广义特征值分解的方法近似求解。最近的研究显示^[20], 直接解决迹比问题是可行的并且比使用广义特征值分解效果更好。

假设目标函数存在一个最小值 ρ^* , 当投影矩阵为 A^* 时可以达到这个最小值, 因此对于满足式 (7) 的任意投影矩阵 A , 有以下关系:

$$\frac{\text{tr}(A^T S_w A)}{\text{tr}(A^T S_b A)} \geq \rho^* \quad (8)$$

因此可以得到

$$\text{tr}(A^T S_w A) - \rho^* \text{tr}(A^T S_b A) \geq 0 \quad (9)$$

为了优化目标函数, 本文定义以下函数:

$$f(\rho) = \min_A \text{tr}(A^T S_w A - \rho A^T S_b A) \quad (10)$$

式中 $f(\rho)$ 具有引理 1 中的属性, 已经在文献 [21-22] 中得到证明。

引理 1 1) f 是关于 ρ 的一个非增函数; 2) 当 $\rho = \rho^*$ 时, $f(\rho) = 0$ 。

由上述 $f(\rho)$ 的属性可知, f 是关于 ρ 的一个非增函数。对于函数 $f(\rho)$, 当 ρ 为负数或者较小的正数时, $f(\rho) > 0$; 当 ρ 为较大的正数时, $f(\rho) < 0$, 所以 ρ^* 一定存在, 因此可以通过交替迭代更新得到 ρ^* 和相对应的 A^* 。

首先保持 ρ 不变并更新 A , 然后保持 A 不变并更新 ρ 。LLRDA 中线性回归系数 $\tilde{\beta}$ 在更新 A 之前进行计算, 并且在接下来的计算过程中固定不变。但是实际上 $\tilde{\beta}$ 与 A 是相关的, 不是独立的, $\tilde{\beta}$ 的更新会对 A 产生影响。为了解决这个问题, 利用 $\tilde{\beta}$ 与 A 之间的关系, 使用解析解形式:

$$\tilde{\beta} = (\tilde{X}_i^T A A^T \tilde{X}_i)^{-1} \tilde{X}_i^T A A^T y \quad (11)$$

来消除 $\tilde{\beta}$ 的影响。

1) 固定 A , 更新 ρ

对于一个给定的 ρ , 用 $A(\rho)$ 表示式 (10) 的解。固定 A , 对 ρ 进行求导:

$$f'(\rho) = -\text{tr}(A(\rho)^T S_b A(\rho)) \quad (12)$$

已知 $A(\rho)$, 可以通过 $\rho_{\text{new}} = \rho - \eta_1 \frac{f(\rho)}{f'(\rho)}$ 得到目标函数的根, 其中 η_1 表示步长。

2) 固定 ρ , 更新 A

存在约束条件 $A^T A = I$ 的情况下, 可以通过求解下面问题来求解 A 。

$$\begin{aligned} \min_A J(A) &= \min_A \text{tr}(A^T S_w A - \rho A^T S_b A) = \\ &= \min_A (\text{tr}(A^T S_w A) - \rho \text{tr}(A^T S_b A)) = \\ &= \min_A \left(\frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} J_i^{(w)}(A) - \frac{\rho}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(x_i^j)} J_{im}^{(b)}(A) \right) \quad (13) \\ \text{s.t. } &A^T A = I \end{aligned}$$

其中:

$$J_i^{(w)}(A) = \text{tr}(A^T (x_i^j - \tilde{X}_i^j \tilde{\beta}_i^j) (x_i^j - \tilde{X}_i^j \tilde{\beta}_i^j)^T A)$$

$$J_{im}^{(b)}(A) = \text{tr}(A^T (x_i^j - \tilde{X}_{im}^j \tilde{\beta}_{im}^j) (x_i^j - \tilde{X}_{im}^j \tilde{\beta}_{im}^j)^T A)$$

因为目标函数要在正交约束条件下进行求解, 所以首先要对其进行求导。 $J(A)$ 对 A 求导得

$$\frac{dJ(A)}{dA} = \frac{1}{n} \sum_{i=1}^C \sum_{j=1}^{n_i} \frac{dJ_i^{(w)}(A)}{dA} - \frac{\rho}{nK} \sum_{i=1}^C \sum_{j=1}^{n_i} \sum_{m \in l(x_i^j)} \frac{dJ_{im}^{(b)}(A)}{dA} \quad (14)$$

对于每一个 x_i^j , 可以计算 $J_i^{(w)}(A)$ 的导数, 为了方便表示, 用 x 来替代 x_i^j , 用 \tilde{X} 来替代 \tilde{X}_i^j :

$$\frac{dJ_i^{(w)}(A)}{dA} = (2xx^T + 6GG^T + R + R^T - H - H^T)A \quad (15)$$

式中: $G = \tilde{X}(\tilde{X}^T A A^T \tilde{X})^{-1} \tilde{X}^T A A^T x$, $R = 2(x - G)G^T$, $H = 4Gx^T$ 。

同样地, 可以计算 $J_{im}^{(b)}(A)$ 的导数, 为了方便表示, 用 x 来替代 x_i^j , 用 \tilde{X} 来替代 \tilde{X}_{im}^j :

$$\frac{dJ_{im}^{(b)}(A)}{dA} = (2xx^T + 6GG^T + R + R^T - H - H^T)A \quad (16)$$

式中: $G = \tilde{X}(\tilde{X}^T A A^T \tilde{X})^{-1} \tilde{X}^T A A^T x$, $R = 2(x - G)G^T$, $H = 4Gx^T$ 。

求完导数之后, 定义一个斜对称矩阵 M :

$$M = \frac{dJ(A)}{dA} A^T - A \left(\frac{dJ(A)}{dA} \right)^T \quad (17)$$

然后, 可以通过 Crank-Nicolson-like 方法, 得到更新后的投影矩阵 A_{new} 。

$$A_{\text{new}} = A - \frac{\tau}{2} M (A + A_{\text{new}}) \quad (18)$$

式中: τ 表示步长, 可以通过解析解的形式得到 $A_{\text{new}} = \left(I + \frac{\tau}{2} M \right)^{-1} \left(I - \frac{\tau}{2} M \right) A$, I 是单位矩阵, 通过这种方法求得的 A_{new} 有很好的特性^[23], 即 $A_{\text{new}}^T A_{\text{new}} = A^T A$ 。

通过前面对 LLRC-DA 的描述, 可以得到 LLRC-DA 的算法:

输入 训练样本矩阵 X , 特征子空间维数 d , 最近邻个数 k , 最近邻类 K , 最大迭代次数 t_1, t_2 , 步长 η_1, τ ;

输出 投影矩阵 $A_{\text{LLRC-DA}}$ 。

1) 如果原始样本的维数大于样本总数, 先使用 PCA 方法将原始样本投影到低维子空间, 其中 A_{PCA} 是使用 PCA 得到的投影矩阵;

2) 固定 ρ , 对于每个样本, 计算它在每个类中的 k 个最近邻, 使用 $A_{\text{new}} = \left(I + \frac{\tau}{2} M \right)^{-1} \left(I - \frac{\tau}{2} M \right) A$ 来更新 A , 直到收敛或者达到迭代次数 t_1 ;

3) 固定 A , 对于每个样本, 计算它在每个类中的 k 个最近邻, 使用 $\rho_{\text{new}} = \rho - \eta_1 \frac{f(\rho)}{f'(\rho)}$ 来更新 ρ 。

4) 重复 2)、3), 直到 $f(\rho) = 0$ 或者达到迭代次数 t_2 。

5) 最终的投影矩阵为 $A_{LLRC-DA} = A_{PCA}A_0$ 。

3 实验

为了验证 LLRC-DA 的有效性, 本文分别在 FERET 和 ORL 人脸库上进行实验, 并与 PCA^[13]、LDA^[14]、LPP^[15]、RDA^[24] 和 LLRDA^[19] 等先进的维数约简算法进行比较。在维数约简之后, 分别使用 NNC^[25]、MDC^[26]、LRC^[17] 和 LLRC^[18] 分类器来比较每种维数约简方法在不同分类器下的分类性能。本文中的实验通过 MATLAB 编程实现, 硬件环境为酷睿 i7 处理器、主频 2.6 GHz、内存 8 GB。

3.1 FERET 人脸库上的实验

FERET 人脸库包含 1 400 幅人脸图像, 共计 200 类, 其中每类 7 幅图像, 包括两张面部表情图, 两张左侧图, 两张右侧图和一张光照图, 图 1 所示为一个人的 7 幅图像。实验中所有的图像均被裁剪为 80 像素×80 像素。在实验之前, 首先使用 PCA 方法将原始图像降到 180 维, 然后在每类中选择前 4 幅图像作为训练集, 剩余的作为测试集, 最近邻类 $K = 101$, 每类中的最近邻个数 $k = 3$, 在 FERET 上的识别率如表 1 所列。



图 1 FERET 库中一个人的 7 幅图像

Fig. 1 Seven images of a person from the FERET database

表 1 FERET 库上的识别率

Table 1 Recognition accuracy when using the FERET database

方法	NNC	MDC	LRC	LLRC
PCA	0.671 7	0.521 7	0.695 0	0.693 3
LDA	0.758 3	0.733 3	0.735 0	0.741 7
LPP	0.530 0	0.456 7	0.615 0	0.601 7
RDA	0.748 3	0.733 3	0.750 0	0.763 3
LLRDA	0.755 0	0.745 0	0.756 7	0.771 7
LLRC-DA	0.753 3	0.693 3	0.788 3	0.795 0

通过表 1 可以看出: 1) 与 PCA、LPP 等非监督学习方法相比, 监督学习的降维方法明显具有更好的性能, 因为监督学习方法利用了类别信息, 使识别率显著提高; 2) 同一种降维方法, 在不同的分类器上的实验结果不同, 甚至会有很大差异, 例如 LLRC-DA 在 MDC 与 LLRC 这两种分类器上的识别率相差很大, 这也说明选择合适的分

类器对于识别性能至关重要; 3) 在所有的降维方法与分类器的组合中, LLRC-DA 与 LLRC 的组合具有最高的识别率, 由于使用了 LLRC 的决策准则, 因此和其他降维方法相比, LLRC-DA 学习的特征子空间与 LLRC 更加契合; 4) LLRC-DA 在 LRC、LLRC 上的识别率均明显高于 LLRDA, 证明本文提出的方法对于提高分类率是有效的。

3.2 ORL 人脸库上的实验

ORL 人脸库的 400 幅图像来自 40 个人, 其中每人 10 幅图像, 图像在不同条件下进行采集, 如光照、面部表情、面部细节等, 图 2 所示为一个人的 10 幅图像。实验中所有的图像均压缩成 56 像素×46 像素。在实验之前, 首先使用 PCA 方法将原始图像降到 50 维, 然后在每类中选择前 5 幅图像作为训练集, 剩余的所有图像作为测试集。



图 2 ORL 库中一个人的 10 幅图像

Fig. 2 Ten images of a person from the ORL database

在 ORL 人脸库上, 本文对 LLRC-DA 在不同参数下的性能进行评估。首先, 对每类中的最近邻参数 k 进行实验, 使用 NNC、MDC、LRC、LLRC 分类器, 固定 $K = 21$, k 的值为 1~4, 实验结果如图 3 所示。从图 3 中可以看出, 在 NNC、LRC、LLRC 分类器下, k 的取值为 3 时, 识别率最高, 这表明利用样本之间的邻域结构, 选择 k 而不是全部样本来表示测试样本, 可以提高识别性能。

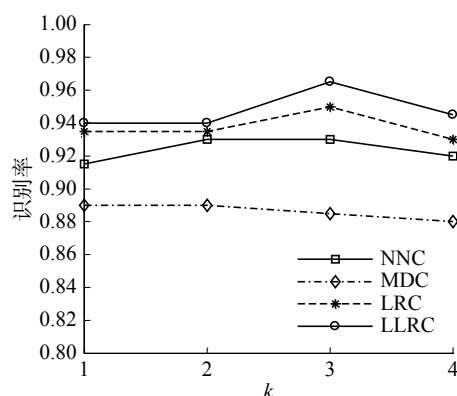


图 3 LLRC-DA 在不同 k 下的识别率

Fig. 3 Recognition rate of LLRC-DA with varied k

接着,对近邻类参数 K 进行实验,使用 NNC、MDC、LRC、LLRC 分类器,选择 $k=3$, K 的值为 3~39,实验结果如图 4 所示。从图 4 中可以看出, K 为 20 时,各分类器均可取得较高的识别率。

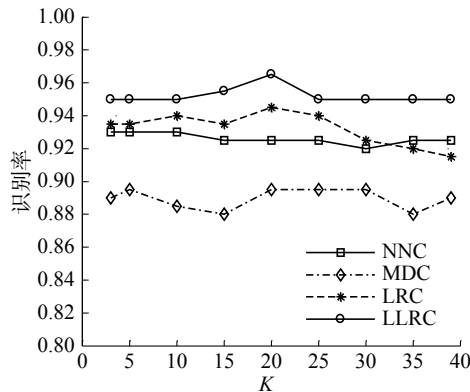


图 4 LLRC-DA 在不同 K 下的识别率

Fig. 4 Recognition rate of LLRC-DA with varied K

最后,比较在不同分类器下 LLRC-DA 的性能,设置 $K=21$, $k=3$,在 ORL 上的识别率如表 2。通过表 2 可以看出:1) 在所有的降维方法与分类器的组合中,LLRC-DA 与 LLRC 的组合具有最高的识别率,证明本文提出的方法是有效的;2) RDA 与 LRC 这一组合也有着较高的识别率,这是因为 RDA 是采用 LRC 决策准则产生的,证明根据分类器的决策准则来设计维数约简算法是很有必要的。然而它的识别率依然比 LLRC-DA 与 LLRC 组合的低,原因是 LLRC 决策准则考虑了数据的邻域结构,使用 k 个最近邻而不是全部样本,这对于提高识别率是有效的。

表 2 ORL 库上的识别率

Table 2 Recognition accuracy when using the ORL database

方法	NNC	MDC	LRC	LLRC
PCA	0.930 0	0.860 0	0.950 0	0.950 0
LDA	0.925 0	0.940 0	0.925 0	0.920 0
LPP	0.885 0	0.850 0	0.920 0	0.920 0
RDA	0.940 0	0.890 0	0.950 0	0.940 0
LLRDA	0.935 0	0.885 0	0.950 0	0.955 0
LLRC-DA	0.930 0	0.885 0	0.950 0	0.965 0

4 结束语

本文提出了面向局部线性回归分类器的判别分析方法 LLRC-DA,根据 LLRC 的决策准则设计目标函数,通过最大化类间局部重构误差并最小化类内局部重构误差来寻找最优的特征子空间。

本文利用了投影矩阵与线性回归系数之间的关系,消除了线性回归系数的影响,使得目标函数只和投影矩阵相关。LLRC-DA 通过对投影矩阵添加正交约束来消除冗余信息,并得到更好的特征子空间。相比于传统算法使用最大特征值分解来求解迹比问题,本文利用了一种新的迹比优化算法来有效地求解投影矩阵。本文在 FERET、ORL 人脸库上进行了实验,与先进的维数约简算法进行了比较,实验结果表明,本文提出的方法具有更好的性能。对于 LLRC-DA 中的最近邻参数 k 的选择问题,将寻找合适的参数选择技术来确定 k 值。在接下来的工作中,打算引入核方法来研究改进本文的方法。

参考文献:

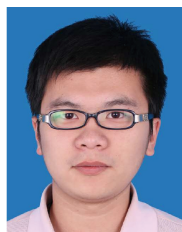
- [1] HARANDI M, SALZMANN M, HARTLEY R. Dimensionality reduction on SPD manifolds: the emergence of geometry-aware methods[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2018, 40(1): 48–62.
- [2] CHOI S, SHIN J H, LEE J, et al. Experimental demonstration of feature extraction and dimensionality reduction using memristor networks[J]. *Nano letters*, 2017, 17(5): 3113–3118.
- [3] YIN Jun, WEI Lai, SONG Miao, et al. Optimized projection for collaborative representation based classification and its applications to face recognition[J]. *Pattern recognition letters*, 2016, 73: 83–90.
- [4] XIA Zhihua, WANG Xinhui, ZHANG Liangao, et al. A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing[J]. *IEEE transactions on information forensics and security*, 2016, 11(11): 2594–2608.
- [5] GONG Yunchao, LAZEBNIK S, GORDO A, et al. Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(12): 2916–2929.
- [6] YUAN Chengsheng, SUN Xingming, LV Rui. Fingerprint liveness detection based on multi-scale LPQ and PCA[J]. *China communications*, 2016, 13(7): 60–65.
- [7] DUBEY R K, GOH J, THING V L L. Fingerprint liveness detection from single image using low-level features and shape analysis[J]. *IEEE transactions on information forensics and security*, 2016, 11(7): 1461–1475.
- [8] YE Jieping, LI Tao, XIONG Tao, et al. Using uncorrelated discriminant analysis for tissue classification with gene expression data[J]. *IEEE/ACM transactions on computational biology and bioinformatics*, 2004, 1(4): 181–190.

- [9] ZOU Quan, ZENG Jiancang, CAO Liujuan, et al. A novel features ranking metric with application to scalable visual and bioinformatics data classification[J]. *Neurocomputing*, 2016, 173: 346–354.
- [10] LI Tao, ZHU Shenghuo, OGIHARA M. Using discriminant analysis for multi-class classification: an experimental investigation[J]. *Knowledge and information systems*, 2006, 10(4): 453–472.
- [11] AMATRIAIN X, PUJOL J M. Data mining methods for recommender systems[M]//RICCI F, ROKACH L, SHAPIRA B. *Recommender Systems Handbook*. Boston, MA: Springer, 2015: 227–262.
- [12] RAUDYS S J, JAIN A K. Small sample size effects in statistical pattern recognition: recommendations for practitioners[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1991, 13(3): 252–264.
- [13] TURK M, PENTLAND A. Eigenfaces for recognition[J]. *Journal of cognitive neuroscience*, 1991, 3(1): 71–86.
- [14] BELHUMEUR P N, HESPAHNA J P, KRIEGMAN D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1997, 19(7): 711–720.
- [15] HE Xiaofei, YAN Shuicheng, HU Yuxiao, et al. Face recognition using laplacianfaces[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2005, 27(3): 328–340.
- [16] ZHANG Lei, YANG Meng, FENG Xiangchu. Sparse representation or collaborative representation: which helps face recognition?[C]//*Proceedings of 2011 International Conference on Computer Vision*. Barcelona, Spain, 2011: 471–478.
- [17] NASEEM I, TOGNERI R, BENNAMOUN M. Linear regression for face recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2010, 32(11): 2106–2112.
- [18] BROWN D, LI Hanxi, GAO Yongsheng. Locality-regularized linear regression for face recognition[C]//*Proceedings of the 21st International Conference on Pattern Recognition*. Tsukuba, Japan, 2012: 1586–1589.
- [19] HUANG Pu, LI Tao, SHU Zhenqiu, et al. Locality-regularized linear regression discriminant analysis for feature extraction[J]. *Information sciences*, 2018, 429: 164–176.
- [20] WANG Huan, YAN Shuisheng, XU Dong, et al. Trace ratio vs. Ratio trace for dimensionality reduction[C]//*Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition*. Minneapolis, USA, 2007: 1–8.
- [21] NGO T T, BELLALIJ M, SAAD Y. The trace ratio optimization problem[J]. *SIAM review*, 2012, 54(3): 545–569.
- [22] JIANG Wenhao, CHUNG Fulai. A trace ratio maximization approach to multiple kernel-based dimensionality reduction[J]. *Neural networks*, 2014, 49: 96–106.
- [23] WEN Zaiwen, YIN Wotao. A feasible method for optimization with orthogonality constraints[J]. *Mathematical programming*, 2013, 142(1/2): 397–434.
- [24] CHEN Yi, JIN Zhong. Reconstructive discriminant analysis: a feature extraction method induced from linear regression classification[J]. *Neurocomputing*, 2012, 87: 41–50.
- [25] COVER T, HART P. Nearest neighbor pattern classification[J]. *IEEE transactions on information theory*, 1967, 13(1): 21–27.
- [26] ZHANG Daoqiang, CHEN Songcan, ZHOU Zhihua. Learning the kernel parameters in kernel minimum distance classifier[J]. *Pattern recognition*, 2006, 39(1): 133–135.

作者简介:



朱换荣,女,1994年生,硕士研究生,主要研究方向为机器学习、人脸识别。



郑智超,男,1992年生,博士研究生,主要研究方向为人脸识别、子空间学习。



孙怀江,男,1968年生,教授,博士生导师,主要研究方向为神经网络与机器学习、人体运动分析与合成、多媒体与虚拟现实、图像处理与计算机视觉。曾主持或参与完成国家级项目3项,省部级项目3项,获省部级科技进步二等奖1项。发表学术论文80余篇。