

• 通信与信息工程 •

基于分数本征特征的手写数字识别

孟庆宇, 刘本永, 姚宏达

(电子科技大学电子工程学院 成都 610054)

【摘要】特征提取是手写体数字识别研究中的重要问题,有效、稳定的特征是提高识别率和识别精度的关键。该文提出了一种基于分数本征特征和核非线性分类器的手写数字识别方法,首先找到时频平面的一个轴进行分数傅里叶变换,使不同类别样本在这个轴上最大限度地分开,然后用主元分析进行降维,得到比较稳健的低维特征,再将常用分类器用于特征分类,实现对手写数字的识别。对实际数据进行实验,结果表明上述本征特征与核非线性分类器相结合有较高的识别率和训练、分类效率。

关键词 分数傅里叶变换; 主元分析; 特征提取; 手写数字识别; 核非线性分类器
中图分类号 TN911.7; TP391 **文献标识码** A

Handwritten Numeral Recognition Based on Fractional Eigenfeatures

MENG Qing-yu, LIU Ben-yong, YAO Hong-da

(School of Electronic Engineering, Univ. of Elec. Sci. & Tech. of China Chengdu 610054)

Abstract Feature extraction is an important part in handwritten numeral recognition. Efficient and robust feature is a key to improving recognition rate and efficiency. This paper adopts fractional Fourier transform and principal component analysis to extract robust and compact features called fractional eigenfeatures. In classification, five kernel-based nonlinear classifiers, Parzen and robust Parzen classifiers, radial basis function classifier, support vector classifier, and kernel-based nonlinear representor are applied and compared. Experimental results show the effects and efficiency of the proposed algorithm.

Key words fractional Fourier transform; principal component analysis; feature extraction; handwriting numeral recognition; kernel-based nonlinear classifiers

手写数字识别是利用计算机对快速输入的手写数字进行自动识别分类,一直是图像处理、模式识别领域的研究热点,在邮政编码数字自动分拣信件系统、统计报表、财务报表、银行票据等方面得到广泛应用^[1-5]。在长期的工作中,研究人员提出了多种识别方法。按使用特征的不同,这些方法大体分为两大类:基于字符统计特征(密度、矩、特征区域)的方法和基于字符结构特征(圈、端点、交叉点、笔画、轮廓)的方法。由于采用了不同性质的模式特征,所以两者各有优势:字符统计特征相对容易获得,并且能很好地描述一类模式的本质特性,在样本统计特性一致性比较好时具有较高的识别率;基于字符结构特征的方法精确地描述了字符的细节特征,对书写结构较规范的字符有较高的识别率,但算法比较复杂。本文提出一种基于分数傅里叶变换(Fractional Fourier Transform, FrFT)和主元分析(Principal Component Analysis, PCA)的新的基于统计特征提取方法。

1 基于FrFT的特征提取

1.1 FrFT

FrFT是传统傅里叶变换在分数级次上的推广,与傅里叶变换紧密联系而又具有许多特殊性质,已广泛用于信号过滤、相位恢复、透镜设计等方面。

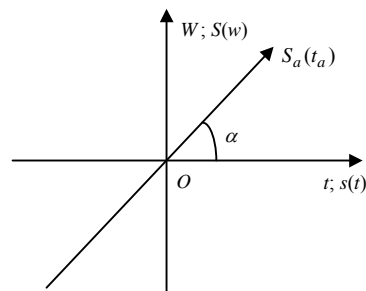


图1 分数傅里叶变换和傅里叶变换的关系

收稿日期: 2005-10-08

基金项目: 教育部科技研究重点项目(105150); ATR国防科技重点实验室项目

作者简介: 孟庆宇(1982-), 男, 硕士生, 主要从事信号处理与模式识别方面的研究。

在时频平面内傅里叶变换可看作信号表示轴在时频平面旋转($\pi/2$) rad, 如图1所示; a 阶FrFT则是在时频平面内旋转任意角度, 定义为:

$$S(t_a) = \int_{-\infty}^{+\infty} K_a(t_a, t) s(t) dt \quad (1)$$

$$K_a(t_a, t) = A_\phi \exp[j\pi(t_a^2 \cot \phi - 2t_a t \csc \phi + t^2 \cot \phi)]$$

$$\phi = a \frac{\pi}{2}, A_\phi = |\sin \phi|^{-2} \exp[-j(\frac{\pi}{4} \operatorname{sgn}(\phi) - \frac{\pi}{2})], K_{\pm 1}(t_a, t) = \delta(t_a - t), K_{\pm 2}(t_a, t) = \delta(t_a + t)$$

式中 $K_a(t_a, t)$ 为积分核并将其与积分算子离散化得到离散分数傅里叶变换(Discrete FrFT, DFrFT)^[6]: $s^a = F_a s$, 其中 s 是信号序列; F_a 是DFrFT的变换矩阵, 由变换系数 a 和序列点数 N 确定; s^a 是变换后的序列。

1.2 分数傅里叶变换在手写数字特征提取的应用

在对样本进行像素平均和按行拼接后进行DFrFT。设 s_i^j 是第 i 类第 j 个样本信号, $x_i^j = F_a s_i^j$ 是经DFrFT变换后的信号(N 维列向量)。进行DFrFT之前要先确定 a 值, 本文采用两种方法确定 a 的值: 首先是根据Fisher准则找到一个 a 值, 使变换后的结果具有最大的类间散度和最小的类内散度, 即使矩阵 $S = S_w^a S_b^a$ 的最大本征值最大, 其中 S_w 和 S_b 分别为类内散布矩阵和类间散布矩阵^[7]; 其次是找到一个 a 值, 使不同类别训练样本经过DFrFT后错误分类样本数最小(这里采用简单的欧式距离分类器)。

2 主元分析(PCA)

PCA是近年来应用较多的一种统计特性提取方法^[5, 8], 它充分利用数据中的二阶统计信息对数据进行特征提取和降维处理, 其主分量对应于数据均方误差重建误差曲面的最小点, 因此对数据信息具有较强的描述能力, 并且算法简单、运算量小。

PCA有多种不同的数值计算方法, 本文通过对训练样本协方差矩阵 C 进行本征分解得到正交变换矩阵 W 。对 M 个训练样本, 其协方差矩阵为 $C = \sum_{j=1}^M (s_i^j - \bar{s}_i)(s_i^j - \bar{s}_i)^T / M$ 。其中 $\bar{s}_i = \sum_{j=1}^M s_i^j / M$ 。设 C 的 N 个本征值按大小顺序排列为 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$, 对应的本征向量为 $u_1, u_2, u_3, \dots, u_N$, 取本征值较大的 p 个本征值对应的特征向量组成特征空间的一组正交基, 其中 p 由压缩比确定^[8]。将所有样本投影到这个特征空间, 得到坐标 $y_i^j = [u_1^T s_i^j \ u_2^T s_i^j \ \dots \ u_p^T s_i^j]^T$ (p 维列向量)。

3 分类器

分类器设计是模式识别中的一个重要问题。一般线性分类器(如欧式距离分类器)只适用于不同类别在特征空间线性可分的情况, 对于有些线性不可分问题要采用非线性分类器(如Parzen分类器^[9])。本文应用常用分类器并比较其分类效果和效率, 包括核非线性表示子(Kernel-based Nonlinear Representor, KNR)^[10], 非线性支持向量分类器(Support Vector Classifier, SVC)^[11], Parzen分类器, 稳健Parzen分类器, 径向基函数神经网络(Radial Basis Function Neural Network, RBF-NN), 以及欧式距离、最近邻、 k -近邻等分类器。其中前五种非线性分类器的核函数均采用高斯核 $k(x, y) = \exp(-\|x - y\|^2 / \sigma^2)$, 参数 σ 的值由文献[9]中的方法取得。由于SVC分界面针对的是两类问题, 解决多类问题时可以用一对一和一对多两种方法, 本文采用前者——每两个类别建立一个分类器, 然后用举手表决的方法进行判决。

4 实验结果及分析

实验数据是荷兰邮政系统提取的手写体字符(“0”~“9”), 每一字符有200个(共2 000个)样本。每个样本是分辨率为30×48的二值图像, 预处理中先取2×3无交叠窗内像素平均, 如图2所示, 再按行拼接得到240维的矢量。在预处理之后, 取每类前10个(共100个)样本, 按上述的两种方法确定 a 值, 得到的结果如图3所示。可以看出, 两种方法所确定的 a 值都是0.05。以下实验中, 固定每类后100个(共1 000个)样本作为测试集, 在每类前100个样本中任取50个作为训练集进行实验。实验中对以下三种特征提取方法的效果: 以0.05为变换系数进行DFrFT, 结果取模值后进行PCA(即DFrFT+PCA); 直接PCA; 先进行离散傅里叶变换(Discrete Fourier Transform, DFT), 结果取模值后进行PCA(即DFT+PCA)。上述PCA中信息压缩比取值为0.95。同时, 对比以下分类器的效果和效率: 基于高斯核的KNR、基于高斯核的非线性SVC、基于高斯核的Parzen和稳健

Parzen(Robust Parzen)、RBF网络、欧式距离(Euclidean)、最近邻、 k -近邻(k -NN)等8种分类器。

对以上特征提取和分类方法,每次用相同训练集和测试集进行实验,共10次无重复实验。对1 000个测试样本和10次实验的平均识别率如表1所示,而各分类器所需要训练时间和测试时间如表2所示。训练时间指训练分类器参数所需要的时间,测试时间是分类器对1 000个测试样本分类时所需时间。



图2 预处理之后的结果

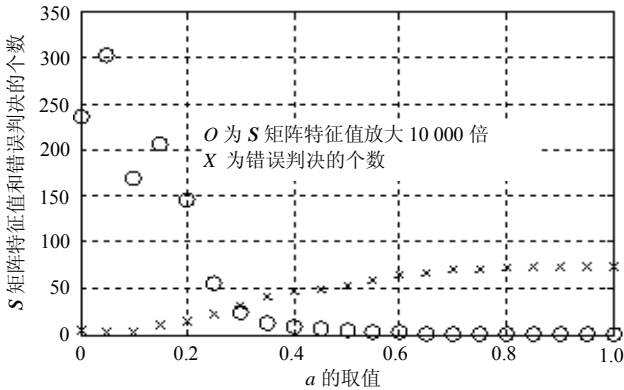


图3 交叉验证的结果

表1 各种特征提取方法和分类器的平均识别结果

	KNR	SVC	Parzen	Robust Parzen	RBF	Euclidean	NN	k -NN
DfFrFT+PCA/(%)	96.95	97.26	96.01	96.73	93.68	93.68	96.20	95.68
直接PCA/(%)	96.54	97.06	94.63	94.63	96.46	90.84	96.33	94.78
DFT+PCA/(%)	79.28	82.70	74.89	75.17	47.90	71.97	78.69	78.25

表2 各种分类器的效率

	KNR	SVC	Parzen	Robust Parzen	RBF	Euclidean	NN	k -NN
训练时间 /s	0.312 0	14.760 0	3.078 0	6.516 0	7.109 0	0.015 0	0	0
测试时间 /s	6.766 0	99.460 0	7.625 0	6.828 0	11.563 0	0.156 0	19.297 0	20.297 0

从表1和表2可以看出:用DfFrFT和PCA提取特征,可以得到较高的识别率,识别结果也比较稳定;同时,除了RBF之外,其他分类器识别率都有提高,而KNR分类器在识别率、效率方面综合性能最高。

5 结 论

利用分数傅里叶变换和主元分析对手写数字进行特征提取,设计了一种新的手写数字的特征提取的方法,与基于傅里叶变换和基于主元分析的特征提取方法比较,实验表明了这种特征提取方法可以得到维数较低而且稳健、有效的特征。同时,用多种分类器进行分类的实验结果表明,与常用分类器如Parzen、稳健Parzen、欧式距离、最近邻、 k -近邻等相比,SVC、RBF、KNR的分类效果、效率等综合性能较好。

参 考 文 献

[1] Hu j, Yan h. Structural primitive extraction and coding for handwritten numeral recognition[J]. Pattern Recognition, 1998, 31(5): 493-509.

[2] 吕 岳, 芮剑明, 余德华. 邮政编码数字识别系统的设计与实现[J]. 计算机工程与应用, 1999, 35(6): 128-129.

[3] 张保轩, 房世辉. 基于外形的手写字体数字识别[J]. 山东通信技术, 1995, 15(1): 28-31.

[4] 朱曼丽, 孙志锋, 孙志林. 一种基于神经网络的数字识别方法[J]. 计算机工程与应用, 2002, 38(12): 75-78.

[5] 芮 挺, 沈春林. 基于主分量分析的手写数字字符识别[J]. 小型微型计算机系统, 2005, 26(2): 289-292.

[6] Candan G, Kutav M A. The discrete fractional Fourier transform[J]. IEEE. Trans. on Signal Processing, 2000, 48(5): 1 329-1 337.

[7] 刘本永. 子空间法雷达目标一维像识别研究[J]. 电子与信息学报, 2004, 26(7): 1 137-1 143.

[8] 孙 鑫, 刘本永. 基于局部本征谱的人脸识别[J]. 电子科技大学学报, 2005, 34(4): 493-496.

[9] Kraaijveld Martin A. A parzen classifier with an improved robustness against deviations between training and test data[J]. Patten Recognition Letter, 1996, 17(7): 679-689.

[10] Zhang J, Liu B Y, Tan H. A kernel-based nonlinear represantor for eigenface classification[J]. J. Electronic Science and Technology of China, 2004, 2(2): 19-22.

[11] 张 莉. 支撑矢量机与核方法研究[D]. 中国优秀硕博学位论文全文数据库, <http://202.115.24.13/kns50/index.aspx>, 2002, 9: 20-36.

编 辑 徐安玉