

一种有效的提高车牌首字符识别率的方法

吕文强 杨 健

(南京理工大学计算机科学与技术学院 南京 210094)

摘 要 针对车牌识别系统中由于低质车牌首字符特征提取困难而导致车牌首字符识别率不高的问题,提出了一种新的车牌汉字特征提取方法。该方法首先对车牌首字符的二值图像进行网格化处理,并对每一块网格区域提取字符笔画所在像素的占比、散度和质心 3 个特征分量,接着将提取到的所有的特征向量用支持向量机分类器进行训练,最终可以得到一组鲁棒性很强的分类器。实验结果表明,该特征提取方法与支持向量机分类器结合可以较大地提高车牌首字符的识别率。

关键词 车牌识别,支持向量机,字符识别,形状参数

中图法分类号 TP391.41 **文献标识码** A

Effective Way of Improving the Recognition Rate of License Plate's First Character

LV Wen-qiang YANG Jian

(College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract This paper offers a new method of extracting feature for solving the problem of the low Chinese character recognition rate resulted from the poor quality of the Chinese character image in the license plate recognition system. Firstly, the binary Chinese character image that has been segmented is divided into many blocks. Secondly, this paper extracts three stroke pixel's feature components that include the proportion of stroke pixels in the block, the divergence and the centroid for each block. Thirdly, this paper combines the new feature extraction method with the SVM classifier. At last, a group of robust classifiers are obtained. The experimental results show that the Chinese character recognition rate can be improved greatly.

Keywords License plate recognition, Support vector machine, Character recognition, Shape parameter

1 引言

随着智能交通系统的不断发展和完善,车牌识别系统越来越受到了人们的重视,目前该系统主要应用于道路监控和停车场安全等方面,具有广泛的应用前景。车牌识别系统主要由 3 大部分构成:车牌定位、字符分割和字符识别。其中车牌字符的识别是车牌识别系统中的最后一个环节,该环节中首字符识别率的高低决定着整个车牌识别系统的性能。

我国的汽车牌照主要由 7 个字符构成,其中首字符为各省的汉字简称,第 2 个字符为英文字母,其余 5 个字符为英文字母和阿拉伯数字的混合排列^[1]。在字符识别中,车牌首字符的识别已经成为影响整个车牌识别率的关键因素。由于车牌的首字符为汉字,笔画较多,当获取的车牌图像由于光照、污渍、天气、距离等原因导致图像分辨率较低时,汉字的笔画特征会变得非常模糊;又加上车牌首字符的类别较多,因此用一般的方法来识别首字符,会很难保证最终的整体识别率。

到目前为止,有关车牌识别系统的文献中大多数讨论的是英文字母和数字的识别,而单纯涉及车牌汉字识别的文献并不多见。归纳起来,涉及车牌汉字识别的算法主要分为两种:一种是基于模板匹配的汉字识别方法^[2-4],另外一种是基于

于字符笔画结构特征分析的车牌汉字识别方法^[5]。

基于模板匹配的汉字识别方法的基本原理是将待识别的汉字图像与已经准备好的字符模板进行匹配,然后根据它们两者之间的相似度来确定汉字字符的类别。这种方法的优点是实现起来较为简单,应用比较广泛;不足之处是实际拍摄的图像由于车牌污损和光照不均等因素会造成二值车牌汉字图像中存在粘连和笔画缺损,而这种情况会使得模板匹配的方法产生很大的局限性,最终识别效果不理想。基于汉字笔画结构分析的方法主要包括正交投影法^[6]、全像素法、粗网格像素统计法^[7]、基于笔画、轮廓和骨架特征法^[8]、小波矩特征提取方法等。这些方法同样对一些简单的汉字字符具有良好的识别效果,但对于一些笔画复杂或是存在粘连的图像识别能力就会很差,而且由于全像素法提取出的特征维数较大,因此会导致识别速度慢等问题。

针对以上字符特征提取方法的不足,本文提出了一种基于粗网格化的结构特征和统计特征相结合的特征提取方法,该方法的主要思想是对粗网格化后的二值首字符图像按网格顺序提取字符笔画的散度、质心和占比 3 种组合特征,由于该方法对图像进行了网格化处理,因此构造的特征空间的维数较小,非常适合利用一些机器学习的方法进行分类。本文

吕文强(1986—),男,硕士,主要研究方向为数字图像处理、模式识别, E-mail: lvwen_126@126.com; 杨 健(1973—),男,教授,博士生导师,主要研究方向为数字图像处理、人脸识别。

采用一些常用的机器学习方法,比如神经网络^[9-11]、支持向量机^[12]、K近邻算法^[13]等来对汉字字符进行了分类试验,并且与模板匹配法、正交投影法和小波矩特征等特征提取方法进行了对比。两组对比试验分别验证了不同机器学习方法的性能和本文提出的特征提取方法的可行性。

2 车牌首字符特征提取

2.1 首字符图像尺寸归一化

为了提高字符识别的准确率和整个车牌识别系统的性能,在对字符的特征进行提取之前,需要将分割后的不同分辨率的字符进行尺寸的归一化。尺寸归一化是指将分割得到的字符图像在分辨率上进行统一,使得所有即将进行特征提取的字符图像具有相同的宽度和高度。

目前,对字符的尺寸进行归一化的方法主要有两种:一是将字符图像的外边框按规定的大小和比例进行合理的缩放,该方法按宽高比例进行缩放,它不会受边框的影响。二是在图像的垂直和水平方向上对目标像素个数的分布按规定进行归一化,这种方法由于要计算整幅图像的散度,因此对边框等的噪声非常敏感,最终可能会干扰到字符图像的缩放。

鉴于双线性插值方法相比其他插值方法的优越性,本文采用双线性插值的算法来对首字符图像按规定的大小和比例进行合理的缩放,而且该方法同时也解决了缩放过程中的非整数问题。设原始字符图像的宽为 w ,高为 h ,模板字符图像的宽为 w_1 ,高为 h_1 ,则由原始图像和模板字符图像的宽高比可得宽度的缩放比例为 $p_w = \frac{w_1}{w}$,高度的缩放比例为 $p_h = \frac{h_1}{h}$ 。若像素点 (x, y) 对应原始图像中的任意一点,则该点经过缩放后对应在新图像中的位置 (x_1, y_1) 的变换矩阵为:

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} p_w & 0 & 0 \\ 0 & p_h & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

在本文中,将字符图像的尺寸归一化为 32×66 ,即图像宽度为32个像素,高度为66个像素。对部分分割出的首字符进行尺寸归一化的实验结果如表1所列。

表1 部分字符尺寸归一化的实验结果

序号	原图像	尺寸归一化后的图像
1		
2		
3		

2.2 字符特征提取

字符特征的提取是指从经过预处理的字符图像中提取到字符的关键特征,使得特征的维数尽可能少、正确分类的精度尽可能高。关键特征提取的方法能直接影响到最终的识别系统的性能。由于目前字符特征提取的好与坏并没有一个统一

的评价标准,因此对待不同的字符图像,只能通过实验对比的方式来判定哪种特征提取方法最优越。

本文提出了一种基于粗网格化的两种特征相结合的特征提取方法,其主要思想是对粗网格化后的字符图像按网格顺序提取字符笔画的散度、质心和占空比3种组合特征。具体实现方法如下。

(1) 散度

散度的大小在字符的二值图像中代表某一矩形区域 M 内字符笔画所包含的像素点相对区域中心的发散程度。散度值越大,表示字符笔画发散程度越大;散度值越小,表示字符笔画的发散程度越小。本文中计算的散度值包括垂直散度和水平散度两种。

二值图像中某一区域 M 内字符笔画结构的垂直散度和水平散度的计算公式如下。

垂直散度:

$$DivA_{Ver} = \sum_{\substack{y_1 \leq y \leq y_2 \\ y \in Y}} \sum_{\substack{x_1 \leq x \leq x_2 \\ x \in X}} |c_{_y} - (y - y_1) + 1| \quad (1)$$

水平散度:

$$DivA_{Hor} = \sum_{\substack{y_1 \leq y \leq y_2 \\ y \in Y}} \sum_{\substack{x_1 \leq x \leq x_2 \\ x \in X}} |c_{_x} - (x - x_1) + 1| \quad (2)$$

式中, (x, y) 为矩形区域 M 内字符笔画像素的坐标集, y_1, y_2, x_1, x_2 分别为区域 M 在整幅图像中上下左右的边界值。 $c_{_y}$ 和 $c_{_x}$ 分别为相对于 M 区域的垂直方向和水平方向的中心值。

最后本文将两个方向上的散度值合并得到 M 区域的总体散度值:

$$DivA_M = DivA_{Ver} \times width_M + DivA_{Hor} \quad (3)$$

式中, $width_M$ 为 M 区域的宽度。

(2) 质心

在字符二值图像中,质心的位置可以用来表示所处理的某一区域 M 中字符笔画整体的集中位置。在具体的实现中,物质系统中质点的位置可用所处理矩形区域 M 内的目标像素点的相对位置来表示,将目标像素点对应的灰度值作为该质点的质量。在字符二值图像中,设定目标像素的灰度级为1,背景像素的灰度级为0。本文中对质心的计算包括垂直质心和水平质心两种。

二值图像中某一区域 M 内字符笔画结构的垂直质心和水平质心的计算公式如下。

垂直质心:

$$CenR_{Ver} = \frac{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} \{(y - y_1) \times g(x, y)\}}{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} g(x, y)} \quad (4)$$

水平质心:

$$CenR_{Hor} = \frac{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} \{(x - x_1) \times g(x, y)\}}{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} g(x, y)} \quad (5)$$

式中, y_1, y_2, x_1, x_2 分别为区域 M 在整幅图像中上下左右的边界值。 $g(x, y)$ 表示坐标 (x, y) 处的像素的灰度值。

最后本文将两个方向上的质心的位置合并得到 M 区域

的整体质心的位置:

$$CenR_M = CenR_{Ver} \times width_M + CenR_{Hor} \quad (6)$$

式中, $width_M$ 为 M 区域的宽度。

(3) 占空比

占空比是指在字符二值图像中, 某一区域 M 内的目标像素的总个数与区域 M 内总像素的个数的比值。

字符区域 M 内目标像素占空比的计算公式为:

$$Ra_M = \frac{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} g(x, y)}{\sum_{y=y_1}^{y_2} \sum_{x=x_1}^{x_2} 1} \quad (7)$$

式中, $g(x, y)$ 表示坐标 (x, y) 处的像素的灰度值, 在这里设目标像素的值为 1, 背景像素的值为 0, y_1, y_2, x_1, x_2 分别为区域 M 在整幅图像中上下左右的边界值。

2.3 车牌首字符图像的粗网格化

网格化是指将预处理后的字符二值图像均匀地划分为多个子网格区域, 对每一个子区域进行多种特征的提取。提取出的网格化后的特征体现出了字符笔画相对于整体的一种分布情况。通过实验结果对比可知, 本文将预处理后的首字符二值图像均匀地划分为 6 行 4 列的 24 个网格时的系统性能最优。这样每幅车牌汉字图像可得到的特征维数为 $24 \times 3 = 72$ 维。

二值图像粗网格化的示意图如图 1 所示。



图 1 首字符二值图像粗网格化

3 车牌首字符识别

3.1 常用的字符识别方法

目前常用的字符识别方法主要有模板匹配法、人工神经网络、支持向量机等。

模板匹配的优点是方法简单, 对处理固定格式的汉字字符样本是很适用的; 缺点是对汉字字符笔画的粗细、分布变化非常敏感, 适应性比较差。

人工神经网络的优点是具有很强的学习能力和容错性能, 从而能较好地解决汉字字符中所提出的因字符残缺不全而无法识别的问题; 缺点是由于汉字字符类别数较多, 网络的结构和局部极小点非常难确定, 而且还会出现过学习和欠学习的情况, 导致汉字字符识别率较低^[14]。

支持向量机(SVM)是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的。首先, 在拥有少量汉字字符的训练样本集的前提下, 通过支持向量机可得到具有代表性的少数的关键样本(支持向量), 从而剔除了大量冗余的字符样本, 加快了分类速度; 其次该方法与汉字字符样本的维数无关, 并且当样本线性不可分时, 可通过引入一个核函数将线性不可分样本变为线性可分的样本^[15]。

假设目前有两类已知数据, 用 $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ 表示, 其中, l 为样本个数。以下给出 SVM 的决策函数:

$$f(x) = \text{sign}\left\{\sum_{i=1}^l \alpha_i^* y_i K(x_i \cdot x) + b^*\right\} \quad (8)$$

式中, x 为观测样本, x_i 为待测样本, $y_i \in (-1, +1)$ 为样本类别标号, $i=1, 2, \dots, l$, $\alpha_i (i=1, 2, \dots, l)$ 为拉格朗日乘子, $K(\cdot)$ 为核函数。 b^* 为偏差常数项, 可以用任意一个支持向量根据式 $y_i(\omega^* \cdot x_i + b^*) = 1$ 求得。

3.2 支持向量机分类

3.2.1 核函数的选取

SVM 是从线性可分情况下的最优分类面发展而来的, 但在实际情况下, 并不能保证每个样本都线性可分, 为了避免提供的汉字样本线性不可分, 在这里引入径向基核函数(RBF)(式(9))将样本的特征映射到高维空间, 使不可分的样本变为线性可分, 然后进行训练和分类^[16]。

$$K(x \cdot x_i) = \exp\left\{-\frac{|x - x_i|^2}{\sigma^2}\right\} \quad (9)$$

3.2.2 训练方法的选取

SVM 是一个两类分类器, 但是将要进行分类的汉字的类别总数多达几十种(本文中假设要分类的类别共有 N 种), 因此需要找到一种合适的训练方法, 使得 SVM 能够适用于多类别的分类。实际上, 到目前为止, 这种训练方法主要有“一对多”和“一对一”^[17]。

一对多: 训练时, 选取一个类别为正类, 其余所有的类别为负类。该方法的优点: 产生的分类器的数目少, 优化问题的规模较小, 分类速度快(只要进行 N 次判别就可判定待分类样本的所属类别); 缺点: 训练速度慢, 识别率相对不高。

一对一: 训练时, 选取一个类别为正类, 选取另一个类别为负类。优点是, 训练速度快识别率高; 缺点: 分类速度相对较慢(需进行 $\frac{N \times (N-1)}{2}$ 次判别, 但 DAGSVM 除外)。

DAGSVM(决策树 SVM)是“一对一”训练方法的一种, 该方法明显的一个优点就是分类速度快(只要进行 $N-1$ 次判别)。本文选用 DAGSVM 的训练方法来对样本进行训练。对任意两类样本集 A, B , 构造两组相同样本数目的训练样本, 一组正样本 $A_1 \in A$, 一组负样本 $B_1 \in B$ 。构造了训练样本之后, 便可对样本进行双线性插值大小归一化、OTSU 二值化和特征提取。

3.2.3 SVM 参数的寻优

由于使用了径向基(RBF)核函数, 因此本文需要确定的参数主要有两个, 惩罚因子 C 和方差 σ 。其中 C 决定着对误差的容忍度, C 越大, 越不能容忍误差。 σ^2 是核函数中唯一的参数, 它控制着在高维空间中的样本特征数据的复杂度。因此, 参数 C 和 σ 的值直接决定最终的分类器的分类性能。

本文决定选用网格法来寻找每个分类器最佳的 C 和 σ 。假设 C 的取值分别为 $[2^{-2}, 2^{-1}, \dots, 2^{11}, 2^{12}]$, σ^2 的取值分别为 $[2^{-5}, 2^{-4}, \dots, 2^8, 2^9]$, 共 225 个 C, σ^2 组合。将每组 C, σ 导入到 SVM 训练器中, 采用测试样本的识别率最高的一组参数来构造 SVM 分类器。

在选取了每个类别的训练样本并且找到了这个类别对应的分类器的最优参数 C 和 σ 之后, 便可得到该分类器的支持向量, 根据决策函数最终构造出了 $\frac{31 \times (31-1)}{2}$ 个车牌首字符的分类器^[18]。

4 实验结果及分析

参考文献

4.1 实验结果

本文中的实验使用的是在路口抓拍的车辆图像,对车牌进行定位、对字符进行分割以及对车牌首字符预处理之后,得到了 1045 幅大小为 32×66 的二值化后的车牌首字符图像。这些首字符中主要包括 31 个省的简称,因此目标是构建一个 31 类的多类分类器。本文将 1045 幅首字符图像分成两组,一组是训练样本集,包含 310 幅首字符图像;另外一组是测试样本集,共 735 幅图像。

本文实验中采用了多种机器学习中的分类算法,主要包括支持向量机、K 近邻、神经网络 3 种。将这 3 种分类算法在所构建的训练样本集上进行训练,并且对同一测试集进行测试识别。各个分类算法的参数为:支持向量机的径向基核函数的参数 σ 和惩罚因子 C 采用网格法获得,每个分类器具有不同的 σ 和 C ;K 近邻选取近邻个数 $K=5$;神经网络采用的是三层的网络结构。为了能够较好地分析实验结果,将 3 种算法对每一个首字符的识别率进行了统计,并记录了每一个首字符被误判为其他字符的概率。

为了验证本文提出的特征提取方法的有效性,在同一个测试样本集上对不同识别方法:模板匹配法,网格法和小波矩特征提取方法进行了实验对比。

4.2 实验结果分析

由表 2 可知,利用本文的特征提取方法 SVM, KNN 和神经网络对训练集的平均识别率分别为 98.5%, 97.2% 和 99.9%。说明这 3 种方法对训练样本集的拟合度都非常好。但对测试样本集进行识别后发现, SVM 分类方法的整体平均识别率为 95.3%, 在总体性能上要优于 KNN 和神经网络。

表 2 3 种分类算法在训练集和测试集上的性能对比

分类算法	平均训练误差(%)	平均测试误差(%)	平均识别耗时(ms)
SVM	98.5	95.3	22
KNN	7.2	93.9	116
神经网络	99.9	94.1	45

表 3 给出了本文首字符识别方法与模板匹配法、网格法、小波矩方法的实验对比结果。实验结果表明,本文中的识别方法要明显优于其他 3 种方法。

表 3 本文方法与其他识别方法的对比

识别方法	本文特征提取方法+SVM	模板匹配	网格法	小波矩
平均识别率(%)	95.30	84.21	90.62	82.84

结束语 本文主要讨论了车牌识别系统中车牌上汉字的识别问题,提出了一种新的粗网格化的特征提取方法。首先采用线性插值的方法对低分辨的车牌汉字图像的大小进行归一化,然后利用 OTSU 求阈值法对图像进行二值化,提取出图像的特征并训练 SVM 分类器,最后通过测试得到测试样本的识别率。由最后的字符识别率实验对比结果可知,与其他现有的车牌汉字识别算法相比,本文方法对大部分字符的识别率要优于其他算法。

- [1] 韩立明,王波涛. 车牌识别中关键技术的研究与实现[J]. 计算机工程与设计, 2010, 31(17): 3919-3923
- [2] Comelli P, Ferragina P, Graninri M, et al. Optical recognition of motor vehicle license plates[J]. IEEE Transactions on Vehicular Technology, 1995, 44(4): 790-799
- [3] Ko M, Kim Y. License plate surveillance system using weighted template matching[C]// Proceeding of the 32nd Applied Imagery Pattern Recognition Workshop. 2003: 269-274
- [4] Huang Y, Lai S, Chuang W. A template-based model for license plate recognition[C]// IEEE International Conference on Networking, Sensing and Control. 2004, 2: 737
- [5] Mori S, Suen C, Yamamoto K. Historical review of OCR research and development[J]. Proceedings of the IEEE, 1992, 80(7): 1029
- [6] Glauber M. Character recognition for business machines[J]. Electronics, 1956, 29(2): 132-136
- [7] Bokser M, inc C, Sunnyvale C. Ommidocument technologies[J]. Proceedings of the IEEE, 1992, 80(7): 1066-1078
- [8] Due T, Jain A, Taxt T. Feature extraction methods for character recognition-a survey [J]. Pattern Recognition, 1996, 29(4): 641-662
- [9] Koval V, Turchenko V, Kochan V, et al. Smart license plate recognition system based on image processing using neural network [C]// IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing System. 2003(8): 125-128
- [10] Pan X, Ye X, Zhang S. A hybrid method for robust ear plate character recognition[J]. Engineering Applications of Artificial Intelligence, 2005, 18(8): 963-972
- [11] Wang W. License plate recognition algorithm based on radial basis function neural networks[C]// Proceedings of the 2009 International Symposium on Intelligent Ubiquitous Computing and Education. 2009: 38-41
- [12] Chi X, Dong J, Liu A, et al. A simple method for Chinese license plate recognition based on support vector machine[C]// Proc of International Conference Communication, Circuits and Systems Proceeding. 2006: 2141-2145
- [13] 薛磊,杨晓敏,吴炜,等. 一种基于 KNN 与改进 SVM 的车牌字符识别算法[J]. 四川大学学报: 自然科学版, 2006, 43(5): 1031-1036
- [14] 边肇祺,张学工. 模式识别(第二版)[M]. 北京: 清华大学出版社, 2000
- [15] Theodoridis S, Koutroumbas K. Pattern Recognition[M]. 北京: 电子工业出版社, 2010
- [16] 奉国和. SVM 分类核函数及参数选择比较[J]. 计算机工程与应用, 2011, 47(3): 123-124
- [17] 刘健,刘忠,熊鹰. 改进的二叉树支持向量机多类分类算法研究[J]. 计算机工程与应用, 2010, 46(33): 117-120
- [18] 何健. SVM 及其在车牌字符识别中的应用[D]. 成都: 四川大学, 2005