

流形判别分析

刘忠宝^{*①} 潘广贞^① 赵文娟^②

^①(中北大学电子与计算机科学技术学院 太原 030051)

^②(山西大学商务学院信息学院 太原 030031)

摘要: 传统降维方法主要有两种思路: 一是利用样本的全局特征, 保证降维前后样本的全局特征不变; 二是尽量保证相邻样本在降维前后的相对关系不变。传统方法由于未能充分利用样本的已有信息, 因此降维效率有限。鉴于此, 在 Fisher 准则和局部流形保持的基础上, 该文提出流形判别分析。该方法首先定义了基于流形的类内离散度 MWCS 和类间离散度 MBCS, 然后利用 Fisher 准则找到最佳投影方向, 该方向满足 MBCS 与 MWCS 之比最大。该方法不仅继承了传统降维方法的优势, 而且进一步提高了降维效率。标准数据集上的实验结果表明该文所提方法的有效性。

关键词: 模式识别; 流形学习; 降维; Fisher 准则; 全局特征

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2013)09-2047-07

DOI: 10.3724/SP.J.1146.2012.01552

Manifold-based Discriminant Analysis

Liu Zhong-bao^① Pan Guang-zhen^① Zhao Wen-juan^②

^①(School of Electronics and Computer Science Technology, North University of China, Taiyuan 030051, China)

^②(School of Information, Business College of Shanxi University, Taiyuan 030031, China)

Abstract: Researches on current Dimensionality Reduction (DR) methods are mainly based on two ways. One attempts to ensure the stabilities of global features of high-dimensional samples, the other tries to make the local manifold structure between data before and after dimension reduction be as invariant as possible. As the existed information is not fully utilized by current DR methods, the DR efficiencies are restricted. Based on the above analysis, Manifold-based Discriminant Analysis (MDA) is proposed based on Fisher criterion and manifold preserving. The global features and local structure are both taken into consideration by MDA. It defines two scatters: Manifold-based Within-Class Scatter (MWCS) and Manifold-based Between-Class Scatter (MBCS). According to Fisher criterion, the optimal projection satisfies the ratio of MBCS and MWCS is maximized. MDA not only inherits the superiorities of current DR methods, but further improves the DR efficiencies. Experiments on some standard datasets verify the effectiveness of the proposed method MDA.

Key words: Pattern recognition; Manifold learning; Dimensionality reduction; Fisher criterion; Global features

1 引言

线性判别分析(Linear Discriminant Analysis, LDA)^[1]是当前主流的特征降维方法之一, 目前已广泛应用于图像处理、语音识别、指纹识别等诸多领域。LDA 的基本思想是保证降维后的样本类内尽可能紧密, 而类间尽可能远离。LDA 的最佳投影方向由 $S_W^{-1} S_B$ 的特征向量决定, 其中 S_W 和 S_B 分别称为类内离散度和类间离散度。实际应用中, LDA 往往

面临类内离散度 S_W 奇异的问题, 我们称此问题为小样本问题(Small Sample Size, SSS)^[2]。为了解决上述问题, 科学家们提出众多改进算法: Friedman^[3]提出的正则化判别分析(Regularized Discriminant Analysis, RDA)有效地解决了协方差矩阵奇异问题; 文献[4]提出的 2 维线性判别分析(Two-Dimensional LDA, 2DLDA)利用样本矩阵直接提取特征; 文献[5]提出的正交化线性判别分析(Orthogonal LDA, OLDA)通过求解一组正交基解决了广义特征值问题; 文献[6]提出的直接线性判别分析(Direct LDA, D-LDA)通过白化类间离散度矩阵以及对角化类内离散度矩阵有效地解决小样本问题; 文献[7]提出最大间隔准则(Maximum Margin

2012-11-30 收到, 2013-02-06 改回

国家自然科学基金(61202311)和山西省自然科学基金(2012011011-3)资助课题

*通信作者: 刘忠宝 liu_zhongbao@hotmail.com

Criterion, MMC)通过最大化特征空间中的类间离散度与类内离散度的差,从而避免出现小样本问题。此外,常见的LDA改进算法还有伪逆判别分析(Pseudo-inverse LDA, PLDA)^[8],零空间判别分析(Null-space LDA, NLDA)^[9],两级判别分析(Two-stage LDA)^[10],惩罚性判别分析(Penalized Discriminant Analysis, PDA)^[11],增强型Fisher线性判别分析(Enhanced Fisher linear discriminant Model, EFM)^[12]等。近年来笔者亦进行了相关研究,先后提出了标量化的线性判别分析算法(Scalarized LDA, SLDA)^[13]、基于多阶矩阵组合改进的LDA算法(Modified LDA based on linear combination of k-order matrices, MLDA)^[14]以及基于矩阵指数的线性判别分析算法(Matrix Exponential LDA, MELDA)^[15]。

当面临小样本问题时,上述方法常用的策略是先解决类内离散度奇异性问题,然后再利用LDA算法求解。然而,LDA算法主要考虑的是样本的全局特性,对于样本的流形结构考虑甚少。另一方面,当前主流流形学习方法,如局部保持投影(Locality Preserving Projection, LPP)^[16],局部线性嵌入(Local Linear Embedding, LLE)^[17],等距映射(ISOmetric feature MAPping, ISOMAP)^[18],无监督判别映射(Unsupervised Discriminant Projection, UDP)^[19]等往往关注的是样本的局部结构,忽略了样本的全局特征。

鉴于此,为了能充分利用样本的全局信息和局部信息,进一步提高降维效率,本文提出流形判别分析(Manifold Discriminant Analysis, MDA)。该方法将Fisher准则和局部流形保持有机地结合起来,充分利用样本的已有信息,有效地提高了降维效率。

本文第2节介绍背景知识;第3节提出流形判别分析MDA;第4节通过标准数据集上的实验验证MDA的有效性;第5节总结全文。

2 背景知识

2.1 LDA

LDA是一种有效的特征降维方法,其目标是寻找一最优投影方向,使样本在该方向上具有最优的可分度。设有 N 个 d 维样本 x_1, x_2, \dots, x_N ,类别数为 c ,各类样本数为 $N_i (i=1, 2, \dots, c)$ 。类间离散度和类内离散度分别定义如下:

$$S_B = \sum_{i=1}^c \frac{N_i}{N} [\bar{x}_i - \bar{x}][\bar{x}_i - \bar{x}]^T \quad (1)$$

$$S_W = \sum_{i=1}^c \sum_{j=1}^{N_i} \frac{1}{N} [x_{ij} - \bar{x}_i][x_{ij} - \bar{x}_i]^T \quad (2)$$

其中 \bar{x} 为所有样本均值, \bar{x}_i 为第 i 类样本均值。

Fisher准则函数定义如下:

$$J(W_{\text{opt}}) = \max_W \frac{W^T S_B W}{W^T S_W W} \quad (3)$$

式(3)等价于

$$\max_W W^T S_B W \quad (4)$$

且

$$\min_W W^T S_W W \quad (5)$$

其中 W 为投影矩阵。

通过Fisher准则求得的最优投影方向 W_{opt} 保证样本在该方向上具有最大的类间离散度和最小的类内离散度。由Lagrange乘子法可知:当 S_W 非奇异时, W_{opt} 满足等式 $S_W^{-1} S_B W = \lambda W$ 的解。

由以上分析可知:LDA通过Fisher准则最大化类间离散度 S_B 和类内离散度 S_W 之比,尽量保证样本降维前后的全局特征不变。 S_W 表明各类样本与其类中心之间的远近关系,并未考虑到样本的流形结构。上述问题严重制约着LDA降维效率的提高。

2.2 流形与流形学习

微分几何中首先对空间 M 上每一点的无穷小邻域定义与欧式空间某个开集的微分同胚,加上这些邻域的连接信息组成微分结构,空间 M 连同这个微分结构称为流形。流形学习是一种新兴的机器学习方法,其主要目标是从高维数据中恢复低维流形结构,并求出相应的嵌入映射,以实现降维。流形学习是从观测到的现象中寻找事物的本质,寻求数据产生的内在规律的有效途径,目前已成为数据挖掘、模式识别、机器学习等领域的研究热点。

2.3 LPP

LPP是一种典型的流形学习方法,其目标是保持样本局部流形结构不变,即高维空间的邻近样本在低维空间的相对关系保持不变。上述思想可表示为式(6)、式(7)组成的最优化问题:

$$\left. \begin{aligned} \min_W \sum_{i,j} (W^T x_i - W^T x_j)^2 S_{ij} \\ \text{s.t.} \sum_i W^T x_i D_{ii} x_i^T W = 1 \end{aligned} \right\} \quad (6)$$

其中 W 为投影矩阵; S_{ij} 为权重函数,其表征样本的相似程度; $D_{ii} = \sum_j S_{ij}$ 。

利用线性代数理论可将上述最优化问题转化为

$$\left. \begin{aligned} \min_W W^T X L X^T W \\ \text{s.t.} W^T X D X^T W = 1 \end{aligned} \right\} \quad (7)$$

其中 $L = D - S$, S 为权重矩阵, D 为对角阵且 $D_{ii} = \sum_j S_{ij}$, X 为样本集合。

LPP 的最佳投影方向可通过求解 $XLX^T W = \lambda XDX^T W$ 的特征向量获得。

综上, LPP 通过保持降维前后样本的相对关系不变实现降维。然而, LPP 并未考虑到样本的全局特性, 当样本中含有噪声时, LPP 的分类性能受到很大影响。

3 流形判别分析

3.1 基本思想

特征降维是当前解决大规模高维数据的一种经典处理技术, 目前已广泛应用于人脸识别、文本分类、基因图谱分析等领域。然而传统降维方法难以充分利用样本的全局特性和局部流形结构, 降维效率有待进一步提高。针对传统方法的不足, 本文提出流形判别分析(Manifold Discriminant Analysis, MDA)。该方法引入两个重要概念: 基于流形的类内离散度(Manifold-based Within-Class Scatter, MWCS)和基于流形的类间离散度(Manifold-based Between-Class Scatter, MBCS)。在 Fisher 准则的基础上, 通过最大化 MBCS 与 MWCS 之比实现降维。

3.2 基于流形的类间离散度

受流形学习启发, 首先创建邻接图 $G_D = \{X, D\}$, 其中 X 为样本集合, D 表示异类样本间的权重。 X 中的任意两个样本 x_i 和 x_j , 其异类权重函数定义如下:

$$D_{ij} = \begin{cases} \exp(-d/\|x_i - x_j\|^2), & l_i \neq l_j \\ 0, & l_i = l_j \end{cases} \quad (8)$$

其中 $l_i, l_j (i, j=1, 2, \dots, N)$ 表示样本的类别标签, d 为常数。

异类权重函数 D_{ij} 表明: 当样本 x_i 和 x_j 异类时, 两者间距较大, 则两者间的权重较大; 当样本 x_i 和 x_j 同类时, 则两者间的权重值为 0。

为了保持异类样本的局部流形结构, 在高维空间彼此远离的异类样本 x_i 和 x_j 降维后仍应保持原有特性。基于上述分析, 可得式(9)最优化表达式:

$$\max_W \sum_{i,j} (y_i - y_j)^2 D_{ij} \quad (9)$$

其中 $y_i = W^T x_i$, W 为投影矩阵, $x_i \in X$ 。

对 $\sum_{i,j} (y_i - y_j)^2 D_{ij}$ 进行代数变化可得

$$\begin{aligned} & \frac{1}{2} \sum_{i,j} (y_i - y_j)^2 D_{ij} \\ &= \frac{1}{2} \sum_{i,j} (W^T x_i - W^T x_j)^2 D_{ij} \\ &= \sum_{i,j} (W^T x_i D_{ij} x_i^T W - W^T x_i D_{ij} x_j^T W) \\ &= W^T X D' X^T W - W^T X D X^T W \\ &= W^T X (D' - D) X^T W = W^T S_D W \end{aligned} \quad (10)$$

其中 $S_D = X(D' - D)X^T$, D' 为对角阵且 $D' = \sum_j D_{ij}$ 。

将(10)式代入式(9)中, 可得

$$\max_W W^T S_D W \quad (11)$$

由前面分析可知: 式(4)反映了各类样本之间的全局特性, 式(11)表明样本的局部流形结构。为了充分利用样本的全局特性和局部流形结构, 综合式(4)和式(11)可得

$$\begin{aligned} & \max_W \alpha W^T S_B W + (1 - \alpha) W^T S_D W \\ &= \max_W W^T [\alpha S_B + (1 - \alpha) S_D] W \\ &= \max_W W^T M_B W \end{aligned} \quad (12)$$

其中 α 为常数; $M_B = \alpha S_B + (1 - \alpha) S_D$, 本文称之为基于流形类间离散度 MBCS。

3.3 基于流形的类内离散度

与基于流形的类间离散度类似, 首先定义同类权重函数定义如下:

$$S_{ij} = \begin{cases} \exp(-\|x_i - x_j\|^2/s), & l_i = l_j \\ 0, & l_i \neq l_j \end{cases} \quad (13)$$

其中 $l_i, l_j (i, j=1, 2, \dots, N)$ 表示样本的类别标签, s 为常数。

同类权重函数 S_{ij} 表明: 当样本 x_i 和 x_j 同类时, 赋予较大的权重; 否则, 权重为 0。为了保持降维前后相邻样本间的相对关系不变, 则找到的最佳投影方向应保证满足式(14)的优化问题:

$$\min_W \sum_{i,j} (y_i - y_j)^2 S_{ij} \quad (14)$$

其中 $y_i = W^T x_i$, W 为投影矩阵, $x_i \in X$ 。

对 $\sum_{i,j} (y_i - y_j)^2 S_{ij}$ 进行代数变化可得

$$\begin{aligned} & \frac{1}{2} \sum_{i,j} (y_i - y_j)^2 S_{ij} \\ &= \frac{1}{2} \sum_{i,j} (W^T x_i - W^T x_j)^2 S_{ij} \\ &= \sum_{i,j} (W^T x_i S_{ij} x_i^T W - W^T x_i S_{ij} x_j^T W) \\ &= W^T X S' X^T W - W^T X S X^T W \\ &= W^T X (S' - S) X^T W = W^T S_S W \end{aligned} \quad (15)$$

其中 $S_S = X(S' - S)X^T$, S' 为对角阵且 $S' = \sum_j S_{ij}$ 。

将式(15)代入式(14)中, 可得

$$\max_W W^T S_S W \quad (16)$$

为了有效利用样本的全局特性和局部流形结构, 综合式(5)和式(16)可得

$$\begin{aligned} & \max_W \beta W^T S_W W + (1 - \beta) W^T S_S W \\ &= \max_W W^T [\beta S_W + (1 - \beta) S_S] W \\ &= \max_W W^T M_W W \end{aligned} \quad (17)$$

其中 β 为常数; $M_W = \beta S_W + (1 - \beta) S_S$, 本文称之为基于流形类内离散度 MWCS。

3.4 流形判别分析

借鉴 LDA, 在 Fisher 准则的基础上, 通过最大化 MBCS 与 MWCS 之比实现降维。上述思想可转化为如式(18)的优化问题:

$$J = \max_W \frac{M_B}{M_W} = \max_W \frac{W^T(\alpha S_B + (1 - \alpha) S_D)W}{W^T(\beta S_W + (1 - \beta) S_B)W} \quad (18)$$

由 Lagrange 乘子法可知: 式(18)中投影矩阵 W 是满足等式 $M_B W = \lambda M_W W$ 的解。

由式(18)可以看出: MDA 不仅充分考虑了样本的全局特性, 而且保持了样本的局部流形结构。MDA 继承了 LDA 和 LPP 的优势, 并在一定程度上提高了降维效率。当 $\alpha = \beta = 1$ 或 $d = s = \infty$ 时, MDA 等价于 LDA; 当 $\alpha = \beta = 0$, $d = \infty$ 且 $s < \infty$ 时, MDA 等价于 LPP。

在实际应用中, M_W 往往奇异, 无法通过上述优化问题求解。为了方便, 采用扰动法解决 M_W 奇异性问题。

基于上述分析, MDA 算法可简述如下:

输入: 样本集 X 和降维数 d

输出: 样本集 X 对应的低维嵌入集 $Y = [y_1, y_2, \dots, y_d]$

步骤 1 创建邻接图 $G_D = \{X, D\}$ 和 $G_S = \{X, S\}$, 其中 $X = \{x_1, x_2, \dots, x_N\}$ 表示样本集, D 和 S 分别表示异类和同类样本间的权重。当两个样本点 x_i 和 x_j 异类时, 则在两者之间新增一条边, 形成异类邻接图; 同理形成同类邻接图。

步骤 2 计算异类权重 D 和同类权重 S 。若异类样本点 x_i 和 x_j 之间有边相连, 则利用式(10)计算异类权重 D ; 若同类样本点 x_i 和 x_j 之间有边相连, 则利用式(15)计算同类权重 S 。

步骤 3 分别计算类间离散度 S_B 、类内离散度 S_W 、基于流形的类间离散度 M_B 以及基于流形的类内离散度 M_W 。

步骤 4 解决 M_W 奇异性问题。当 M_W 奇异时, 采用扰动法解决该问题, 即在其主对角线上增加一个很小的正数 δ 。设增加扰动后的 M_W 为 M'_W 。

步骤 5 计算最佳投影矩阵 W 。最佳投影矩阵 W 满足等式 $M_W^{-1} M_B W = \lambda W$ 或 $M'_W^{-1} M_B W = \lambda W$ 的解。上式前 d 个最大非零特征值对应的特征向量构成投影矩阵 $W = [w_1, \dots, w_d]$ 。

步骤 6 对样本进行降维。对于任意样本 $x_i \in X$, 经降维后可得 $y_i = W^T x_i$ 。

3.5 MDA 与传统降维方法的关系

传统降维方法主要有两种思路: 一是利用样本

的全局特征, 保证降维前后样本的全局特征不变, 典型代表为 LDA; 二是尽量保证相邻样本在降维前后的流形结构不变, 典型代表为 LPP。LDA 在 Fisher 准则下选择最优的投影向量, 使得样本的类间离散度最大而类内离散度最小。LDA 充分利用样本的类别信息, 有效地提高了算法的识别率。由于 LDA 重点考虑的是样本的线性可分性问题, 往往忽略样本的局部流形结构, 因此降维效率有限。以 LPP 为代表的流形学习方法试图保持流形的局部邻域结构信息并利用这些信息构造全局嵌入。流形学习方法能够有效地探索非线性流形分布数据的内在规律与性质。但是在实际应用中该方法对噪声和离群值较为敏感, 这极大限制了其鲁棒性及泛化能力的提高。

流形判别分析 MDA 在 Fisher 准则的基础上, 借鉴流形学习思想, 通过最大化基于流形的类内离散度 MWCS 与基于流形的类间离散度 MBCS 之比实现降维。与传统降维方法相比, MDA 最大优势在于充分利用了样本的全局和局部信息, 不仅保证样本在全局上线性可分, 而且使得样本的局部流形结构尽量保持不变。

4 实验分析

通过与主流降维方法 PCA, LPP, LDA 比较, 验证 MDA 的有效性。实验环境为 3 GHz Pentium4 CPU, 2 G RAM, Windows XP 及 Matlab 7.0。MDA 的降维效率与参数选择有关。参数通过 5 倍交叉验证获取。参数 α 和 β 分别在网格 $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ 中搜索选取。实验包括 UCI 2 维可视化实验和人脸识别实验。

实验步骤如下:

步骤 1 将样本分为训练样本和测试样本;

步骤 2 利用 MDA 求最佳投影方向;

步骤 3 将测试样本投影到最佳投影方向上;

步骤 4 将投影后的测试样本通过最近邻分类器与训练样本进行特征识别, 得到识别结果。

4.1 UCI 2 维可视化实验

实验选取 UCI 中的 Wine 数据集。该数据集包含 3 类 178 个样本, 样本维数为 13。分别在数据上运行 PCA, LPP, LDA, MDA 等降维方法, 并将样本降至 2 维。实验结果如图 1 所示。

由图 1 可以看出: PCA 降维后的 3 类样本的重叠率较高, 降维效率较低。LPP, LDA, MDA 基本能完成降维任务, 但三者效率差别较大。三者中, LPP 的识别率最低, 主要表现在样本分布松散且 3 类样本边界附近重叠率较高, 主要原因在于 LPP 关注样本的局部流形结构, 对于样本的全局特性, 特别是样本的类内离散度和类间离散度重视不够。

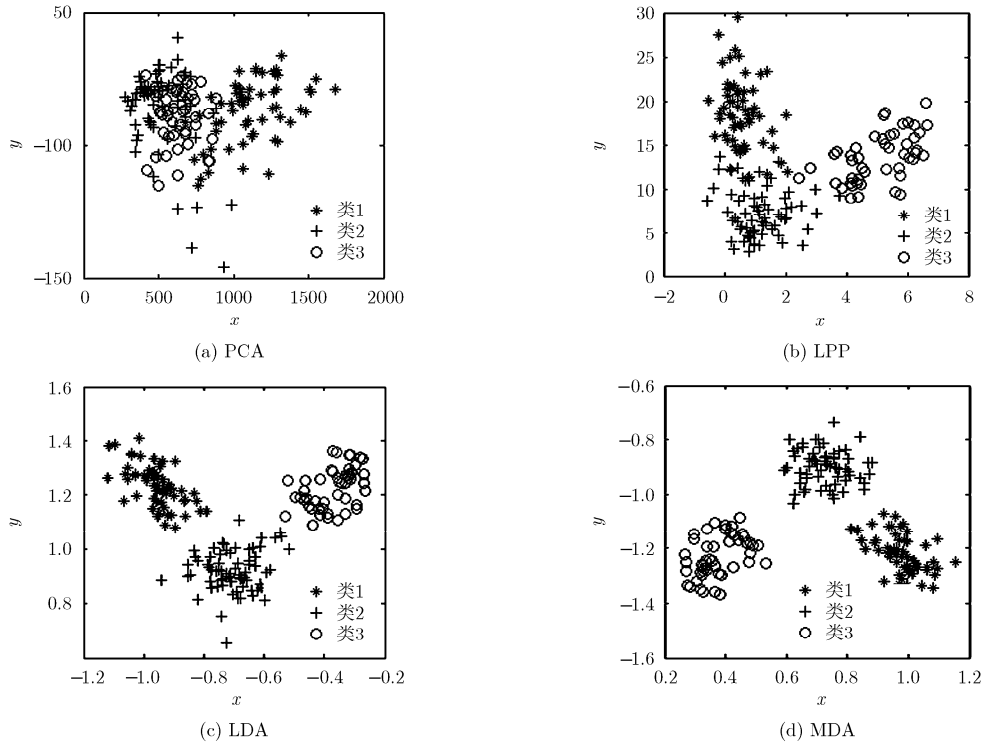


图 1 2 维可视化实验结果

LDA 和 MDA 均能达到较高的识别率, 但从样本的分布性状看, MDA 显然比 LDA 更优。主要原因在于 MDA 同时考虑了样本的全局特性和局部流形结构, 保证降维过程中尽可能保持样本的原有特性。而 LDA 关注的是样本的全局特性, 即 LDA 中的类间离散度保证各类样本尽量分开, 类内离散度反映各类样本内部的紧密程度, 但两者均未考虑相邻样本在降维前后相对关系的稳定性, 即样本的局部结构。

4.2 人脸数据集上的实验

实验数据集采用 ORL 数据集和 Yale 数据集。实验分别考察识别率与训练样本数和降维数的关系, 从而说明 MDA 较之传统方法的优越性。

4.2.1 识别率与训练样本数的关系 实验选取 ORL 中每人前 $k(k=3,4,5,6,7)$ 幅照片作为训练样本, 剩下的作为测试样本; 选取 Yale 中每人前 $k(k=3,5,7,9)$ 幅照片作为训练样本, 剩下的作为测试样本。分别在上述数据集上运行 PCA, LPP, LDA, MDA, 得到的实验结果如表 1 所示, 其中括号外的值表示最佳识别率, 括号内的值表示取得最佳识别率时的维数。实验中 LDA 实际上为 PCA+LDA。

由表 1 可以看出: 在 ORL 数据集上, 与 PCA, LPP, LDA 相比, MDA 具有最优的识别率; 在 Yale 数据集上, 除 $k=8$ 外, MDA 具有最优的识别率。

表 1 PCA, LPP, LDA 和 MDA 在人脸数据集上的识别率

Data sets	k	PCA	LPP	LDA	MDA
ORL	3	0.711(28)	0.789(28)	0.814(30)	0.875(20)
	4	0.808(28)	0.867(30)	0.875(30)	0.954(18)
	5	0.845(22)	0.890(24)	0.905(30)	0.950(21)
	6	0.863(22)	0.906(30)	0.950(30)	0.963(25)
	7	0.892(20)	0.917(22)	0.925(26)	0.958(20)
	8	0.873(20)	0.925(30)	0.938(26)	0.963(23)
Yale	4	0.619(12)	0.733(14)	0.667(14)	0.733(12)
	5	0.667(14)	0.763(14)	0.767(14)	0.767(13)
	6	0.653(12)	0.770(14)	0.747(10)	0.787(14)
	7	0.750(12)	0.833(12)	0.833(14)	0.900(14)
	8	0.800(10)	0.899(14)	0.822(14)	0.867(14)

4.2.2 识别率与降维数的关系 实验选取 ORL 中每人前 5 张照片作为训练样本, 剩下的作为测试样本; Yale 中每人前 9 幅照片作为训练样本, 剩下的作为测试样本。实验结果如图 2 所示。

由图 2 可以看出: 随着降维数的增大, 识别率基本呈上升趋势(个别情况除外)。与 PCA, LPP, LDA 相比, MDA 的识别率最优。

表 1 及图 2 均表明 MDA 的识别率优于 PCA,

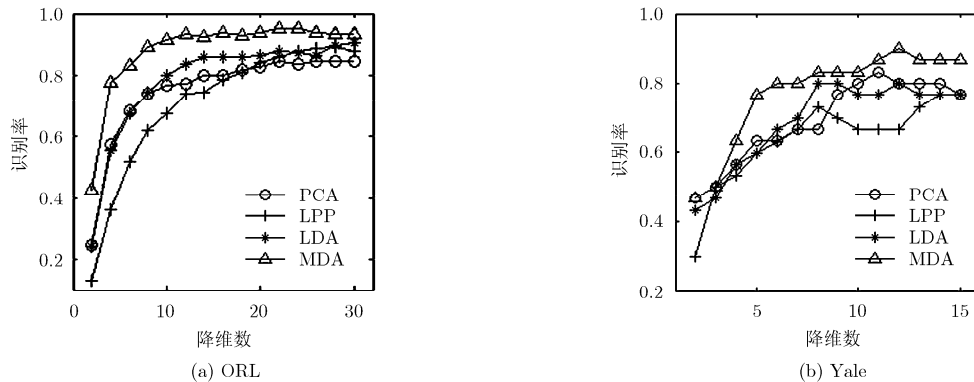


图 2 识别率与降维数的关系

LPP, LDA 等方法。其原因归纳如下:

(1)PCA 试图在样本空间中找到最能表征样本特征的主成分, 而其往往忽略同类样本间的相似度以及异类样本间的差异性。

(2)LDA 关注的是样本的全局特性, 分别用类内离散度和类间离散度表示同类样本间的相似度和异类样本间的差异性。该方法对样本的局部流形结构重视不够, 无法保持降维前后样本的局部特征。

(3)LPP 的基本思想是尽量保证相邻样本在降维前后相对关系不变。该方法很好地保持样本的局部结构, 但其对全局信息考虑不足, 因此在降维过程中易受噪声点或奇异点的影响。

(4)MDA 充分利用样本的全局信息和局部信息, 有机地将 Fisher 准则和局部流形保持结合起来, 有效地提高了降维效率。

5 结束语

当前主流降维方法基本上围绕两种思路提出: 一是保证降维前后样本的全局特征不变, 典型代表有 LDA; 二是保持降维前后样本的局部流形结构不变, 典型代表有 LPP。在分析已有方法不足的基础上, 本文提出流形判别分析 MDA。该方法充分利用样本的全局特性和局部流形结构, 将 Fisher 准则和局部流形保持有机地结合起来, 有效地提高了降维效率。标准数据集上的实验结果表明本文所提方法 MDA 的有效性。实际应用往往面临线性不可分的情况, 如何解决非线性问题成为特征降维领域的研究热点之一。本文所提方法 MDA 是针对线性空间提出的, 当线性不可分时, 如何进行特征降维将是本文下一步的工作。

参考文献

- [1] Martinez A M and Kak A C. PCA versus LDA[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(2): 228-233.
- [2] Alibeigi M, Hashemi S, and Hamzeh A. DBFS: an effective density based feature selection scheme for small sample size and high dimensional imbalanced data sets[J]. *Data & Knowledge Engineering*, 2012, 81/82: 67-103.
- [3] Friedman H. Regularized discriminant analysis[J]. *Journal of the American Statistical Association*, 1989, 84(405): 165-175.
- [4] Li M and Yuan B. 2D-LDA: a novel statistical linear discriminant analysis for image matrix[J]. *Pattern Recognition Letters*, 2005, 26(5): 527-532.
- [5] Ye J P and Xiong T. Computational and theoretical analysis of null space and orthogonal linear discriminant analysis[J]. *Journal of Machine Learning Research*, 2006, 7: 1183-1204.
- [6] Yu H and Yang J. A direct LDA algorithm for high-dimensional data with application to face recognition [J]. *Pattern Recognition*, 2001, 34(11): 2067-2070.
- [7] Wan M H, Lai Z H, and Jin Z. Feature extraction using two-dimensional local graph embedding based on maximum margin criterion[J]. *Applied Mathematics and Computation*, 2011, 217(23): 9659-9668.
- [8] Ji S W and Ye J P. Generalized linear discriminant analysis: a unified framework and efficient model selection[J]. *IEEE Transactions on Neural Networks*, 2008, 19(10): 1768-1782.
- [9] Chen L F, Liao H Y M, Ko M T, et al. A new LDA-based face recognition system which can solve the small sample size problem[J]. *Pattern Recognition*, 2000, 33(10): 1713-1726.
- [10] Belhumeur P N, Hespanha J P, and 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.
- [11] Hastie T, Tibshirani R, and Tibshirani R. Penalized discriminant analysis[J]. *The Annals of Statistics*, 1995, 23(1): 73-102.
- [12] Liu C J and Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. *IEEE Transactions on Image Processing*, 2002, 11(4): 467-476.
- [13] 刘忠宝, 王士同. 改进的线性判别分析算法[J]. *计算机应用*, 2011, 31(1): 250-253.

Liu Zhong-bao and Wang Shi-tong. Improved linear discriminant analysis method[J]. *Journal of Computer*

- Applications*, 2011, 31(1): 250-253.
- [14] 刘忠宝, 王士同. 多阶矩阵组合 LDA 及其在人脸识别中的应用[J]. 计算机工程与应用, 2011, 47(12): 152-155.
Liu Zhong-bao and Wang Shi-tong. Modified LDA based on linear combination of k-order matrices and its application to face recognition[J]. *Computer Engineering and Applications*, 2011, 47(12): 152-155.
- [15] 刘忠宝, 王士同. 一种改进的线性判别分析算法在人脸识别中的应用[J]. 计算机工程与科学, 2011, 33(7): 89-93.
Liu Zhong-bao and Wang Shi-tong. An improved LDA algorithm and its application to face recognition[J]. *Computer Engineering and Science*, 2011, 33(7): 89-93.
- [16] He X F and Niyogi P. Locality preserving projections[C]. *Advances in Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2003: 153-160.
- [17] Roweis S T and Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, 290(5500): 2323-2326.
- [18] Park H. ISOMAP induced manifold embedding and its application to Alzheimer's disease and mild cognitive impairment[J]. *Neuroscience Letters*, 2012, 513(2): 141-145.
- [19] Yang J, Zhang D, Yang J Y, et al. Globally maximizing locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(4): 650-664.
- 刘忠宝: 男, 1981 年生, 讲师, 研究方向为模式识别与机器学习.
潘广贞: 男, 1969 年生, 副教授, 研究方向为智能信息处理.
赵文娟: 女, 1983 年生, 助教, 研究方向为机器学习.