

## 鉴别投影嵌入及其在人脸识别中的应用

严 严 章毓晋

(清华大学电子工程系 北京 100084)

**摘要:** 该文提出了一种新的监督线性降维方法,称为鉴别投影嵌入(Discriminant Projection Embedding, DPE)。和常用的线性鉴别分析相比,鉴别投影嵌入可以更好地保留类内的局部几何位置信息和提取类间的鉴别结构信息。在人脸识别公用数据库上进行了一系列的实验,实验结果表明了该文方法的可行性和有效性。

**关键词:** 人脸识别; 监督线性降维方法; 图像差值模型; 鉴别投影嵌入

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2008)12-2902-04

## Discriminant Projection Embedding with Its Application to Face Recognition

Yan Yan Zhang Yu-jin

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

**Abstract:** A new supervised linear dimensionality reduction method called Discriminant Projection Embedding (DPE) is proposed. Compared with widely-used Linear Discriminant Analysis (LDA), DPE can preserve the within-class neighboring geometry and extract between-class relevant structures for classification more efficient. Experimental results on public face databases show the feasibility and efficiency of DPE.

**Key words:** Face recognition; Supervised linear dimensionality reduction; Image difference model; Discriminant projection embedding

### 1 引言

近年来,各种监督降维方法(Supervised Dimensionality Reduction, SDR)和非监督的降维方法(Unsupervised Dimensionality Reduction, UDR)受到了越来越多的关注<sup>[1]</sup>。前者主要有线性鉴别分析(Linear Discriminant Analysis, LDA)等,后者主要有主成分分析(Principal Component Analysis, PCA),局部保持投影(Locality Preserving Projection, LPP)等。这些方法的基本假设是:高维空间中的数据,如人脸,在空间中实际上属于一个低维的流形<sup>[2]</sup>。由于利用了类别的标签信息,监督降维方法通常在模式分类中的性能更好。其中,监督线性降维(Supervised Linear Dimensionality Reduction, SLDR)方法由于计算方便和简单,得到广泛的研究和应用<sup>[3]</sup>。

大多数的 SLDR 方法通过最大化下面的准则函数寻找最佳的线性投影矩阵  $P_{opt}$ <sup>[1]</sup>:

$$P_{opt} = \arg \max_P \frac{\text{tr}(P^T B P)}{\text{tr}(P^T W P)} \quad (1)$$

其中  $B$  和  $W$  通常代表类间度量矩阵和类内度量矩阵。

当  $W$  非奇异时,SLDR 的最佳投影矩阵等于  $W^{-1}B$  的特征矩阵。但在实际应用中(如人脸识别)会经常遇到所谓的

“小样本集”(Small Sample Size, 3S)问题。当训练的样本数少于样本的维数时,3S 问题就会发生。在这种情况下,类内度量矩阵  $W$  变成奇异矩阵。因此无法直接对 SLDR 计算投影矩阵。

现有的文献中对 3S 问题的处理包括下面一些方法。Zhao 等<sup>[4]</sup>通过对  $W$  的特征值增加一个小的扰动使得  $W$  变成非奇异矩阵进行降维。更通常的方法是利用 PCA 首先对高维的数据进行降维,在降维的空间中  $W$  变成非奇异矩阵,从而可求得投影矩阵。Bellhumeur 等<sup>[5]</sup>利用 PCA+LDA 的框架进行人脸识别。Yan 等<sup>[1]</sup>提出了 PCA+边缘 Fisher 分析(Marginal Fisher Analysis, MFA)的方法进行数据的降维。Bressan 等<sup>[6]</sup>采用 PCA+非参数鉴别分析(Nonparametric Discriminant Analysis, NDA)进行模式分类。Chen 等<sup>[7]</sup>使用 PCA+局部鉴别嵌入(Local Discriminant Embedding, LDE)的方法进行人脸识别。另外一种方法是利用  $W$  的零空间<sup>[8]</sup>或者  $B$  的主元空间<sup>[9]</sup>进行降维。上面的方法也可以推广到对应的核(Kernel)形式和张量(Tensor)形式<sup>[1]</sup>。

本文首先利用图像差值模型分析了监督线性降维方法是如何提取对分类有用的鉴别信息,并指出更好地刻画类内度量矩阵和类间度量矩阵反映变换差分量和身份差分量的变化是监督线性降维方法的关键。针对人脸识别问题,本文提出了一种新的监督线性降维方法,称为鉴别投影嵌入(Discriminant Projection Embedding, DPE)。和常用的 LDA

方法相比, DPE 可以更好地保留类内的局部几何位置信息和提取类间的鉴别结构信息。为了验证 DPE 的可行性和有效性, 本文将 DPE 应用于 3 个公共的人脸数据库。在不同人脸数据库上本文的方法都取得了很好的效果。

## 2 监督线性降维方法分析

下面从最近邻法(Nearest Neighbor, NN)分类的角度出发分析 SLDR 是如何提取对分类有用的鉴别信息。本文以人脸识别为例, 可以自然地推广到其他模式分类问题中。

假定有一个查询的人脸样本待分类, 设为  $\mathbf{x}$ 。最近邻法首先计算  $\mathbf{x}$  和数据库中各个人脸类别的差值, 把具有最小范数的差值(即最小人脸距离)的人脸类别作为识别的结果(分类的结果)。有两种不同的人脸图像差值: 人脸类内差值  $\Delta^I$  (Intra-class difference)和人脸类间差值  $\Delta^E$  (Inter-class difference)。对于人脸的类内差值而言, 包括了变换差分量  $\mathbf{T}$ (对应了不同的光照, 姿态, 表情下的人脸变化)和噪声分量  $\mathbf{N}$ 。对于人脸的类间差值, 包括了身份差分量  $\mathbf{I}$ (对应了不同的身份变化), 变换差分量  $\mathbf{T}$ , 噪声分量  $\mathbf{N}$ 。为简单起见, 记  $\Delta^I = \mathbf{T} + \mathbf{N}$  以及  $\Delta^E = \mathbf{I} + \mathbf{T} + \mathbf{N}$  (实际上是各种因素复杂的作用, 而不仅仅是加性的作用)。如果待识别人脸的类内距离小于类间距离, 那么人脸将被正确地识别和分类。但在人脸识别中, 由于噪声, 光照, 姿态, 表情等各种因素的影响, 类内距离可能会大于类间距离从而导致识别的错误。因此直接的距离比对并不是最优的方法。对于人脸识别而言, 本质的特征是身份的不同(身份差分量  $\mathbf{I}$ , 即试图提取的鉴别信息), 而光照, 姿态, 表情等变化因素(变换差分量  $\mathbf{T}$  和噪声分量  $\mathbf{N}$ )都会影响识别或者分类的结果, 是应该去除的。通常的方法是通过训练集寻找某个子空间(对应投影矩阵)。在子空间中, 类内差值和类间差值中的变换差分量和噪声分量被降低和去除, 而类间差值的身份差分量被很好地保留和增强。一般类内度量矩阵主要用来刻画变换差分量(一般也包括噪声分量), 而类间度量矩阵用来刻画身份差分量(一般也包括一定的变换差分量和噪声分量)。大部分的 SLDR 方法都是通过最大化投影后类间度量矩阵和类内度量矩阵的迹的比值(身份差分量增强, 变换差分量和噪声分量减少)求得投影子空间。一般把上面的模型称为图像差值模型<sup>[10]</sup>。

从上面的分析中可以看出, SLDR 的关键是如何刻画类内度量矩阵更好地反映变换差分量的变化以及刻画类间度量矩阵更好地反映身份差分量的变化。

## 3 鉴别投影嵌入

针对实际问题的不同, 应该利用不同的方法分别刻画类内和类间度量矩阵。针对人脸识别问题, 本文提出一种新的监督线性降维方法, 称为鉴别投影嵌入(DPE)。DPE 基于如下的考虑: Roweis 等人<sup>[11]</sup>的实验表明同一个人的人脸图像由于不同光照和姿态的变化在高维空间中呈现局部线性流形。一张人脸图像可以用若干距离相近的人脸线性表出。因此本

文采用邻域保持嵌入(Neighborhood Preserving Embedding, NPE)的方法<sup>[12]</sup>刻画模式的类内变化。采用 NPE 可以很好地保持同一类模式之间的几何位置关系。同时 NPE 更关注类内的局部结构信息。对于多数的实际分类问题, 局部结构信息更有利于分类<sup>[13]</sup>。对于复杂的类间变化, 本文采用 NDA 的思想<sup>[6,13]</sup>, 即求每个样本和其类间最近邻的差值组成的协方差矩阵, 提取类间的鉴别结构信息。

### 3.1 类内度量矩阵的刻画

Roweis 和 Saul 提出的局部线性嵌入(Locally Linear Embedding, LLE)算法<sup>[11]</sup>。LLE 算法假设数据中的每个点都可以通过周围的  $K$  近邻的线性组合表示出来。降维后, 通过保持这种线性组合的关系可以保留数据之间的几何位置关系。因此, LLE 算法可以很好地保留类内各个样本之间的几何位置关系。LLE 通过最小化式(2)得到权重矩阵  $\mathbf{A}$ 。

$$\min \sum_i \left\| \mathbf{x}_i - \sum_j \alpha_{ij} \mathbf{x}_j \right\|^2 \quad (2)$$

约束条件是不同类的节点之间的权重  $\alpha_{ij} = 0$ , 同类节点之间的权值和  $\sum_j \alpha_{ij} = 1$ 。

在上述约束条件下, 权重  $\alpha_{ij}$  对旋转, 尺度变换和平移具有不变性。原始的 LLE 算法只定义在训练数据上, 无法推广至测试数据。NPE 是 LLE 的线性化形式, 可以同时定义在训练数据和测试数据上面。假定  $\mathbf{P}$  是线性投影矩阵, 为了保证投影后的向量之间保持原始高维空间之间的空间位置关系, 损失函数相应的变成了:

$$\begin{aligned} J_{\min}(\mathbf{P}) &= \min \sum_i \left\| \mathbf{y}_i - \sum_j \alpha_{ij} \mathbf{y}_j \right\|^2 \\ &= \min_{\mathbf{P}} \sum_i \left\| \mathbf{P}^T \mathbf{x}_i - \sum_j \alpha_{ij} \mathbf{P}^T \mathbf{x}_j \right\|^2 \\ &= \min_{\mathbf{P}} \text{tr}(\mathbf{P}^T \mathbf{X}(\mathbf{I} - \mathbf{A})^T (\mathbf{I} - \mathbf{A}) \mathbf{X}^T \mathbf{P}) \\ &= \min_{\mathbf{P}} \text{tr}(\mathbf{P}^T \mathbf{W} \mathbf{P}) \end{aligned} \quad (3)$$

其中  $\mathbf{y}$  是投影后的数据,  $\mathbf{P}$  是投影矩阵,  $\mathbf{A}$  是权重矩阵。 $\mathbf{W} = \mathbf{X}(\mathbf{I} - \mathbf{A})^T (\mathbf{I} - \mathbf{A}) \mathbf{X}^T$  表示类内度量矩阵。 $\text{tr}(\cdot)$  代表了求迹运算,  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$  是原始的数据矩阵,  $n$  是总的训练样本数目。 $\mathbf{I} = \text{diag}(1, 1, \dots, 1)$ 。

### 3.2 类间度量矩阵的刻画

通常类间度量矩阵采用的是  $\mathbf{S}_B = \sum_{i=1}^c (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T$ ,

其中  $c$  是总的类别数目,  $\boldsymbol{\mu}_i$  是第  $i$  类的均值向量,  $\boldsymbol{\mu}$  而代表总的均值向量。由于采用参数化的散度矩阵(均值是二阶统计量), 无法处理超过二阶的统计量, 因此无法很好地表达复杂的数据分布(非单高斯分布)。Fukunaga<sup>[13]</sup>提出了一种非参数的鉴别分析方法试图解决这个问题。在 NDA 中由于采用了非参数的结构形式, 因此可以保留对分类有用的结构信息<sup>[6]</sup>。在 NDA 中, 类间度量矩阵的定义如下:

$$\mathbf{B} = \frac{1}{n} \sum_{i=1}^n w_i (\Delta_i^E) (\Delta_i^E)^T = \frac{1}{n} \Delta^E \Delta^E \quad (4)$$

其中  $\Delta^E = [\sqrt{w_1}\Delta_1^E, \sqrt{w_2}\Delta_2^E, \dots, \sqrt{w_n}\Delta_n^E]$ 。  $\Delta_i^E$  是  $x_i \in C_k$  和其最近邻  $x^E$  (或者  $K$  近邻均值) 的差值, 即  $\Delta_i^E = x_i - x^E$ 。  $x^E$  定义为  $x^E = \{x' \in \bar{C}_k / \|x' - x_i\| \leq \|z - x_i\|, \forall z \in \bar{C}_k, x_i \in C_k\}$ 。  $w_i$  代表第  $i$  个样本的权重。

### 3.3 鉴别投影嵌入

综合上面对类内和类间度量矩阵的刻画, 可以通过最大化准则函数式(5)得到投影矩阵。

$$\begin{aligned} P_{\text{opt}} &= \arg \max_P \frac{\text{tr}(P^T B P)}{\text{tr}(P^T W P)} \\ &= \arg \max_P \frac{\text{tr}\left(\frac{1}{n} P^T \Delta^E \Delta^E P\right)}{\text{tr}(P^T X (I - A)^T (I - A) X^T P)} \quad (5) \end{aligned}$$

由于利用邻域保持嵌入刻画类内的模式变化, 同时引入了鉴别分析的思想, 本文把上述监督线性降维方法称之为鉴别投影嵌入(Discriminant Projection Embedding, DPE)。利用本文提出的 DPE 由于可以很好地保留类内的局部几何位置信息和提取类间的鉴别结构信息, 从而更有效地提取有利于分类结果的鉴别信息。利用上面的图像差值模型, DPE 实际上利用了类内的局部流形变化有效地刻画变换差分变化, 而利用类间的最近邻差值变化刻画身份差分变化。最小化类内的局部流形变化可以有效地降低变换差分变化, 而最大化类间的最近邻差值变化可以使得投影轴(子空间)更关注和区分距离接近的人脸类别, 增强身份差分变化。并且和线性鉴别分析等方法求得投影向量个数小于总类别数不同, 由于类间度量矩阵采用非参数化的形式(式(4)), DPE 求得投影向量个数大于总的类别数, 从而可以得到更多的鉴别矢量。

## 4 实验结果及讨论

为了验证 DPE 的可行性和有效性, 本文在 3 个公共人脸数据库上进行了一系列的实验, 包括了 ORL 数据库, Bern 数据库和 FERET 数据库。比较了各种典型的监督降维方法, 包括了 LDA<sup>[5]</sup>, NDA<sup>[6]</sup>, LDE<sup>[7]</sup>, NDP<sup>[14]</sup>。为了克服上述所有方法中类内度量矩阵的奇异性问题, 本文利用 PCA 首先对人脸数据进行降维, 以保证类内度量矩阵的非奇异性。为了保证所有的算法在公平的测试环境下进行, 本文对数据库中的所有人脸做如下的处理: 首先大小归一化至  $64 \times 64$ , 然后进行直方图均衡化, 所有的方法均采用基于二范数的最近邻法分类。

### 4.1 实验设置和实验结果

ORL 库共有 40 个人组成, 每个人 10 幅图像。在 ORL 库中存在一定的光照, 表情(睁眼, 闭眼, 微笑等), 饰物的变化(眼镜)。Bern 库共有 30 个人组成, 每个人 10 幅图像。每个人包括了一定表情和姿态变化。ORL 库和 Bern 库的图像结构类似, 本文将之合并成一个数据库。合并后的数据库(ORL+Bern 库)共有 70 个人, 每个人 10 幅图像。

FERET 库是由美国军方发起的专门的人脸评测所采集

的人脸数据库。本文的算法在 FERET 的子库上进行。子库一共包含了 200 个人的 1400 幅人脸图像, 每个人有 7 幅图像。图像中包含了光照, 姿态, 表情等各种因素的变化。

本文对 ORL+Bern 库和 FERET 库都采用下面两种方式选取训练样本。

(1)对数据库中每一类人脸随机选取  $m$  张人脸( $m=4$ )作为训练, 剩下的图像作为测试。对于给定的  $m$ , 本文总共进行 20 轮实验, 每一轮的训练人脸随机选取。最后的结果是 20 轮的平均。图 1 中给出了鉴别矢量个数不同时各种方法人脸识别率变化情况。

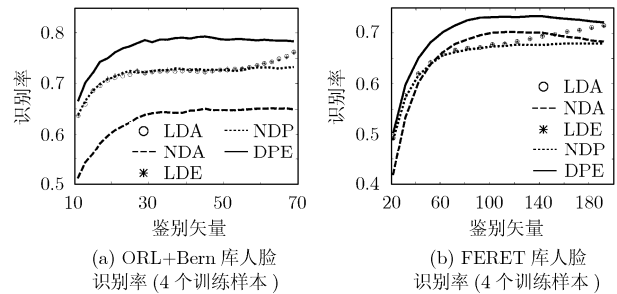


图 1 ORL+Bern 库和 FERET 库上各种方法的识别率随鉴别矢量个数的变化(方式(1))

(2)对数据库中每一类选取若干典型代表样本(包括姿态, 光照, 表情的变化)作为训练, 剩下的图像作为测试。图 2 给出了鉴别矢量个数不同时各种方法人脸识别率变化情况(选择的典型训练样本数为 4)。

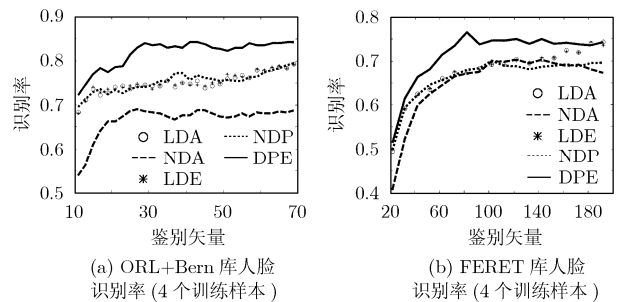


图 2 ORL+Bern 库和 FERET 库上各种方法的识别率随鉴别矢量个数的变化(方式(2))

为了定量的比较各种方法的性能, 表 1 给出了采用方式(1)和方式(2)时各种方法在 ORL+Bern 库和 FERET 库的最佳识别率和对应的鉴别向量个数。

### 4.2 分析与讨论

从实验结果可以看出, 在 ORL+Bern 库和 FERET 库上, DPE 的分类结果都要优于 LDA, NDA, LDE, NDP 等监督降维方法, 表明了 DPE 在人脸识别中具有很好的推广性。由于 DPE 利用了邻域保持嵌入刻画类内的模式变化, 不同的训练样本对最终的识别结果有很大的影响。从表 1 可

表 1 各种方法的最佳识别率和对应的鉴别向量的个数(方式(1)和方式(2))

训练方式	方式(1)		方式(2)	
算法	ORL+Bern 库(4个训练 样本)	FERET 库(4 个训练样本)	ORL+Bern 库(4个训练 样本)	FERET 库(4 个训练样本)
LDA	0.7626(69)	0.7160(192)	0.7929(69)	0.7433(192)
NDA	0.6829(209)	0.7029(112)	0.7190(209)	0.7033(122)
LDE	0.7615(69)	0.7151(192)	0.7929(69)	0.7433(182)
NDP	0.7329(69)	0.6803(172)	0.7952(69)	0.7017(102)
DPE	<b>0.7927(45)</b>	<b>0.7339(142)</b>	<b>0.8429(41)</b>	<b>0.7650(82)</b>

可以看出,方式(2)的DPE识别结果要明显优于方式(1)的DPE识别结果。选择具有代表性的典型训练样本可以极大提高DPE的识别率。

DPE方法和其他方法的不同在于同时利用了类内的局部结构信息和类间的全局鉴别结构信息,能够很好地提取对分类和识别最有利的鉴别信息。而LDA方法则是保持全局的欧式结构(Global Euclidean Structure)信息。NDA方法提取了类间最近邻之间信息,但是没有考虑类内样本之间的位置关系。和LDE,NDP方法不同,DPE更关心可能对分类造成影响的相近的人脸类别。

## 5 结束语

本文从最近邻分类的角度出发,分析了SLDR是如何提取有效的鉴别信息。SLDR的本质是寻找一定的子空间。在子空间中,类内差值和类间差值的变换差分量和噪声分量被降低和去除,而类间差值的身份差分量被很好地保留和增强。如何更好地刻画类内度量矩阵(反映变换差分量的变化)和类间度量矩阵(反映身份差分量的变化)是SLDR的关键问题。针对人脸识别问题,本文提出了一种新的SLDR方法,称为鉴别投影嵌入(DPE)。DPE利用了类内的局部流形变化有效地刻画变换差分量变化,而利用类间的最近邻差值变化刻画身份差分量变化。在不同的人脸数据库上表明了DPE的有效性。在未来的工作中,我们将在DPE的基础上进一步研究各种寻找投影子空间的方法,包括利用 $\mathbf{W}$ 的零空间<sup>[8]</sup>或者 $\mathbf{B}$ 的主元空间<sup>[9]</sup>进行降维的方法。

## 参 考 文 献

[1] Yan S C, Xu D, Zhang B, and Zhang H J. Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Trans. on Pattern Analysis and Machine*

- Intelligence*, 2007, 29(1): 40-51.
- [2] Moghaddam B. Principal manifolds and probabilistic subspace for visual recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, 24(6): 780-788.
- [3] Zhao W, Chellappa R, Phillips P J, and Rosenfeld A. Face recognition: A literature survey. *ACM Computing Surveys*, 2003, 35(4): 399-459.
- [4] Zhao W, Chellappa R, and Krishnaswamy A. Discriminant analysis of principal components for face recognition. Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998: 336-341.
- [5] Belhumeur P N, Hefanha J P, and Kriegman D J. Eigenfaces vs Fisherfaces: recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711-720.
- [6] Bressan M and Vitria J. Nonparametric discriminant analysis and nearest neighbor classification. *Pattern Recognition Letters*, 2003, 24(15): 2743-2749.
- [7] Chen H T, Chang H W, and Liu T L. Local discriminant embedding and its variants. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005: 846-853.
- [8] Chen L, Liao H, Ko M, Lin J, and Yu G. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognition*, 2000, 33(10): 1713-1726.
- [9] Yu H and Yang J. A direct LDA algorithm for high-dimensional data with application to face recognition. *Pattern Recognition*, 2001, 34(10): 2067-2070.
- [10] Wang X G and Tang X O. A unified framework for face recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2004, 26(9): 1222-1228.
- [11] Roweis S and Saul L K. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, 290(5500): 2323-2326.
- [12] He X, Cai D, Yan S, and Zhang H J. Neighborhood preserving embedding. Proceedings of IEEE International Conference on Computer Vision, Beijing, China, 2005: 1208-1213.
- [13] Fukunaga K. Introduction to Statistical Pattern Recognition. New York: Academic Press, 1990: 466-479.
- [14] You Q, Zhen N, Du S, and Wu Y. Neighborhood discriminant analysis for face recognition. *Pattern Recognition Letters*, 2007, 40(8): 2283-2291.

严 严: 男, 1984年生, 博士生, 研究领域为人脸识别、计算机视觉。

章毓晋: 男, 1954年生, 教授, 研究领域为图像工程(图像处理、图像分析、图像理解及其技术应用)。