

基于小波的车牌汉字特征提取

潘翔 叶修梓 张三元

(浙江大学 CAD&CG 国家重点实验室, 杭州 310027)

摘要 车牌识别系统是智能交通不可或缺的一部分。在车牌识别系统中, 车牌首位汉字的特征提取和识别是一个难点。为此, 将小波的多分辨率特性应用于车牌汉字特征的提取, 提出了一种直接从灰度图象提取车牌汉字特征的提取方法。该方法首先提取图象的小波矩和基于小波分解的区域密度特征, 然后以识别率好坏为最优依据, 进行特征分量排序和选择, 最后把特征矢量送入 BP 神经网络进行字符识别。该方法避免了一些传统汉字特征提取方法需要对图象进行二值化操作而造成的汉字字符结构信息丢失。提取的特征有效地反映了车牌汉字的局部和全局特征。实验结果表明该特征提取方法可以得到比较好的识别效果。

关键词 模式识别(520·2040) 车牌汉字 矩 小波变换 特征提取

中图法分类号: TP391 **文献标识码:** A **文章编号:** 1006-8961(2003)10-1218-05

The Car Plate Chinese Character Feature Extraction Based on Wavelet

PAN Xiang, YE Xiu-zi, ZHANG San-yuan

(State Key Laboratory of CAD&CG Zhejiang University, Hangzhou 310027)

Abstract The car plate recognition system is an indispensable part of intelligent traffic system. In car plate recognition system, one of the most difficult problems is the feature and recognition of car plate Chinese character. The multiresolution property of wavelet is applied to extract car plate Chinese character feature extraction, and a method that directly extracts feature from gray scale image is proposed. The method firstly extracts two kinds of features: one is wavelet moments in different multiresolution, and the other is the zoning density of wavelet-decomposed image. Then, two kinds of features are combined and a feature selection algorithm based recognition rate is presented to choose better features. Finally, the selected features are as the input of BP neural network, which is adopted to recognize car plate Chinese character. As a result, the proposed method avoids the binary operation used in some traditional Chinese character feature extractions that will seriously destroy the Chinese character structure. Furthermore, the extracted features can describe the local and global property of the character. Compared with some other feature extraction methods, the proposed method can achieve better recognition performance.

Keywords Pattern recognition, Car plate chinese character, Moment, Wavelet transform, Feature extraction

0 引言

近年来, 图象处理、通信技术和计算机技术在交通领域得到了广泛的应用, 其中车牌识别技术是研究的热点之一。车牌识别技术被广泛应用于过路费系统和交通管理系统, 具有巨大的经济价值和现

实意义。一个完整的车牌自动识别系统包括图象采集、车牌定位、字符切分以及字符识别4个部分。国内外对这方面的研究虽然取得了一定的进展, 但是还存在一些问题, 特别是对于国内车牌的识别。和国外汽车牌照识别系统比较, 国内车牌识别系统有其特殊性。首先国内车牌不仅有英文字母和阿拉伯数字, 而且还有汉字; 其次, 由于国内部分道路条件较

基金项目: 国家自然科学基金(60073026); 国家高技术研究发展(863)计划重点项目(2002AA411010)

收稿日期: 2002-10-16; **改回日期:** 2003-06-09

差,车牌污染比较严重,从而使部分车牌图片质量很差,其中车牌中的汉字更难识别。因此如何提取车牌汉字特征,成了车牌识别的难点问题,也限制了车牌识别系统的进一步应用。

对于车牌汉字的特征提取,传统的方法是先对图象作二值化,然后再提取特征。但是对于车牌汉字,由于其笔划多、分辨率低和车牌受污染严重,并且图象强度区域分布不均匀,若进行二值化处理,会造成大量字符特征信息的丢失。为了避免这种情况,可以采用自适应阈值和局部二值化等方法来提高二值化效果,但是,无论采用哪一种二值化方法,总是会造成汉字笔划信息的丢失,从而使字符结构特征不明显^[1]。

小波变换提供了一种字符图象的多分辨率分析,它不同于传统的短时傅立叶变换,它可以分析图象在时域和频域的局部特性,因此小波变换特别适合于提取图象局部的细节特征。小波变换在模式识别中已经得到了成功的应用^[2]。

将小波理论应用于车牌汉字特征提取,提出了一种直接从灰度图象提取车牌汉字特征的方法。该方法避免了由于二值化过程而造成的字符信息丢失,提取的特征能够有效地反映车牌汉字的局部和全局特征。

1 车牌汉字图象预处理

在进行特征提取以前,要对图象进行预处理,因为不同的灰度图象,其大小不一致,灰度有很大的差别。图象预处理可以使得在不同环境下得到的字符灰度比较统一和大小一样,从而可以提高特征提取的质量。在这里采用反差增强算法进行字符灰度调整,即采用直方图统计其灰度分布,然后自适应得到门限值 T_1 和 T_2 ,对于不同的灰度范围采取不同的拉伸

$$f(i, j) = \begin{cases} 0 & x \leq T_1 \\ \frac{255}{T_2 - T_1}(x - T_1) & T_1 < x < T_2 \\ 255 & x \geq T_2 \end{cases} \quad (1)$$

其中, x 为图象点 (i, j) 的灰度值, $f(i, j)$ 为拉伸后得到的灰度值。对于大小归一化,为了尽可能防止图象失真,采用双线性插值方法把图象归一到 32×16 大小。图1是“浙”和“闽”字的预处理效果图,和原图比较,处理后的图片字符和背景的对比更加明显,不同字符归一化后的灰度比较统一。



图1 预处理效果图

2 特征提取

在经过图象预处理后,一个关键的问题是如何提取图象的特征,这是得到高性能识别系统的关键。对于提取出来的特征,要保证类内聚类性和类间可分性,同时避免维数冗余性。对于汉字,其特征主要是横、竖、撇和捺4种笔划。因此,对于车牌汉字特征的提取,其关键是提取出来的特征是否能够有效地反映出汉字笔划的这些特性。横笔划,在灰度图象中表现出来的是灰度变化沿水平方向为低频,沿垂直方向为高频;竖笔划刚好相反。小波的多分辨率特性为这些特征提取提供了很好的理论基础^[3]。本文从小波理论出发,提取如下两部分特征:

(1) 不同层次的小波矩 小波矩能够反映出图象的局部细节特征。

(2) 基于小波分解的区域密度 该特征能够反映图象全局的特征,并且对一般噪声不敏感。

2.1 小波矩

在模式识别中,矩特征是一种比较有效的特征描述方法,其具有旋转、平移、大小不变性,但是,计算量大和高阶矩受噪声影响较大一直困扰着矩特征的应用。小波矩揭示了图象在多分辨率分解中不变性特征的本质,解决了小波不变量及矩的多尺度表示问题^[4]。

设 $f(x, y)$ 代表灰度图象,其对应的极坐标式为 $f(r, \theta)$,则对于一个矩不变量,可以采用下式表示

$$F_{pq} = \iint f(r, \theta) g_p(r) e^{jq\theta} dr d\theta \quad (2)$$

其中, F_{pq} 表示第 pq 阶矩, $g_p(r)$ 是一个 r 的函数,用于控制不同的矩表示量,可以证明 F_{pq} 的模具有旋转不变性,令

$$S_q(r) = \int f(r, \theta) e^{jq\theta} d\theta \quad (3)$$

则有

$$F_{pq} = \int S_q(r) \cdot g_p(r) r dr \quad (4)$$

令 $g_p(r)$ 为小波基函数^[4]

$$\phi_{a,b}(r) = \frac{1}{\sqrt{a}} \psi\left(\frac{r-b}{a}\right) \quad (5)$$

则可以得到小波矩,为了简化计算,在这里,母小波

采用三次B样条的高斯逼近形式^[5]

$$\psi(r) = \frac{4\beta^{n+1}}{\sqrt{2\pi(n+1)}} \sigma_w \cos(2\pi f_0(2r-1)) \times \exp\left\{-\frac{(2r-1)^2}{2\sigma_w^2(n+1)}\right\} \quad (6)$$

其中, n 为 B 样条的阶次, f_0 为调制系数, σ_w^2 为离差量, 若使 $n=3$, $\beta=0.697\ 066$, $f_0=0.409\ 177$, $\sigma_w^2=0.561\ 145$, 则该函数可以达到很好的逼近效果, 其逼近误差小于 3%。

a 和 b 的取值如下

$$\begin{cases} a = 0.5^m & m = 0, 1, 2, 3 \\ b = 0.5n \cdot 0.5^m & n = 0, 1, \dots, 2^{m-1} \end{cases} \quad (7)$$

把母小波写成一般的形式

$$\psi_{m,n}(r) = 2^{m/2} \psi(2^m r - 0.5n) \quad (8)$$

则小波矩不变量如下计算

$$\|F_{m,n,q}\| = \left\| \int S_q(r) \cdot \psi_{m,n}(r) r dr \right\| \quad (9)$$

其中, $m=0, 1, 2, 3$, $n=0, 1, \dots, 2^{m-1}$, $q=0, 1, 2, 3$ 。

可以得到 136 维小波矩特征。

2.2 区域密度

在 OCR(光学字符识别)系统中, 区域密度是一种比较常用的全局特征, 其优点是对一般噪声不敏感。在这里, 首先利用塔式分解算法对汉字图象作一次小波分解(图 2)。

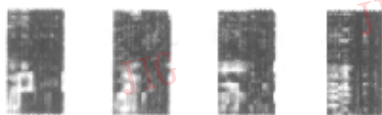


图 2 不同汉字图象的…次小波分解

对分解后的图象, 按照区域大小 3×3 对每个区域进行密度统计, 可以得到 66 维特征(因图象大小是 32×16 , 故部分区域的大小不是 3×3)。

3 特征选择

对于提取出来的特征, 要进行特征选择。特征选择是指在学习训练过程中, 对特征分量进行测试。从而在识别过程中, 只提取那些最能反映类别属性的特征分量来作为特征输入神经网络进行识别。特别是在实时性非常强的车牌识别系统中, 减少要提取的特征维数, 可以减少系统运行的时间, 提高系统运行的效率。因此, 对于提取出来的 136 维小波矩和 66 维区域密度特征需要进行特征选择。

特征分量选择, 就是判断每一维特征分量的优劣。一个识别系统, 判断特征分量好坏的最终决定因素是该维分量对总体识别性能的影响有多大。如果采用该维分量能够得到更好的识别率, 则认为该维分量优于其他维特征分量。

基于该思想, 提出一种简单有效的特征选择方法。该方法以识别率为最优性能评价标准, 对于不同的特征维数输入, 统计其对于训练集合的识别率。根据识别率的好坏来确定系统所需的输入特征。具体步骤如下:

(1) 均值向量计算 对训练集合, 计算每一类的均值向量

$$C_i^*, i = 1, 2, \dots, N, k = 1, 2, \dots, K$$

其中, K 为类别总数, N 为特征总维数, 在这里 $K=24$, $N=202$ 。

(2) 特征分量排序 对特征向量的每一维分量

$$F_i, i = 1, 2, \dots, N$$

单独作为特征输入, 对应的均值向量的该维分量

$$C_i^*, k = 1, 2, \dots, K$$

作为聚类中心。采用模版匹配方法, 统计其对于训练集合的识别率, 按照识别率从高到低对特征分量进行排序, 得到特征分量排序结果

$$\{F_1, F_2, \dots, F_N\}$$

(3) 特征维数选择 在经过特征分量排序后, 按照特征维数递增的方法(每次特征维数加 1), 选择不同维数的特征作为输入

$$\{F_1, F_2, \dots, F_i\}$$

其中, $i=1, 2, \dots, N$

计算其相应的识别率

$$Rate_i, i = 1, 2, \dots, N$$

图 3 给出了随着特征维数的增加, 识别率的变化情况。

从图中可以看出, 随着特征维数的增加, 识别率并

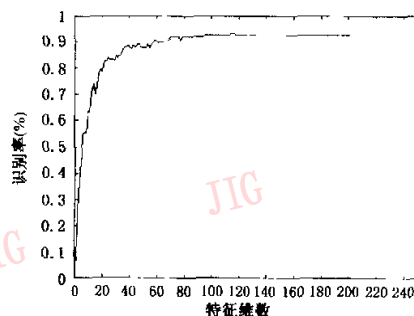


图 3 特征维数和识别率

不是单调增加,而是当维数增加到某一维时,识别率会慢慢趋向饱和.在这里,选择前面 81 维特征;其中小波矩的特征个数是 27,区域密度的特征个数是 54.

4 BP 神经网络分类器

将选择得到的特征送入 BP 神经网络进行字符识别.采用的网络结构为 3 层神经网络结构:其输入节点数为提取的特征维数,输出节点为 24 个(主要是各种常用省份的简称),两个隐层的节点数都为 50 个.训练误差为 0.01,学习最大次数为 10 000 次.

BP 神经网络训练算法如下:每次产生一个 1 到 60 之间的随机数和一个汉字索引的随机数,组合而成得到一个相应的汉字图象样本的文件索引,提取该样本的特征输入到神经网络中训练,直到该样本的训练误差小于给定误差,则认为完成一次学习;然后转入下一个样本学习.若单个样本的学习次数超过 300 次后(忽略网络前 100 个样本的学习),其误差若还大于训练误差,则认为网络不能收敛,中止学习.

5 实验结果

为了验证该方法的有效性,把本文提出的方法和以下两种特征提取方法作了比较:一种是对灰度图象进行 Kirsch 边缘检测,然后以区域大小 4×4 划分得到边缘图象,对每个区域,统计 4 个方向的边缘强度,得到 128 维特征;另一种是方向像素密度特征,首先对图象进行二值处理,然后以区域大小 4×4 进行划分,再计算每个区域 4 个方向的像素密度,也得到 128 维特征.实验数据库是从不同地方不同地点在不同环境下采集过来的车牌中分割出来的汉字,字符总数为 2 880,有 24 类汉字,每个汉字的个数为 120 个.图 4 是经过预处理后的部分汉字(为了增强对比度,进行了反色处理).



图 4 部分汉字样本

对于方向像素密度特征提取,需要对汉字进行二值化处理.采用的二值化算法是基于 LOG 算子的自

适应二值化方法,得到的二值化效果如图 5 所示.



图 5 灰度汉字和相应的二值化结果

从图 5 可以发现,对于笔划比较少,图象质量比较好并且结构简单的汉字,二值化效果还可以,但是对于结构复杂的汉字,或是图片不清晰,二值化以后,图象失真严重.

样本库中的字符被分成两组,每组 1 440 个.一组用于 BP 神经网络的学习训练;另外一组用于测试识别性能.

对于不同的特征输入,BP 神经网络的学习曲线如图 6 所示.

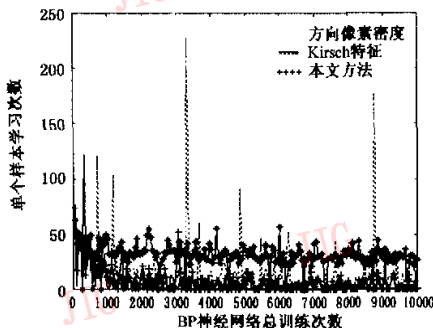


图 6 BP 神经网络学习曲线

从 BP 神经网络的收敛曲线可以看出:和其他两种特征提取方法比较,本文方法提取的特征,随着训练次数的增加,单个样本学习的收敛速度更快.

表 1 给出了不同方法提取出来的特征送入神经网络的识别性能.

表 1 不同特征的识别率(不考虑拒取率)

	单位: %		
	本文算法	Kirsch	方向像素密度
训练集合	95.76	84.6	88.5
测试集合	92.86	80.3	85.8

若神经网络输出取阈值 0.9,对于训练集合和测试集合中的汉字,得到如表 2 所示的识别性能.

表 2 不同特征的识别率(考虑拒取率)

	单位: %		
	识别率	拒取率	误识率
本文算法	90.23	8.12	1.65
Kirsch	78.54	6.4	15.06
方向像素密度	83.03	6.3	10.67

图7是用不同方法对不同汉字的识别结果,其中横坐标索引对应的汉字原型如表3所示。

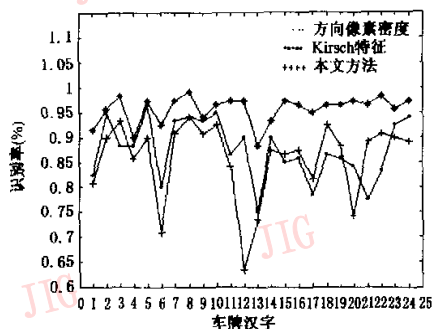


图7 不同车牌汉字的识别率(不考虑拒取率)

表3 汉字原型及对应的索引

1	2	3	4	5	6	7	8	9	10	11	12
赣	晋	浙	苏	湘	鄂	辽	吉	黑	甘	陕	豫
13	14	15	16	17	18	19	20	21	22	23	24
冀	鲁	新	川	青	云	桂	闽	粤	京	沪	皖

可以发现,该方法保证了对于不同汉字的识别性能。特别是对于一些相似字,如“浙”和“湘”,“粤”和“闽”等,本方法也保证了很高的识别率,其识别率都为99%左右。

部分识别错误的字符如图8所示。



图8 部分识别错误的汉字

识别错误的原因,主要集中在以下几个方面:在车牌定位过程中,汉字没有被完全定位出来,如“闽”字;部分汉字结构特别复杂且特别模糊,如“鄂”字,即使是人也很难分辨出来;把车牌边缘作为字符的一部分定位进来,如“苏”字。对于这些字符的识别,需要和车牌识别的其他部分如图象采集、字符分割等结合起来进一步研究,从而得到质量更好的字符图象。

6 结 论

本文提出了利用小波变换的边缘检测和局部特征提取特性,结合矩理论,直接从灰度图象中提取车牌汉字特征的方法。该方法避免了车牌汉字由于二值化而造成特征信息的丢失,并且提出利用特征向

量的每一维分量对识别率的贡献程度来进行特征选择。实验结果表明该方法可以得到较好的识别性能。

不同的特征有不同的表征意义。如何采取有效的算法,利用不同特征之间的互补性,从而集成各种特征的优点和抑制彼此的缺点,获得更好的识别性能,将是进一步研究的内容。

参 考 文 献

- 1 Lee Seong-Whan, Kim Young Joon. Direct extraction of topographic feature for gray scale character recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(7): 724~729.
- 2 Huang Lei, Huang Xiao. Multiresolution Recognition of (Offline) Handwritten Chinese Character With Wavelet Transform [A]. In: Proceedings of Sixth International Conference on Document Analysis and Recognition [C], Seattle, Washington, 2001: 631~634.
- 3 Mallat S G. A Theory for Multiresolution signal decomposition: The Wavelet Representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(7): 674~693.
- 4 Shen Ding gang, Ip Horace H S. Discriminative wavelet shape descriptors for recognition of 2-D patterns [J]. Pattern Recognition, 1999, 32(2): 151~165.
- 5 Unser M, Aldroubi A, Eden Murray *et al.* On the asymptotic convergence of B-spline wavelets to Gabor function [J]. IEEE Transactions on Information Theory, 1992, 38(2): 864~872.

潘翔 1977年生,2001年获浙江工业大学控制理论和控制工程硕士学位,现为浙江大学图形图象实验室博士研究生。研究方向为图象处理、模式识别和多媒体信息检索。



叶修梓 1966年生,教授,博士生导师,长江学者,美国 SolidWorks 首席科学家,1994年获德国柏林工业大学 CAD/CAM 博士学位。研究方向为计算机图形/图象、CAD、生物信息等。发表 SCI 论文 18 篇。



张三元 1966年生,教授,浙江大学图形图象实验室副主任,1992年获浙江大学 CAD/CG 博士学位。研究方向为计算机辅助设计、计算机图形学、网络图论等。发表学术论文 40 余篇。

