

# 手写字符识别的非线性主动判别函数

孙广玲

(上海大学通信与信息工程学院, 上海 200072)

**摘要** 提出了用于手写字符识别的非线性主动判别函数,是线性主动判别函数在手写字符非线性变化情况下的推广。该方法利用 Kernel PCA 分析捕捉和表示这种非线性变化。将输入空间非线性映射为特征空间,在特征空间的主子空间中生成最优主动原型模板,其与字符特征向量在特征空间主子空间的投影之间的距离即为非线性主动判别函数;同时,基于最小分类错误准则对该函数进行了优化。实验结果表明,非线性主动判别函数获得了比线性主动判别函数更高的识别率。

**关键词** 字符识别 非线性主动判别函数 Kernel PCA

中图法分类号:TP39 文献标识码:A 文章编号:1006-8961(2008)10-1853-04

## Nonlinear Active Discriminant Functions for Handwritten Character Recognition

SUN Guang-ling

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

**Abstract** A generalization of Linear Active Discriminant Functions named as Nonlinear Active Discriminant Functions (nonlinear ADF) to deal with nonlinear deformations of handwritten character is proposed. In Nonlinear ADF, Kernel PCA is applied to capture and represent the nonlinear deformations. Input space is mapped to feature space through nonlinear mapping. Then an optimal active prototype model is produced in principal subspace of the feature space and the distance between it and the projection of character feature vector in the principal subspace is defined as Nonlinear ADF. In addition, the Nonlinear ADF is further optimized using Minimum Classification Error criterion. Experimental results demonstrated that Nonlinear ADF has achieved a higher recognition rate than that of Linear ADF.

**Keywords** handwritten character recognition, nonlinear active discriminant functions, Kernel PCA

## 1 引言

手写字符识别是一个古老的话题,近几十年来,学者们提出了众多方法。其中,“变形模板”是一个独特的并且非常有效和有前途的技术,已在各种图像处理 and 模式识别的问题中得到广泛应用,它擅长于处理非刚性的、变化规律可遵循一定形式的对象。对于手写字符识别而言,也已提出很多基于变形模板的方法<sup>[1-4]</sup>。然而,文献[1]、[2]提出的方法只适用于拓扑结构简单和主要是曲线笔划的西方字符,如阿拉伯

数字和英文字母等,而并不适用于拓扑结构复杂和以直线笔划为主的东方字符,如中国汉字和日文汉字等;文献[3]、[4]的方法既适用于简单字符也适用于复杂字符。在学习阶段,学习字符变化的知识;在识别阶段,应用学习阶段获得的知识,构造主动手写模板以最佳地匹配待识别字符。实验结果表明,主动手写模板方法对于识别手写字符比较有效。但是,主动手写模板仍然存在若干缺点。首先,该方法依赖学习样本的形状信息捕捉形状的变化,而形状通过字符关键点坐标描述。不幸的是,关键点需要手工标注,这不仅是一项非常困难和枯燥的工作,而且会因某些主

观因素额外地引入误差;其次,该方法是基于字符图像骨架的,而骨架是必须通过细化算法提取出来的,然而,细化过程将不可避免地降低计算效率;再有,文献[3]、[4]的方法在识别汉字时,主动手写模板是对汉字的部件建模,而汉字部件的定义及选择都是比较困难的问题,这不可避免地将影响到该方法对于手写汉字识别的效果。在以前的工作中,曾提出主动判别函数用于手写字符的识别<sup>[5]</sup>。相对于文献[1]、[2]提出的方法,主动判别函数的优势是它既可识别简单字符也可识别复杂字符,并且计算效率较高;相对于文献[3]、[4]提出的方法,主动判别函数的优势是学习样本的描述形式是统计特征向量,可由计算机自动提取,不再需要人工标注获取;并且它是基于字符整体的策略,因而对识别汉字更为有效。然而正如文献[4]所指出的,手写字符普遍存在着因人因时的不同而导致的高度非线性变化,而以前提出的主动判别函数仅依赖于手写字符特征线性变化的假设,从而极有必要将其推广至更一般的非线性的情况。若仍利用 PCA (principal component analysis) 分析方法捕捉和表示此非线性变化,将自然地引入非线性 PCA 技术。Kernel PCA 是一种基于核方法的非线性 PCA,当核函数是两向量的内积时,就退化为线性 PCA<sup>[6]</sup>。

## 2 基于 Kernel PCA 的非线性最优主动原型模板

### 2.1 非线性主动原型模板

类似于文献[5]中的方法,仍在学习样本和利用 PCA 分析产生的本征空间中生成主动原型模板,区别在于此本征空间是基于非线性 Kernel PCA 产生的,因此,称为非线性主动原型模板。因为样本的能量集中在前  $k$  个主成分上,因此,只考虑在由相应的前  $k$  个本征向量张成的主子空间中生成主动原型模板。用  $\Gamma'$  表示这个主子空间中的变形模板,它具有下面的形式:

$$\Gamma' = \Psi + b$$

$$\Psi = [\Phi(\varphi)]^T V, \quad V^T V = I_{k \times k} \quad (1)$$

$\Phi$  是 Kernel PCA 中的非线性映射  $\Phi: R^{(N)} \rightarrow F$ , 即由输入空间  $R^{(N)}$  映射至特征空间  $F$ ;  $\varphi$  是学习样本在输入空间的均值  $\varphi = \left( \sum_{i=1}^m x_i \right) / m$ ;  $V$  是特征空间的主子空间,由  $k$  个列向量组成,每个列向量是特征空间的特征向量,在组成矩阵  $V$  时,按照本征值由大到小排序,其计算方法参见文献[6];  $\Psi$  是

$\varphi$  在特征空间主子空间的投影;注意到主子空间的各维分量是互不相关的,因此,调整参数  $b$  的各维分量可独立地设置或求取。为避免  $\Gamma'$  偏离  $\Psi$  过远,  $b$  应被约束在一定范围之内,即

$$-\theta_j \leq b_j \leq \theta_j, \theta_j > 0, j = 1, 2, \dots, k \quad (2)$$

### 2.2 非线性最优主动原型模板

将字符样本在输入空间的特征向量  $x$  在某类别特征空间主子空间中的投影与主动原型模板向量之间的乘方范数距离作为优化目标,即搜索最优主动原型模板,使这距离最小。下式是对这个乘方范数距离的定义:

$$\|x' - \Gamma'\|_{p/r} = \left( \sum_{j=1}^k |x'_j - \Gamma'_j|^p \right)^{1/r} \\ x'_j = [\Phi(x)]^T V^{(j)} \quad (3)$$

将式(1)代入到式(3)中,可进一步得到:

$$\|x' - \Gamma'\|_{p/r} = \left( \sum_{j=1}^k |\bar{x}'_j - b_j|^p \right)^{1/r} \\ \bar{x}'_j = [\Phi(x) - \Phi(\varphi)]^T V^{(j)} \quad (4)$$

参数  $p$  是控制各维之间差异的权重,参数  $r$  控制了当向量之间逐渐分开时,它们之间的距离增长的速度。

现在的任务是在式(2)的约束下,搜索  $b^*$ , 使其最小化  $\|x' - \Gamma'\|_{p/r}$ 。如上所述,主子空间的各维分量是互不相关的,所以  $b_j^*, 1 \leq j \leq k$  可独立搜索,不难得到:

$$b_j^* = \begin{cases} \bar{x}'_j & \text{if } -\theta_j \leq \bar{x}'_j \leq \theta_j \\ -\theta_j & \text{if } \bar{x}'_j < -\theta_j \\ \theta_j & \text{if } \bar{x}'_j > \theta_j \end{cases} \quad \theta_j > 0, 1 \leq j \leq k \quad (5)$$

从而非线性最优主动原型模板  $\Gamma'^*$  为

$$\Gamma'^* = \Psi + b^* \quad (6)$$

## 3 非线性主动判别函数及其优化算法

为了区分不同的类别,将式(6)中的每个符号加以下标表示,例如第  $i$  个类别的最优非线性主动原型模板表示为  $\Gamma'_i^* = \Psi_i + b_i^*, 1 \leq i \leq C, C$  是类别数,从而定义一种新的输入空间向量  $x$  与类别  $i$  的距离:

$$D_i(x; \Lambda_i) = \|x' - \Gamma'_i^*\|_{p/r} \\ = \left( \sum_{j=1}^k f_p(x; \Psi_i, V_i^{(j)}, \theta_{i,j}) \right)^{1/r} \\ \Lambda_i = \{ \Psi_i, V_i, \theta_i \}$$

$$f_p(\mathbf{x}; \Psi_i, \mathbf{V}_i^{(j)}, \theta_{i,j}) = \begin{cases} (\bar{x}'_{i,j} - \theta_{i,j})^p & \theta_{i,j} < \bar{x}'_{i,j} \\ (-\bar{x}'_{i,j} - \theta_{i,j})^p & \bar{x}'_{i,j} < -\theta_{i,j} \\ 0 & -\theta_{i,j} \leq \bar{x}'_{i,j} \leq \theta_{i,j} \end{cases}$$

$$\bar{x}'_{i,j} = [\Phi(\mathbf{x})]^T \mathbf{V}_i^{(j)} - \Psi_{i,j} = \left[ \sum_{l=1}^{m_i} \alpha_{i,l}^{(j)} K(\mathbf{x}, \mathbf{x}_{i,l}) \right] - \Psi_{i,j}$$

$$\Psi_{i,j} = [\Phi(\boldsymbol{\varphi}_i)]^T \mathbf{V}_i^{(j)} \quad (7)$$

$\mathbf{x}_{i,l}$  表示第  $i$  类学习样本,  $m_i$  是第  $i$  类学习样本的个数,  $K(\cdot, \cdot)$  为核函数,  $\alpha_{i,l}^{(j)}$  的含义及计算方法详见文献[6];为了简化计算并且不失一般性,将乘方范数距离中的  $p, r$  参数具体化为  $p = r = 1$ , 此即非线性主动判别函数(ADF),用符号  $g_{\text{ADF},i}(\mathbf{x}; \mathbf{A}_i)$  表示:

$$g_{\text{ADF},i}(\mathbf{x}; \mathbf{A}_i) = \sum_{j=1}^k f_1(\mathbf{x}; \Psi_i, \mathbf{V}_i^{(j)}, \theta_{i,j}), \theta_{i,j} = \sigma_{i,j} \quad (8)$$

显然,搜索边界  $\theta_{i,j}$  值可取为与第  $i$  类特征空间的特征空间各本征值的平方根  $\sqrt{\lambda_{i,j}}$  成正比关系,即  $\theta_{i,j} = a \sqrt{\lambda_{i,j}}, a > 0$ ,不妨就设  $\theta_{i,j} = \sqrt{\lambda_{i,j}}$ 。识别时,  $\mathbf{x}$  所属类别应为  $g_{\text{ADF},i}(\mathbf{x}; \mathbf{A}_i)$  取得最小值对应的类别。

如同文献[5]中提出的基于最小分类错误准则(MCE)优化线性 ADF,这里,仍然可以使用 MCE 准则优化非线性 ADF。为了表述的方便,今后将  $g_{\text{ADF},i}(\mathbf{x}; \mathbf{A}_i)$  简记为  $g_i^*(\mathbf{x}; \mathbf{A}_i)$ 。为了实施 MCE 准则,首先要得到目标函数,其过程参考文献[5]相应内容。

(1) 初始化参数集  $\mathbf{A}_{\text{tot}}$

$\Psi_i$  的初始值如式(1)计算;要求  $\theta_{i,j} > 0$ ,然而随机逼近法只适用于没有约束的问题,为此,引入新的参数  $\bar{\theta}_{i,j}$  且  $\theta_{i,j} = \exp(\bar{\theta}_{i,j}), \bar{\theta}_{i,j} \in \mathbf{R}$ ,此时一定满足  $\theta_{i,j} > 0$ 。 $\theta_i$  的初始值可根据本征值设置,即  $\theta_i = (\lambda_i)^{1/2}$ ,在初始化  $\theta_i$  之后,可根据上述关系得到  $\bar{\theta}_i$  的初始值。

(2) 对各样本依次按下式重复修正  $\Psi_i, \bar{\theta}_i$ ,直至满足收敛条件;在算法收敛之后,再根据  $\theta_i$  与  $\bar{\theta}_i$  之关系得到  $\theta_i$ 。

$$\Psi_{s,j}(t+1) = \Psi_{s,j}(t) + \varepsilon_i \cdot A(\mathbf{x}) \cdot \frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \Psi_{s,j}} \quad \text{for } s = q$$

$$\Psi_{s,j}(t+1) = \Psi_{s,j}(t) - \varepsilon_i \cdot A(\mathbf{x}) \cdot B(\mathbf{x}) \cdot C(\mathbf{x}) \cdot \frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \Psi_{s,j}} \quad \text{for } s \neq q$$

$$\bar{\theta}_{s,j}(t+1) = \bar{\theta}_{s,j}(t) + \varepsilon_i \cdot A(\mathbf{x}) \cdot \frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \theta_{s,j}} \cdot \exp(\bar{\theta}_{s,j}) \quad \text{for } s = q$$

$$\bar{\theta}_{s,j}(t+1) = \bar{\theta}_{s,j}(t) - \varepsilon_i \cdot A(\mathbf{x}) \cdot$$

$$B(\mathbf{x}) \cdot C(\mathbf{x}) \cdot \frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \theta_{s,j}} \times$$

$$\exp(\bar{\theta}_{s,j}) \quad \text{for } s \neq q$$

$$\mathbf{x} \in \boldsymbol{\omega}_q, 1 \leq s \leq C, 1 \leq j \leq k$$

$$A(\mathbf{x}) = \frac{-\zeta \cdot e^{\zeta \cdot [h_q(\mathbf{x}; \mathbf{A}_{\text{tot}}) + \alpha]}}{(1 + e^{\zeta \cdot [h_q(\mathbf{x}; \mathbf{A}_{\text{tot}}) + \alpha]})^2}$$

$$B(\mathbf{x}) = \frac{1}{C-1} \cdot [h_q(\mathbf{x}; \mathbf{A}_{\text{tot}}) + g_q^*(\mathbf{x}; \mathbf{A}_q)]^{1-\eta}$$

$$C(\mathbf{x}) = [g_s^*(\mathbf{x}; \mathbf{A}_s)]^{\eta-1}$$

$$\frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \Psi_{s,j}} = \begin{cases} -1 & \theta_{s,j} < \bar{x}'_{s,j} \\ 1 & \bar{x}'_{s,j} < -\theta_{s,j} \\ 0 & \text{其他} \end{cases}$$

$$\frac{\partial g_s^*(\mathbf{x}; \mathbf{A}_s)}{\partial \theta_{s,j}} = \begin{cases} -1 & \theta_{s,j} < |\bar{x}'_{s,j}| \\ 0 & \text{其他} \end{cases} \quad (9)$$

### 4 实验结果和分析

学习与测试的数据来自于银行票据中的手写账目汉字,如图 1 所示的部分样本,可以看出,样本的变化是非常大的。包括“零壹贰叁肆伍陆柒捌玖拾佰仟万亿元圆角分整正”共 21 个类别,学习与测试的样本都通过手工方式从字符串中分割出单个的字符,每个类别共收集了 400 个样本,并随机分成两部分,其中 300 个是学习样本,100 个是测试样本。

预处理和特征向量提取环节采用了文献[7]

的方法提取了 392 维梯度特征,然后应用“从类平均向量中提取鉴别信息”的线性特征变换方法将原 392 维特征变换为 196 维特征。

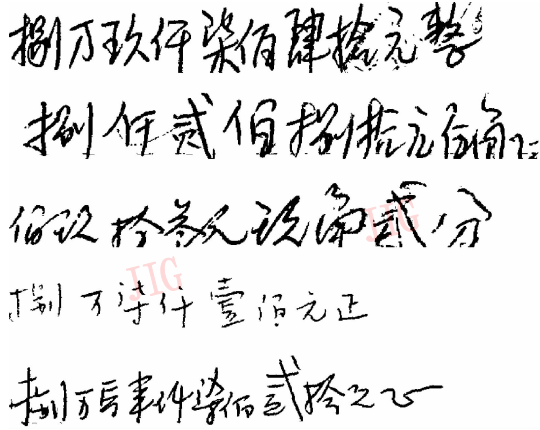


图 1 手写账目汉字样本

Fig. 1 Handwritten account Chinese character samples

基于 MCE 优化准则的优化算法参数设置为  $\zeta = 0.35$ ,  $\alpha = 0$ ,  $\varepsilon_0 = 0.1$ ,  $\varepsilon_t = 1/t$ ,  $t \geq 1$ , 迭代次数设为 20。

核函数选择 3 阶多项式核函数。

在上述实验参数的设置下,得到了经过 MCE 准则优化后的线性 ADF 和非线性 ADF 在不同主成分个数下的识别率,分别如表 1 和表 2 所示。

表 1 线性 ADF 的不同主成分个数识别率

Tab. 1 Accuracy of linear ADF of different principal component number

	单位: %			
主成分个数	32	64	96	128
识别率	99.42	99.46	99.49	99.52

表 2 非线性 ADF 的不同主成分个数识别率

Tab. 2 Accuracy of nonlinear ADF of different principal component number

	单位: %			
主成分个数	64	128	192	256
识别率	99.50	99.54	99.59	99.61

由于 Kernel PCA 可以提取超过原始特征维数的主成分个数,因此,非线性 ADF 在 256 个主成分时获得了最高的识别率。

## 5 结 论

本文提出了非线性主动判别函数,是以前提出的主动判别函数在非线性情况下的推广。首先在利用 Kernel PCA 分析产生的主子空间中,生成了主动原型模板和最优主动原型模板;将字符的特征向量  $x$  在其上的投影与最优主动原型模板的距离定义为非线性主动判别函数;进一步给出了基于最小分类错误准则的非线性主动判别函数的优化算法。实验结果表明,非线性主动判别函数的识别率高于线性主动判别函数,但是是以牺牲计算效率和存储空间为代价的,这是所有基于 Kernel PCA 的应用共同面临的问题;另外,出于计算效率的考虑,特征空间的特征向量  $V^{(i)}$  并未优化。这些都是以后的工作中值得进一步研究的问题。

## 参考文献 (References)

- Jain A K, Zongker D. Representation and recognition of handwritten digits using deformable templates [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(12): 1386 ~ 1391.
- Seiichi Uchida, Hiroaki Sakoe. Eigen-deformations for Elastic Matching based Handwritten Character Recognition [J]. Pattern Recognition, 2003, 36(9): 2031 ~ 2040.
- Shi Da-ming, Gunn S R, Dampier R I. A radical approach to handwritten chinese character recognition using active handwriting models [A]. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition [C]. Kauai, Hawaii, USA, 2001: 670 ~ 675.
- Shi Da-ming, Gunn S R, Dampier R I. Handwritten chinese radical recognition using nonlinear active shape models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(2): 277 ~ 280.
- Sun Guang-ling, Tang Xiang-long, Liu Jia-feng. Active discriminant functions for handwriting recognition [J]. Journal of Software, 2005, 16(4): 166 ~ 176.
- Scholkopf B, Smola A J, Muller K. Kernel principal component analysis [A]. In: Scholkopf B, Burges C J C, Smola A J (Eds.), Advances in Kernel Methods [C], Cambridge, MA, USA: MIT Press, 1998. 327 ~ 352.
- Sun Guang-ling, Tang Xiang-long. Fuzzy gradient feature extraction method based on gradient normalization [J]. Journal of Harbin Institute of Technology, 2006, 38(12): 2030 ~ 2035. [孙广玲, 唐降龙. 基于梯度归一化的模糊梯度特征提取方法 [J]. 哈尔滨工业大学学报, 2006, 38(12): 2030 ~ 2035.]