

特征融合用于手写体汉字识别研究

居 琰* 汪同庆 彭 建 王贵新 刘建胜 袁祥辉

(重庆大学光电工程学院 重庆 400044)

【摘要】分析了手写汉字特征的提取方法，提取具有一定互补性的轮廓方向特征和方向距离分布特征，并进行K-L变换降维处理，用多特征合成一个区分能力更强的新特征。讨论了RBF网络分类器特性，结合特征融合方法和模块RBF神经网络结构有机地构建一个小类别手写体汉字识别系统。实验表明，该系统可行和有效。

关 键 词 特征融合； K-L变换； 模块RBF网络； 手写体汉字识别

中图分类号 TP391.4

Research on Handwritten Chinese Character Recognition Using Feature Fusion and Modular RBF Classifier

Ju Yan Wang Tongqing Peng Jian Wang Guixin Liu Jiansheng Yuan Xianghui

(Institute of Optic-Electronic Engineering, Chongqing University Chongqing 400044)

Abstract This paper analysis the feature extraction method of handwritten Chinese character. The contour direction feature(CDF)and directional distance distributions feature(DDDF) are extracted from Chinese character image as a pair of feature vectors having good complementarity. After dimensions reduction of original feature using Karhunen-Loeve transform, these two feature vectors are combined to produce a new feature vector has high discriminating powers. Basing on the characteristic of RBFNN classifier, a novel architecture which integrates feature fusion and modular RBF neural networks classifier approaches into a small set handwritten Chinese character recognition system is presented. Experiments show that the system has achieved impressive performance and the results are informing.

Key words feature fusion; karhunen-loeve transform; modular radial basis function neural network; handwritten chinese character recognition

汉字特征直接反映了汉字形体整体或局部分布状况，对大多数结构稳定的汉字，只要获得了正确优良的特征集，就可以按照一定的策略和步骤，将待识别汉字正确分类，汉字识别的正确性就能够得到保证。但是手写体汉字书写因人而异，笔画变形大，使属于同一类别样本间的差异较大，而且汉字类别多，相似字多，难以获得稳定的特征进行有效识别分类。

本文根据手写体汉字识别的特点，对手写体汉字特征的获取方法进行了改进，提取具有一定互补性的不同汉字特征，结合待识别汉字的自身特性，进行特征融合，使融合后获取的新特征适应不同汉字类别及不同样本间的书写变化。分类器设计则采用多模块RBF神经网络，其结构简单、算法简便，两者结合应用于小类别手写体汉字识别，取得了令人满意的效果。

2001年7月30日收稿

* 男 29岁 博士生

1 汉字特征提取融合

汉字识别过程中,特征提取是关键的部分,很多学者进行了大量的研究工作,提出了许多不同的特征提取方法^[1,2]。但是,利用这些方法所构建的汉字识别系统识别率均不高,难以达到实用要求。由此可见,由单一特征、单一方案构成的识别系统,其识别率有一定的限度。

手写汉字识别器主要由特征提取器及分类器组成。识别器集成和组合包含以下两方面:1)融合不同分类器分类结果;2)由不同的原始特征向量合成新的特征向量。多分类器结果融合适用于对不同特性的分类器进行决策互补,而新特征向量合成有利于增强特征向量的分类能力。

Favata等用表征字符局部及全局的三种不同特征合成新的特征向量进行手写字符识别^[3],实验证明合成后的特征向量提高了字符的识别率。Mastui等对利用字符笔画背景结构以及轮廓方向特征合成新特征手写数字识别进行了研究^[4]。这些研究工作均证明了特征融合技术可以有效提高字符的识别率。上述研究工作的特征融合对象均为手写数字,合成的规则是简单的叠加。手写汉字与手写数字相比,结构更复杂,字体变形更大,为获得高的识别率,提取汉字特征维数远大于提取数字特征维数。对汉字特征融合进行研究时,简单叠加规则将导致合成后新特征维数加倍放大,并存在大量冗余特征,导致分类性能下降。同时向量维数的增长也使训练更复杂,计算量加大。

在特征空间中,如果次要的分量可忽略不计,那么就达到了所需的降维目的。另外,从统计意义上分析,次要分量有可能来自噪声或者字符样本固有特性以外的外界干扰,因此,消除这些分量是有益处的。主分量分析(PCA)也称为K-L变换(Karhunen-Loeve),它通过引入一种可检测因忽略变换向量的某些分量所造成误差的度量,可知被忽略分量的数目及特性,并能以最小丢失率来降低随机输入向量的维数。

根据K-L变换理论,特征值越大,样本在其所对应的特征向量上的投影就越真实地反映了该样本的主要成分,据此可以实现对特征的降维。

系统中采用了两种特征:轮廓方向特征(contour direction feature, CDF)和方向距离分布特征:(directional distance distributions feature, DDDF)^[5,6]。轮廓方向特征是一种手写体汉字识别特征,主要有横、竖、撇、捺四种基本笔画,其他笔画可看成是由这四种基本笔画复合而成。尽管手写汉字存在种种变形,但是笔画间的相对位置基本保持不变,这是手写体汉字最为稳定的特征。由于不同手写汉字样本笔画宽度变化很大,所以笔画分解时只考虑手写汉字的轮廓。特征提取前汉字图像已归一化为64×64点阵。首先对汉字图像进行8×8非均匀网格划分,非均匀网线根据汉字图像在水平、垂直两个方向上的直方图投影来确定。非均匀网线两两相交即构成了弹性网格,划分后获得的弹性网格数为64,如图1所示。再提取汉字轮廓,设 $a(x,y)$ 表示轮廓点 (x,y) 对应于 x 轴的笔画方向角度, (x,y) 的8邻域定义如图2所示。则有

$$a(x,y)=\tan^{-1}G_x/G_y$$

式中

$$G_x=(z_6+2z_7+z_8)-(z_1+2z_2+z_3)$$

$$G_y=(z_3+2z_5+z_8)-(z_1+2z_4+z_6)$$

笔画方向角度变化范围为0~180°,将此变化范围划分为4个分区,即 G_1 、 G_2 、 G_3 、 G_4 ,分别对应于汉字的横、撇、竖、捺笔画,如图3所示。依据笔画方向角度 $a(x,y)$,将汉字轮廓点 (x,y) 归入4个分区。分别计算64个弹性网格内归属横、撇、竖、捺笔画汉字轮廓点 (x,y) 的数目,每个弹性网格可得到一个四维向量,最终可以得到256(4×64)维数的特征向量。

方向距离分布特征(directional distance distributions feature, DDDF)已被成功运用于手写数字识别,并获得了平均97.3%的识别率^[6]。DDDF有两种基本计算方式,即通过计算每一个黑像素点和白

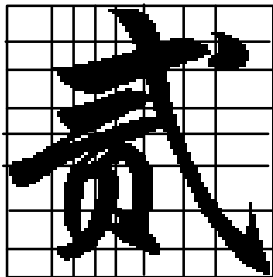


图1 8×8不均匀网格划分

Z_1	Z_2	Z_3
Z_4	(x, y)	Z_5
Z_6	Z_7	Z_8

图2 点 (x, y) 8邻域定义

像素点的8方向或者4方向距离来获取字符的整体信息。为了有效地区分出各类不同的手写汉字,系统采用的是计算8方向距离。将识别汉字均匀划分为36(6×6)网格,计算每一网格内每一像素点的8方向距离分布,求取网格内所有像素点平均值,这样每一网格可得到一个16维向量,最终构成一个576(16×36)维特征向量。

轮廓方向特征反映了手写汉字的局部信息,而方向距离分布特征反映了字符的全局形状信息,提供了整体字符的大致轮廓。因此这两种特征有较好的互补特性,基于这两种原始特征合成的新特征向量将具有更强的区分能力。

如前所述,若特征融合采用的规则是简单的叠加,则新特征的维数将达到832维。过高的特征维数将给后续的分类器研究带来很多不便,并且冗余特征的存在也将降低分类器的分类性能。

采用K-L变换对原始特征进行降维处理后,可将轮廓方向特征从原来的256维特征空间变换到128维特征空间,方向距离分布特征从原来的576维特征空间变换到144维特征空间,进行叠加后新特征的维数也仅为272维。由于K-L变换后所获得的特征维数比变换前要小得多,所以运算速度也快得多。

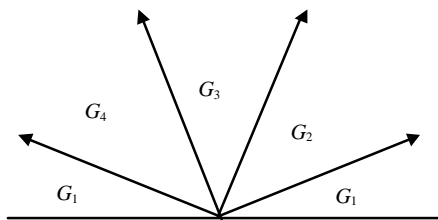


图3 笔画方向角四区间归属

2 模块RBF网络分类器设计

目前在字符识别等典型模式分类问题中,BP以外的模型很少有报道。BP模型虽然简捷,但当样本特征维数增大时,BP模型在判别相似类别的差异时会遇到困难,造成误分而影响精度。

径向基函数神经网络理论(Radial Basis Function Neural Network ,RBF)为多层前馈网络的学习提供了一种新颖而有效的方法。RBF网络不仅具有良好的推广能力,而且计算量少,学习速度也比其他算法快得多。RBFNN的网络结构为3层,中间层表示特征空间按照一定密度分布的中心点,样本输入分别对应各中心点,通过一个非线性的径向基函数得到中间层的竞争输出,中间层的状态包括中间层结点的个数及其状态值,它反映了样本空间中数据的离散分布状况。输出层输出则是对中间层RBF基函数输出的线性映射。应用RBF网络,使变换在隐含层进行,这就允许隐含单元空间具有更高的维数。对于一组输入量 x_j 和权值 w_{ij} 径向基函数神经元的输出为

$$out = g \left[\sum_j^n (x_j - w_{ij})^2 \right]$$

式中 $g()$ 是高斯函数

$$g(r^2) = ce^{-r^2/s^2}$$

虽然RBF网络比通常的感知器(MLP)有更强的局部聚类 and 分类能力,但是应用于手写汉字识别

系统研究中,如用同一功能类型的单个神经网络来完成,必然导致网络结构复杂,造成训练时运算量大而难以保证学习效果。为此,系统采用的是模块化模型。该模型的构思是将复杂的分类分解为若干个子任务,通过综合各子任务的完成结果来实现整个分类过程。这种方法是将一个 K 类问题分解

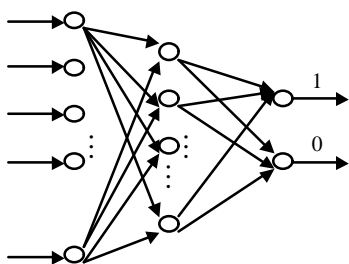


图4 模块RBF子网络

为 K 个两类问题,其模块网络就是 K 个解决两类问题模块的集合。每个模块是一个具有两结点输出的RBF网络,它用来将一类模型和其他类的模型区分开来,即 A 类模块被训练成能够区分 A 类和非 A 类模型。当输入样本是 A 类时,模块的输出值为高,而当输入样本是非 A 类时,模型的输出值为低。图4为模块网络的示意图。在模块网络中,模块可以并行独立地训练。同样,当分类时这些模块也可以并行处理。由于子网络的针对性较强,易于提取细节信息,从而改善分类精度。

RBF网络的训练包括确定函数的中心、宽度、联结隐含层和输出层神经元的权值,常用的算法有聚类算法等。使用模块网络时,每个子网络只解决二类分类问题,因此使用一种学习效率更快且针对二类问题进行优化的网络学习算法是十分必要的。系统采用的是Musavi等人提出的迭代聚类算法^[7],它能使RBF网络隐含层结点的数目最小。

3 实验结果及结论

在小类别数汉字识别系统中,需识别的手写汉字类别数较少,如银行票据(支票、账单等)的大写金额自动识别,待识别的通常是为数不多的几种手写汉字,完整的系统识别流程如图5所示。不失一般性,为有效检验特征融合和模块RBF网络分类器的特性,实验分别构建了三个面向小类别手写体汉字识别系统。在实验中,需要识别的汉字包括零、壹、贰、叁、肆、伍、陆、柒、捌、玖、拾、亿、佰、仟、万、圆、元、角、分、整、正等共21个。第一个实验系统采用特征为256维轮廓方向特征,模块神经网络采用如前所述的21组子RBF网络分类器,将其标记为 $C1$;第二个实验系统采用文中所述的272维融合特征,模块神经网络同样为21组子RBF网络分类器,将其标记为 $C2$;第三个实验系统同样采用模块神经网络设计,与 $C1$ 、 $C2$ 不同的是子网络分类器不采用RBF网络分类器,而采用常见的BP网络分类器,训练算法采用递推最小二乘法(RLS)算法,特征为256维轮廓方向特征,标记为 $C3$ 。三个模块网络的子网络分类器分别用来识别上述汉字的一种,模块网络结构各模块间既相互独立、相互补充,又相互排斥,各自完成识别不同的汉字。

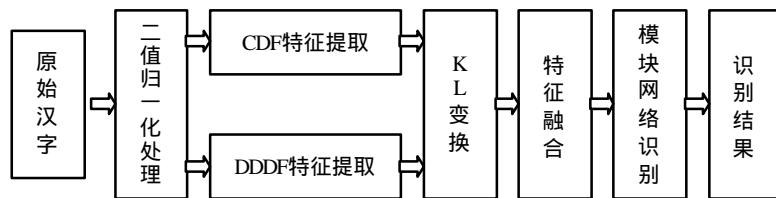


图5 系统识别流程

21种需识别汉字各收集了1 000个不同的书写样本,共有样本数目21 000。收集的每种汉字以600个字作为训练样本,以400个字构成测试样本,实验结果如表1所示。实验表明,系统 $C2$ 的识别率高于 $C1$ 的识别率,说明特征融合技术可以有效提高手写汉字的识别率; $C1$ 的识别率高于 $C3$,则验证了模块RBF网络分类器的良好分类特性。

表1 实验结果

识别系统	训练样本[识别率/(%)]	测试样本[识别率/(%)]
C1	98.11	96.43
C2	98.94	97.75
C3	97.07	95.77

参 考 文 献

1 Wang Guixin, Liu Jiansheng, Ju Yan, *et al.* Handwritten numerals recognition based on “effective-rows” feature. Journal of University of Electronic Science and Technology of China, 2001, 30(3): 287-291[王贵新, 刘建胜, 居 琰, 等. “有效行”特征对手写体字符的识别. 电子科技大学学报, 2001, 30(3): 287-291]

2 Cui Huailin, Zhao Shuxiang. Research on preliminary classification of handprinted Chinese character ecognition. Journal of University of Electronic Science and Technology of China, 1996, 25(3): 311-315[崔怀林, 赵树芎. 手写汉字识别粗分类方法的研究. 电子科技大学学报, 1996, 25(3): 311-315]

3 Favata J T, Srikantan G, Srihari S N. Handprinted character/digit recognition using a multiple feature/resolution philosophy. Proc. IWPFR’ 94, 1994: 57-66

4 Matsuii T, Tsutsunida T, Srihari S N. Combination of stroke/background structure and contour-direction features in handprinted alphanumeric recognition. Proc. of IWFHR’ 94, 1994: 87-96

5 Tseng Yihong, Kuo Chichang, Lee Hsijian. Speeding up Chinese character recognition in an automatic document reading system. Pattern Recognition, 1998, 31(11): 1 601-1 612

6 Oh II Seok, Suen Ching Y. Feature for character recognition based on directional distance distributions. Proceedings of the International Conference on Document Analysis and Recognition, ICDAR Aug 18-20 1997: 288-292

7 Musavi M L, Faris K B, Chan K H, *et al.* On the implementation of RBF technique in neural networks. ACMM ANNA-91. Anal. Neural Network Appl., 1991: 110-115

· 科研成果介绍 ·

X波段LFMCW高线性度信号源

主研人员：汪学刚 张永强 向敬成 宋旭东 黎安尧 黄 河 石 凯

X波段LFMCW高线性度信号源采用了数字式实时闭环线性度校正技术,具有校正后线性度高、适应于不同VCO特性,可同时进行线性度校正与测量等一系列优点。信号源中使用了体声波延时线、大规模可编程等新器件,具有体积小、成本低、应用灵活等特点。该信号源满足了技术指标的要求,提高了距离分辨力,改善了测距精度,使用效果良好。

· 甬 江 ·