

支持向量鉴别分析及在人脸表情识别中的应用

应自炉^{1,2}, 唐京海¹, 李景文¹, 张有为^{1,2}

(1. 北京航空航天大学电子信息工程学院, 北京 100083; 2. 五邑大学信息学院, 广东江门 529020)

摘 要: 模式识别一般首先要对数据进行降维, PCA 和 LDA 及其对应的核化算法是其中应用广泛的方法, 但这些算法的应用前提是假设样本数据为高斯分布, 在少样本训练时它们的推广性能有很大局限. 本文提出了一种基于支持向量机的鉴别分析算法, 该算法首先寻找有限样本情况下最优分类面, 以其法线方向为投影轴对数据进行投影降维, 在 multi-class 情况下提供了极其丰富的方案选择投影轴. 该算法体现了支持向量机的内在优良推广性能, 克服了 PCA 和 LDA 等算法的局限性. 本文将所提算法应用于人脸表情特征提取, 并与 PCA、LDA、KPCA、GDA 等算法进行了比较, 结果表明该算法的有效性.

关键词: 模式识别; 主元分析; Fisher 鉴别分析; 支持向量机; 表情识别

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2008) 04-0725-06

Support Vector Discriminant Analysis and Its Application to Facial Expression Recognition

YING Zi-lu^{1,2}, TANG Jing-hai¹, LI Jing-wen¹, ZHANG You-wei^{1,2}

(1. School of Electronics and Information Engineering, Beihang University, Beijing 100083;

2. School of Information, Wuyi University, Jiangmen, Guangdong 529020)

Abstract: Dimension reduction of data is usually an important preprocessing step in pattern recognition. PCA and Fisher's LDA and their kernelized versions are widely used approaches for dimension reduction. But they have limitations when used for small sample training because of their Gaussian distribution assumption. This paper propose an algorithm for dimension reduction called support vector discriminant analysis (SVDA), which first looks for the optimal separating hyperplane by SVM algorithm and then project data in the corresponding normal direction. In multiclass cases, the algorithm has many choices for selecting projecting axis. The algorithm has the intrinsic nice generalization ability of SVM. The paper applies the algorithm to the feature extraction in facial expression recognition application and compares the results to other algorithms such as PCA, LDA, KPCA and GDA. The results show the effectiveness of the proposed algorithm.

Key words: pattern recognition; principal component analysis (PCA); Fisher linear discriminant analysis (Fisher LDA); support vector machine (SVM); support vector discriminant analysis (SVDA); facial expression recognition

1 引言

在许多统计模式识别或统计学习问题如图像识别中, 由于原始数据的维数很高, 不适合直接进行处理. 数据维数太高会使得算法计算量太大, 鲁棒性变差, 算法太复杂而影响其推广性能等, 因此首先需要对输入数据进行降维. 应用广泛的传统降维方法有主元分析 (PCA) 法和 Fisher 线性鉴别分析 (LDA) 法等^[1]. PCA 选择信号重构时均方误差最小的投影方向, 它是无监督方法, 它没有利用训练样本的类别信息, PCA 有利于进行数据表示. Fisher LDA 是一种有监督方法, 它利用了训练样本

的类别信息, 其选择的投影方向使类间离散度与类内离散度的比值最大, 因此 LDA 充分保留了数据的类别结构, 有利于进行分类. 一般情况下, 为了分类目的, 就 PCA 和 LDA 方法来说, LDA 方法要优于 PCA 方法. 如图 1 所示为高斯分布、包含有两个类别的二维数据, 它们的 PCA 投影轴和 LDA 投影轴如图 1(a) 所示, 一维 PCA 和一维 LDA 投影及其分布如图 1(b) 所示. 从图 1(b) 可以看出, PCA 投影后两个类别的数据很大部分重叠在一起, 已经无法在一维特征上进行正确分类. 而 LDA 投影后两个类别的数据全部分离开, 所以能得到正确的分类.

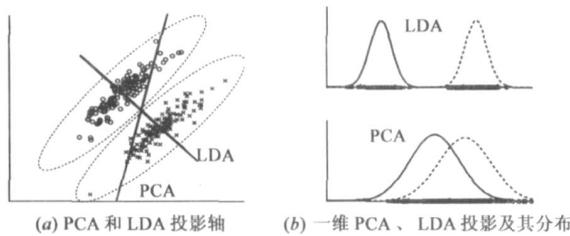


图 1 PCA 投影和 LDA 投影

无论是 PCA 还是 LDA 都是假定样本数据为高斯分布,它们利用了有限样本对均值和方差的估计来代替高斯分布中的均值和方差等参数,在这种估计很符合测试数据分布时,就能得到很好的结果,而如果这种估计与实际参数偏差较大时,就会得到比较差的结果。比如 A M Martinez 和 A C Kak^[2]就研究了很小样本情况下,即使

为了进行分类,PCA 不一定比 LDA 效果差,他们给了一个典型情况如图 2 所示,两类数据分布如图 2 中的虚线椭圆所示,图中小圆圈和小十字分别表示来自两个类别的训练数据,这时

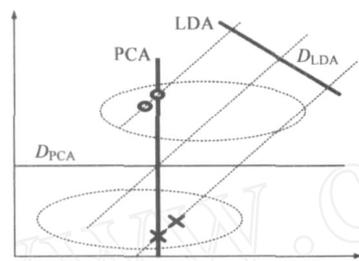


图 2 小样本时,PCA 投影有时也优于 LDA

由两个训练样本得到的投影轴分别为图中标有 PCA 和 LDA 的两条线,图中的 D_{PCA} 和 D_{LDA} 分别为 PCA 和 LDA 算法下的分类线,显然这种情况下 PCA 对整个样本空间的数据能得到更好的分类结果。

一般来说,在图像识别问题中,原始数据表现了非常强的非线性关系,而 PCA 和 LDA 处理的是线性关系,这时,可以利用某种非线性变换,将输入数据映射到某一高维甚至是无穷维的特征空间再利用 PCA 或 LDA 的方法,而这种映射也不需要显式进行,只需用核函数隐式进行就可以,这就得到了核化的 PCA 和 LDA,即 B Scholkopf 等所提出的 KPCA^[3],S Mika 等提出的 KFD^[4],G Baudat 和 F Anouar 提出的 GDA^[5]等方法。

无论是 PCA、LDA 或是核化的 KPCA、KFD 和 GDA 等,或是这些算法的各种变种,如陈松灿等人所提出各种形式的 PCA^[6-8]等,它们有一个共同特点,就是假设数据为正态分布,利用有限样本对均值和方差等的估计值来代替正态分布中的对应均值和方差,所以这些算法只有在这些参数估计比较准确的情况下才适用。图 3 所示为线性可分的两类样本,图 3(a)显示了对应的 PCA 和 LDA 投影轴,图 3(b)为对应的一维 PCA 和 LDA 投影及分布,可以看出,投影后两类数据中有些样本再不能进行正确分类,从图 3(a)可以看出,如果选择图中虚线作分类线,即选择虚线的法线作投影轴,则该例中的样

本是线性可分的,但 PCA 和 LDA 投影都不能找到正确的投影方向。

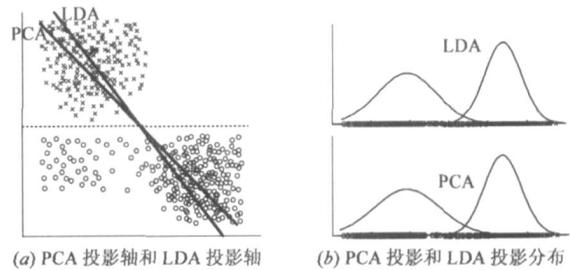


图 3 线性可分数据经 PCA 或 LDA 投影后不能进行正确分类

支持向量机(SVM)是一种有限样本情况下优秀的机器学习方法^[9],该算法从 1995 年 C Cortes 和 V Vapnik^[10]提出以来就得到了及其广泛的应用,其应用主要在模式分类和回归两个方面.支持向量机方法寻找有限训练样本情况下最优分类面,使得分类间隔达到最大,是 V Vapnik 的结构风险最小化原理的具体应用,在有限样本情况下学习结果具有优良的推广性能.本文将支持向量机算法应用在数据降维和特征提取方面,我们称之为支持向量鉴别分析(SVDA),该算法利用支持向量机内在的优良性能,比如核技巧及良好的推广性能等,寻找在有限样本情况下使得类间分离间隔最大的投影方向,得到降维数据具有最大类间分离性,在多类情况下,可以具有极其丰富的方案选择最佳投影轴,并且不需要假设数据的概率分布模型.下面首先介绍该算法,接着介绍该算法在人脸表情识别方面的应用,给出识别结果,并与 PCA、LDA、KPCA、GDA 等算法进行比较分析,最后讨论了有关结果,给出了我们的结论。

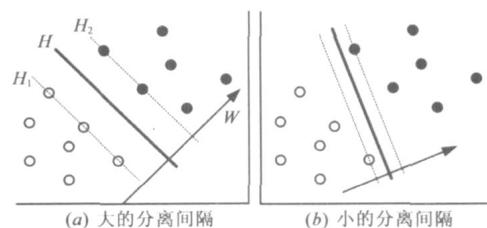


图 4 支持向量鉴别分析投影方向

2 支持向量鉴别分析

2.1 两类情况

设 $\{x_i, y_i\}, i = 1, \dots, l$ 为训练数据, $x_i \in \mathbf{R}^d$ 为 d 维训练样本, $y_i \in \{-1, 1\}$ 为样本 i 的类别信息.根据 SVM 理论^[10,11],为使训练时得到的算法具有对未知测试样本数据具有良好推广性,应该选择决策面 H 离最近训练样本最远,即使得图 4(a)所示 H_1 和 H_2 之间距离最大,而图 4(b)所示的决策面就不如 4(a)图,这就是 SVM 理论非常直观的几何解释.设决策面 H 方程为

$$H: x^T w + b = 0 \tag{1}$$

H_1 和 H_2 的方程分别为,

$$H_1: x^T w + b = -1 \quad (2)$$

$$H_2: x^T w + b = +1 \quad (3)$$

在线性可分的情况下,所有样本必须满足方程

$$x_i^T w + b = +1 \quad \text{for } y_i = +1 \quad (4)$$

$$x_i^T w + b = -1 \quad \text{for } y_i = -1 \quad (5)$$

H_1 和 H_2 之间的距离为 $2/ \|w\|^2$, 使得 H_1 和 H_2 之间距离最大就是使得目标函数 $\frac{1}{2} \|w\|^2$ 最小. 如果不是线性可分的情况, 则目标函数中要加上错分样本的惩罚项, 即最小化下面的函数

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (6)$$

而训练样本必须满足下面条件

$$x_i^T w + b = +1 - \xi_i \quad \text{for } y_i = +1 \quad (7)$$

$$x_i^T w + b = -1 + \xi_i \quad \text{for } y_i = -1 \quad (8)$$

对于支持向量分类, 我们需要求出最优分类面 $H: x^T w + b = 0$, 而进行 SVDA 分析时, 只需求出投影轴 w , 然后对任意样本 x , 可以求得投影坐标为 $y = w^T x$.

方程(6)是一个二次优化问题, 它的对偶问题为:

$$\max L_D = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \quad (9)$$

$$\text{s. t.} \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (10)$$

$$0 \leq \alpha_i \leq C, i = 1, \dots, l \quad (11)$$

其中 $x_i \cdot x$ 表示向量 x_i 和 x 的内积, 这时, 得到的投影轴为 $w = \sum_{i=1}^l \alpha_i y_i x_i$, 对任意样本 x 的投影坐标为

$$y = w^T x = \sum_{i=1}^l \alpha_i y_i (x_i \cdot x) \quad (12)$$

通过解对偶问题(9)得到投影公式(12). 可以看出, 投影运算只涉及到内积运算, 所以式(12)所示的线性投影方程很容易通过核函数方法推广到非线性情况.

根据支持向量鉴别分析的几何解释, 很容易知道, 对图 1、图 2 和图 3 所示的几种样本分布情况, 支持向量鉴别分析都能够找到正确的投影轴, 并且使得投影后不同类别的样本之间间隔达到最大.

2.2 非线性推广

前面介绍的是线性支持向量鉴别分析, 得到线性投影方程(12). 在诸如人脸表情识别这样的图像识别问题中, 其特征表现了非常强的非线性关系, 所以有必要将式(12)推广到非线性情况. 由于对偶问题式(9)及投影方程(12)都只涉及到内积运算, 所以借助于核技巧很容易将上面的线性方法推广到非线性情况. 设 $\phi: x_i$

(x_i) 为非线性映射, 它将输入数据映射到某一高维甚至是无限维的特征空间 H 中, 而在特征空间 H 中两个向量 $(x_i), (x_j)$ 的内积可以用核函数 $k(x_i, x_j)$ 来实

现, 即有

$$k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (13)$$

这时优化问题(9)变为

$$\max L_D = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (14)$$

$$\text{s. t.} \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (15)$$

$$0 \leq \alpha_i \leq C, i = 1, \dots, l \quad (16)$$

问题(14)的解为 $w = \sum_{i=1}^l \alpha_i y_i \phi(x_i)$, 它是高维特征空间 H 中的解向量, 但我们并不需要显式地得到 H 中的投影轴 w , 我们只要对任意样本 x 进行投影, 而这时对任意样本 x 的投影坐标变为:

$$y = w \cdot x = \sum_{i=1}^l \alpha_i y_i k(x_i, x) \quad (17)$$

式(17)只用到核函数. 常见的核函数有:

- (1) 线性核函数: $k(x_i, x_j) = x_i \cdot x_j$;
- (2) 多项式核函数: $k(x_j, x_j) = (x_i \cdot x_j + c)^d$;
- (3) 高斯径向基函数核函数: $k(x_i, x_j) = e^{-\frac{x_i - x_j}{2}^2}$;
- (4) Sigmoid 核函数: $k(x_i, x_j) = \tanh(c(x_i \cdot x_j) + d)$.

2.3 多类情况推广

前面讨论的是二类问题, 这种情况下可以选择的投影轴只有一个, 这类似于 Fisher LDA 分析. 对于多类情况, 我们可以通过下面几种方案选择不同的投影轴, 首先假设我们共有 c 类样本数据.

(1) 数据降为 $c - 1$ 维: 第一步在 c 类数据中任意选择一类数据作为第一类, 剩下的 $c - 1$ 类数据作为第二类, 进行前面介绍的基本二类投影, 这样可以得到一个投影轴, 这个投影把选作第一类的类别数据从其他类中分离出来. 第二步在剩下的 $c - 1$ 类数据中选任意一类作为第一类, 剩下的 $c - 2$ 类数据作为第二类, 进行基本的二类投影, 这样也可以得到一个投影轴, 分离出第二个类别数据. 以此类推, 到最后只有两类数据时, 可以得到一个投影轴, 分离最后两个类别的数据. 显然, 这样一共可以得到 $c - 1$ 个投影轴, 数据降到 $c - 1$ 维. 在每一步中选择作为第一类数据的选择方式不同, 使得这种方式下选出 $c - 1$ 个投影轴的方案也各不相同, 存在有很多方案.

(2) 数据降为 c 维: 选择其中一个类别数据作为第一类, 剩下的其他全部类别数据作为第二类, 进行基本二类投影, 可以得到一个投影轴. 这种方式下, 每个类别都有一次机会作为第一类, 所以一共可以得到 c 个投影轴, 数据降到 c 维. 这类似于分类法中的一对多处理过程.

(3) 数据降为 $c(c - 1)/2$ 维: 每次选择其中任意两

个类进行基本二类投影,而不考虑其他类别数据的影响,可以得到一个投影轴,从 c 个类中选二类出来共有 $c(c-1)/2$ 种方式,这样可以得到共 $c(c-1)/2$ 个投影轴,这类似于分类法中的一对一处理过程.

(4) 数据降为其他维数:选择其中某些类作为第一类,选择其他一些类作为第二类,进行基本二类投影.每一次不同的选择方案都可以得到一个投影,而 c 类数据这样组合成两类数据的方案是非常之多的,所以可以得到很多不同的投影轴,而每种投影轴都代表了对原来数据最大化分离的不同方案,上面(1)、(2)和(3)只是这种情况的特例.在这种选择方式下,要注意使得每一个类都有机会从其他类中分离开,比如可以将这里的方案同(1)、(2)或(3)结合使用.方案(1)、(2)或(3)本身也可以单独使用,或是结合使用.

在多类情况下,第(4)种选择方案给了 SVDA 分析非常丰富的方案选择不同的投影,具体如何选择可以根据问题的性质来决定,而这恰恰是 SVM 分类算法中不能提供的.在分类问题中,一般都是选择一对一或一对多的方式,而进行数据投影降维时,则可以根据我们目的希望将哪些类别从哪些类别中尽量分离开进行不同的组合来选择最优的投影轴.

SVDA 算法的计算复杂度相对样本数而言和 SVM 是一个数量级的,比如方案(2)和(3)就类似于一对多和一对一 SVM,而方案(1)比一对多 SVM 的计算量更少,方案(4)的计算量取决于具体选择的投影轴个数,它们都是和 SVM 在一个计算复杂度上.下面介绍 SVDA 算法在人脸表情识别中的应用.

3 SVDA 算法在人脸表情识别中应用

人脸表情识别是人机自然交互领域的一个热点课题,它在人机交互、安全防范、远程教育、远程商务、计算机游戏、多媒体制作、测谎、可视电话、电视会议、辅助医疗、行为科学、精神分析等领域中有着广泛的应用^[12,13].

本文将 SVDA 算法应用在人脸表情识别研究中,从而验证 SVDA 算法的有效性.实验使用日本女性人脸表情数据库 JAFFE 静态图像库. JAFFE 静态图像库由 10 人共 213 幅图像组成,每人包含愤怒、厌恶、恐惧、高兴、中性、悲伤和惊奇共 7 种表情,图 5 是其中一组惊奇表情图像,图 6 所示为对应于图 5 所剪切出的纯表情人脸图.

首先将图 6 所示表情图像归一化到 30×30 像素大小,亦即每个样本数据为 900 维.在 JAFFE 库上,选择每人每种表情一幅图像作训练,因为 JAFFE 库有 10 个人,所以每种表情共有 10 个样本图像作为训练,7 种表情共 70 个样本数据构成训练集.选择剩下的每人每种表

情二幅图像作测试,即每种表情共有 20 个样本图像作为测试,7 种表情共 140 个样本构成测试集.



图 5 JAFFE 人脸表情库中一组惊奇表情



图 6 剪切后对应图 5 的一组惊奇表情

表 1 列出了我们得到的实验结果,每行表示的是同一种特征提取方法而不同分类器得到的表情识别结果,每一列表示的是同一分类器但不同特征提取算法得到的表情识别结果.其中分类器 KNN1 表示的是 k 最近邻 ($k=1$) 分类器, KNN3 表示的是 k 最近邻 ($k=3$) 分类器, OAASVM 是一一对多方式的多类 SVM 分类算法, OAOSVM 是一一对一方式的多类 SVM 分类算法, BSVM2 是 V. Franc^[15] 等实现的多类 SVM 分类算法.特征提取算法包含 PCA、KPCA、LDA、GDA、SVDA 等算法,各个算法后面的数字表示特征提取后保留的特征个数,即投影轴的个数,亦即降维后数据的维数.比如 PCA(7) 表示用 PCA 方法将原来 900 维样本降为 7 为特征样本.所有用到核函数的算法都是用高斯径向基函数核函数,参数取 $=3$.

表 1 各种特征降维方法与各种分类器组合得到的表情识别率

降维算法	KNN1	KNN3	OAASVM	OAOSVM	BSVM2
PCA(7)	78.57 %	72.14 %	77.14 %	80.00 %	77.86 %
KPCA(7)	80.00 %	76.43 %	50.00 %	46.43 %	52.14 %
LDA(7)	81.43 %	74.29 %	76.43 %	77.14 %	75.71 %
GDA(6)	88.57 %	88.57 %	91.43 %	89.29 %	91.43 %
SVDA(6)	89.21 %	89.21 %	87.14 %	90.00 %	90.07 %
SVDA(7)	91.43 %	90.00 %	91.43 %	90.00 %	91.43 %
PCA(21)	83.57 %	82.14 %	90.00 %	90.00 %	90.71 %
KPCA(21)	81.43 %	80.00 %	72.86 %	77.14 %	82.14 %
LDA(21)	83.57 %	75.00 %	80.71 %	76.43 %	79.29 %
SVDA(21)	91.43 %	89.29 %	91.43 %	91.43 %	90.00 %
SVDA(28)	92.14 %	90.71 %	92.86 %	92.14 %	92.86 %
PCA(70)	83.57 %	82.14 %	91.43 %	92.14 %	92.14 %
KPCA(70)	83.57 %	82.14 %	93.57 %	92.86 %	92.86 %
LDA(70)	85.71 %	79.29 %	93.57 %	93.57 %	92.14 %
SVDA(70)	93.57 %	91.43 %	93.57 %	93.57 %	92.14 %

表 1 中的 SVDA(6) 选择了 6 个投影轴,即 2.3 节中介绍的方案(1). SVDA(7) 选择了 7 个投影轴,即 2.3 节中介绍的方案(2). SVDA(21) 选择的是任意两个类进行二类 SVDA 投影,即 2.3 节中介绍的方案(3). SVDA(28) 选择了 SVDA(7) 和 SVDA(21) 两种方式的结合,共有 28 个投影轴. SVDA(70) 选择的是 SVDA(7) SVDA(21) 中的 28 个投影轴,以及 3 类对剩下 4 类这种选择中的 42 种投影轴. 为了可比性,PCA, KPCA 等选择的投影轴个数与 SVDA 一样,而 GDA 最多只有 $c-1$ 个投影轴,所以只有 GDA(6). 没有选择更少的投影轴是因为对 SVDA 而言,比 6 更少的投影轴个数已不能将 7 个类的类别结构完全反映出来,这时的识别率很低,其实可以看出,对 PCA, KPCA, LDA 等算法,选择 7 个特征时识别率已经比较低了. 表 1 中的特征个数最大只取 70 是因为 KPCA 算法只能取到最大样本个数,在我们实验中,训练样本只有 70 个,所以特征个数最多取 70.

由表 1 可以看出,SVDA 算法的优势是显而易见的. 特别是在特征提取维数较低时,GDA 和 SVDA 相对其他的特征提取方法有明显的优势,但 GDA 只有一种方案. 比较奇怪的是在特征个数比较少的时候,KPCA 方法在我们的实验中识别率一直偏低,甚至不如 PCA 算法. 如果特征维数取到 70 时,各种特征提取方法的优劣区别已经不大,但 SVDA 还是有一些微弱优势,SVDA 算法的优越性一方面来自 SVM 算法的内在优势,另一方面则是来自 SVDA 提供了很多的投影轴选择方案,使得数据的分类信息能尽可能被充分利用. 从表 1 还可以看出,对 SVDA 而言,特征提取后各种分类器的分类效果都差不多,并且 k 最近邻($k=1$)方法基本上已经达到最佳了. 总体来说,SVDA 方法基本上是最优的.

4 结论

模式识别中,进行原始数据降维处理是很重要的一个环节,一方面它可能降低噪声的影响,另一方面可以大大减少后续算法的计算量,同时还可以增强算法的鲁棒性等,所以降维方法的研究一直是模式识别领域中一个热门的研究课题. 本文提出了一种基于 SVM 的数据降维处理算法 SVDA,它具备了 SVM 算法内在的许多性质,比如良好的对未知样本的推广性能,容易借助于核函数进行非线性推广等. 同时,SVDA 又不同于 SVM 分类算法,这可以从几个方面来看:首先 SVDA 不需要寻找决策面,不需要构建决策函数,其次作为特征提取算法所得结果可以结合其他特征提取算法如人脸表情识别中的几何特征提取法相结合,作为后续分类器输入数据,再者 SVDA 算法提供了比 SVM 更为广泛得多的投影轴选择方案,比如我们实验中的 SVDA(6)、SVDA(28) 和 SVDA(70) 等,这是 SVM 算法不具备的. 本文将 SVDA

算法在 JAFFE 人脸表情库上进行了广泛的各种实验,并与其他常用的数据降维算法包括 PCA、KPCA、LDA、GDA 等进行了对比分析,从实验结果看,SVDA 算法是非常有效的,也许最重要的一个原因是它没有对数据的分布作任何假设,而在有限样本情况下要得到原始数据的真实概率分布是比较困难的.

致谢 本文研究得到广东省自然科学基金项目(07010869),北京大学视觉与听觉信息处理国家重点实验室开放课题基金项目(0505),江门市科技攻关项目“人机自然交互中的情感状态识别理论与应用研究”,浙江大学 CAD & CG 国家重点实验室开放课题(A0703)的支持,在此表示感谢,同时 Ptools^[14]和 Stptools^[15]的免费提供使用使得本论文实验结果的获得争取了时间. 在此一并表示衷心的感谢.

参考文献:

- [1] R O Duda, P E Hart, D G Stork. Pattern Classification [M]. John Wiley & Sons. 2nd Edition. 2001.
- [2] A M Martinez, A C Kak. PCA versus LDA [J]. IEEE TPAMI, 2001, 23(2): 228 - 233.
- [3] B Scholkopf, A Smola, K R Muller. Kernel principal component analysis [A]. In W. Gerstner, A. Germond, M. Hasler, and J. - D. Nicoud, editors, Artificial Neural Networks-ICANN '97 [C]. Berlin, 1997, 583 - 588.
- [4] S Mika, G Rätsch, J Weston, B Schölkopf, K R Müller. Fisher discriminant analysis with kernels [A]. In Y. - H. Hu, J. Larsen, E. Wilson, and S. Douglas (Eds.). Neural Networks for Signal Processing IX [C]. IEEE, 1999. 41 - 48.
- [5] G Baudat, F Anouar. Generalized discriminant analysis using a kernel approach [J]. Neural Computation, 2000, 12(10): 2385 - 2404.
- [6] Chen Songcan. Feature Extraction Based on Linear Algebra with applications [OL]. <http://www.iip1.fudan.sh.cn/ml/chensc.ppt>. 2004.
- [7] Chen Songcan, Zhu Yulian. Subpattern-based principle component analysis [J]. Pattern Recognition 2004, 37(5): 1081 - 1083.
- [8] Chen Songcan, Zhang Daoqiang, Zhou Zhihua. Enhanced (PC) 2A for face recognition with one training image per person [J]. Pattern Recognition Letters, 2004, 25(10): 1173 - 1181.
- [9] 王珏, 周志华, 周傲英. 机器学习及应用 [M]. 清华大学出版社, 2006.
Wang Yu, Zhou Zhihua, Zhou Aoying. Machine Learning and Its Application [M]. Tsinghua University Press, 2006. (in Chinese)
- [10] C Cortes, V Vapnik. Support vector networks [J]. Machine Learning, 1995, 20(3): 273 - 297.

- [11] B Scholkopf, A J Smola, Learning with Kernels[M]. The MIT Press, MA, 2002.
- [12] 杨国亮, 王志良, 王国江. 面部表情识别研究进展[J]. 自动化技术与应用, 2006, 25(4): 1 - 6.
Yang Guoliang, Wang Zhiliang, Wang Guojiang. A survey of facial expression recognition [J]. Techniques of Automation and Applications, 2006, 25(4): 1 - 6. (in Chinese)
- [13] P S Aleksic, A K Katsaggelos. Automatic facial expression recognition using facial animation parameters and multistream HMMs[J]. IEEE Trans on Information Forensics and Security, 2006, 1(1): 3 - 11.
- [14] R P W Duin, P Juszczak, P Paclik, E Pekalska, D de Ridder, D M J Tax, PRTools4, A Matlab Toolbox for Pattern Recognition [Z]. Delft University of Technology, 2004.
- [15] V Franc. MATLAB STPRTOOLS [CP]. <http://cmp.felk.cvut.cz/~xfrancv/stprtool/>

作者简介:



应自炉 男, 1966 年 1 月生于江西上饶, 1988 年毕业于北京航空航天大学获硕士学位, 现为北京航空航天大学在职博士, 五邑大学副教授. 主要研究方向是信号与信息处理、人机自然交互、模式识别、人脸表情识别.
E-mail: ziluy@163.com



唐京海 男, 1975 年 1 月生于湖南东安, 2001 年毕业于西安空军工程大学获硕士学位, 现为北京航空航天大学在职博士, 服役空军 94786 部队. 主要研究方向是人机自然交互、模式识别、人脸表情识别.
E-mail: tangjh999@sina.com

李景文 男, 1964 年 8 月生, 北京航空航天大学教授, 博士生导师. 主要研究方向信号与信息处理、目标检测与成像、图像处理、图像识别、信息融合.

张有为 男, 1937 年 7 月生, 五邑大学教授、北京航空航天大学博士生导师. 主要研究方面是检测与估计理论、人机自然交互、语音识别、情态识别.