

文章编号: 1008-374X(2006)01-0043-04

基于轮廓结构和统计特征的字符识别研究

贾 婧, 葛万成, 陈康力

(同济大学 电子与信息工程学院, 上海 200092)

摘 要: 车牌识别系统是智能交通不可缺少的一部分, 在车牌识别系统中, 字符的特征提取和字符识别是这一系统的关键技术. 文章利用字符的轮廓结构特征和统计特征对字符进行识别, 根据字符外部轮廓的上、下、左、右4个方向的特点和一些统计特征, 如字符最大宽度、垂直方向笔画数等作为识别特征, 形成判别树, 再利用判别树对汽车牌照中的数字和字母进行识别. 此方法识别的准确率比较高, 并提高了识别速度, 还适用于其他不同字体的数字和字母识别.

关 键 词: 结构特征; 统计特征; 字符识别

中图分类号: TP 391.41 **文献标识码:** A

0 引 言

文字的发明使得人类的交流更加方便和快捷, 它是信息承载和传递的工具, 使信息在时间和空间上都得以迅速的传播. 简单的几个字符就可以涵盖大量的信息, 通过认知这些字符我们可以了解到相关信息. 而随着计算机技术的发展, 人类进入信息时代, 远不满足直接用眼睛和耳朵来获取文字信息, 并记录在纸上, 人们开始寻求方法, 利用计算机代替人类简单、重复的劳动, 实现文字的自动输入、自动识别、自动输出. 近几年来, 随着计算机、图像处理等技术的发展, 文字识别技术也取得了巨大进步, 在汽车牌照的自动识别^[2], 商标自动翻译, 图书馆图书索引号自动识别等领域得到了一定的应用.

模式识别是在输入样本中寻找特征并识别对象的一种技术. 模式识别主要有2种方法: 一种是根据统计特征进行识别, 另一种是根据对象的结构特征进行识别, 而后者常用的方法为句法识别. 本文以车辆牌照的字符为例, 介绍一种依据字符的轮廓结构特征和统计特征相结合的方法^[1], 利用结构语句识别的方法进行数字和字母的识别, 能够实现车牌数字和字母的准确识别, 并提高了识别速度, 此方法同样适用于其他不同字体的数字和字母识别.

基于字符特征的字符识别算法一般包括以下3个步骤:

- 1) 提取字符的特征, 建立一个字符特征向量空间, 每个字符对应向量空间中的一个向量;
- 2) 定义分类函数;
- 3) 对输入的字符图像自动提取相应的特征, 根据分类函数进行特征匹配, 识别出相应的字符.

用于字符识别的分类特征应满足以下要求:

- 1) 有较强的分类能力, 类内各样本距离应尽量小, 类间距离应尽量大; 特征用于粗分类时, 粗分类后各子集的样本数不宜过多, 其数量应大体相等;
- 2) 具有较高的稳定性和鲁棒性, 尽量减少笔划断裂或粘连的影响;
- 3) 易于提取, 在保证系统性能稳定的前提下, 特征的维数不宜过高, 以减少运算量和存储复杂度.

我们识别的对象是经过图像预处理后的二值化单个字符, 见图1.

由于车牌图像会存在噪声和污损, 而且在进行二值化处理和字符分割后可能使字符变形, 使得我们在提取特征的时候受到干扰. 我们从字符的外部轮廓的顶部、左侧、底部、右侧4个方向进行特征提取, 可以减少以上缺损和变形造成的干扰. 当其中某部位的笔画发生变形时, 不会改变或减少对其他部位特

收稿日期: 2004-11-10

作者简介: 贾 婧(1981-), 女, 山东滕州人, 同济大学硕士研究生; 葛万成(1964-), 男, 湖北襄阳人, 同济大学教授, 博士, 博士研究生导师.

征的影响.



图 1 车牌字符

另外引入轮廓的统计特征,例如水平方向的最大字符宽度,垂直方向的笔画数等,可以区分像 0、8、Q、G 等这些外部结构特征相似的字符.

1 轮廓结构特征

1.1 外部轮廓定义

对于大小为 $m \times n$ 的字符图像, $P(x, y)$ 表示图像中坐标为 (x, y) 的像素点, U 表示字符像素点的集合,

左侧轮廓: $LP(i) = \min\{x | P(x, y) \in U, y = i\} \quad i = 1, 2, \dots, m$

定义为字符最左侧边界像素点的水平方向坐标值.

右侧轮廓: $RP(i) = \max\{x | P(x, y) \in U, y = i\} \quad i = 1, 2, \dots, m$

定义为字符最右侧边界像素点的水平方向坐标值.

顶部轮廓: $TP(j) = \min\{y | P(x, y) \in U, x = j\} \quad j = 1, 2, \dots, n$

定义为字符最高边界像素点的垂直方向坐标值.

底部轮廓: $BP(j) = \max\{y | P(x, y) \in U, x = j\} \quad j = 1, 2, \dots, n$

定义为字符最低边界像素点的垂直方向坐标值.

外部轮廓定义如图 2 所示.

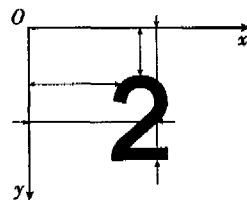


图 2 外部轮廓定义图示

1.2 结构基元

利用轮廓的一阶微分变化趋势,定义构成字符轮廓的基本基元.定义 4 个方向轮廓的一阶微分:

$$LPD = LP(i+1) - LP(i)$$

$$RPD = RP(i+1) - RP(i)$$

$$TPD = TP(j+1) - TP(j)$$

$$BPD = BP(j+1) - BP(j) \quad i = 1, 2, \dots, m-1, j = 1, 2, \dots, n-1.$$

基本基元共设 5 个:左斜(L)、右斜(R)、数值(V)、圆弧(C)、突变(P).另 SL, SV 和 SR 分别为某侧轮廓一阶微分值小于 0、等于 0、大于 0 的个数,若

1) $SL=0, SR=0$, 定义结构为 V, 如图 3(a);

2) $SR=0, SL$ 大于阈值 LT, 则结构为 L, 如图 3(b);

3) $SL=0, SR$ 大于阈值 RT, 则结构为 R, 如图 3(c);

4) $SR > RT, SL > LT$, 则结构为 C, 如图 3(d), 即包含了上升和下降 2 种趋势;

5) 当某侧有一阶微分值超过阈值 PT 时, 表示字符轮廓不连续, 定义结构为 P, 如图 3(e).

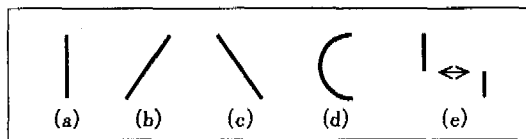


图 3 结构基元图示

1.3 基元的检测

我们以检测到突变为转折, 设 $PD(k)$ 表示某侧轮廓的一阶微分, $k = 1, 2 \sim K$, SL, SV, SR 分别为检测到突变时的微分值小于 0、等于 0 和大于 0 的个数, PT, LT, RT, T 分别为选定的突变, 左斜、右斜和基元检测的阈值, 这里阈值的选取要根据字符的大小, 在实验中不断修正. 具体的检测规则如下:

1) 若 $PD(k) \geq PT$, 在 k 处检测到结构突变(P);

2) 若 $SL + SV + SR \geq T$, 进行基元检测, 如下:

若 $SL < LT, SR < RT$, 检测到结构为竖直(V);

若 $SL > LT, SR < RT$, 检测到结构为左斜(L);

若 $SL < LT, SR > RT$, 检测到结构为右斜(R);

若 $SL > LT, SR > RT$, 检测到结构为圆弧(C);

若 $SL + SV + SR < T$, 记录结构为(P).

3) 检测到突变时, 要对突变前后的轮廓特征分别检测, 即若在 $k1$ 处检测到 P, 则应在 $[1, k1-1]$ 范围内按 2) 进行检测. 在 $k2$ 处发现突变, 则应在 $[k1, k2-1]$ 范围内按 2) 进行检测, 以此类推. 对于突变后的剩余部分, 也要按 2) 进行检测.

设定 T 表示当无法判别结构特征时的结构未定和初始状态, 则以字符“2”为例, 它的结构特征分别为: 顶部: TC; 左侧: TLPL; 右侧: TCP; 底部: TV. 字符“F”为: 顶部: TV, 左侧: TV, 右侧: TPVPV, 底部: TPV. 字符“K”的结构特征为: 顶部: TPL, 左侧: TV, 右侧: TC, 底部: TPR. 字符“5”的结构特征可检测为:

顶部: TV, 左侧: TVPC, 右侧: TPVC, 底部: TC, 如图 4 所示. 其他字符的特征这里就不一一列举了, 而且根据实际的识别对象, 不同的字体, 相应的结构特征会有所不同, 这需要通过实验进行统计归纳.

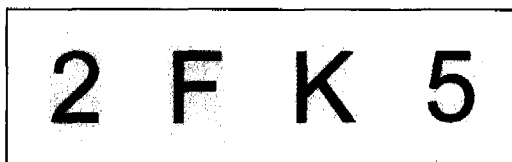


图 4 结构特征举例图

2 轮廓统计特征

单单采用上述的结构基元还不能完全识别车牌的数字和字母, 于是我们引入了轮廓的统计特征.

1) 水平方向的最大字符宽度 W :

$$W = \max\{RP(k) - LP(k)\}$$

当 $W \leq H/2$, 即为数字 1, ($H = m$)

该特征主要用于识别数字 1, 并避免 T, L 和 1 发生混淆.

2) 垂直方向的笔画数 comp:

通过记录字符垂直方向的笔画数, 可以区分 0 和 8, B 和 D 等容易混淆的字符.

具体算法是: 在二值化字符图像里, 从上至下扫描第 $n/2$ 列, 记录像素由 0 到 1 和继而由 1 到 0 的变更次数, 即为这一方向上的笔画数目.

3 数字字母的识别算法

字符的结构特征直观地反映了字符的属性, 利用字符的结构特征识别字符要充分地利用字符的特点. 采用统计特征进行识别则是充分利用了计算机处理数字特征能力强的特点, 用这种特征构造的字符识别系统比较简便.

通过实验, 对大量样本的轮廓特征进行统计和分析, 总结出这 34 个字符的局部轮廓结构特征和轮廓统计特征组成的特征向量. 根据实验情况, 我们的识别算法的基本思想是, 从顶部轮廓特征开始, 根据顶部特征将字符分成几类, 再在每一类中利用其他三边的轮廓结构特征和统计特征进行细分, 形成一个判别树, 直到确定单个字符的特征“模板”, 识别时只要对号入座进行匹配即可. 具体如下:

top、left、right、bottom 表示四边轮廓的结构基元特征数组, XT、XL、XR、XB 分别表示对应数组的元素个数. 例如字符“2”的 XT 为 2, XL 为 4, XR 为 3, XB 为 2; 字符“F”的 XT 为 2, XL 为 2, XR 为 5, XB 为 3; 字符“K”的 XT 为 3, XL 为 2, XR 为 2, XB 为 3; 字符“5”的 XT 为 2, XL 为 4, XR 为 4, XB 为 2.

1) $XT > 2$ 的字符有: 6, L, J, H, K, N, V, W, A, U, X

其中 $XR > 2$ 的字符可能是 L, 6, 若 $left(2) = 'V'$, 则识别为 L, $left(2) = 'C'$, 识别为 6;

若 $XL > 2$, 则字符为 J;

剩余的字符里面, 若 $XB > 2$, 则可能是 H, K, N, V, W. 若 top 中既有 'L', 又有 'R', 则识别为 V 或 W, 他们两者的区别是 bottom 中是否有 'P'. 若 top 中有 'L' 没有 'R', 则为 K, 若有 'R' 没有 'L', 则为 N, 若 top 中有 'V' 的则是 H.

剩下的是 A, U, X. 若 $comp > 1$, 则字符识别为 A. 否则, 若 $left(2) = 'V'$, 则为 U, 否则为 X.

2) $XT = 2$ 且 $top(2) = 'C'$ 的字符有: 0, 2, 8, 9, C, G, Q, S, M, Y

若 $XL = 2, XR = 2$, 可能的字符为 0, 8, M, Q, Y. 若 $XB > 2$ 则为 M, Y, 如果 bottom 中既有 'L' 又有 'R',

R', 则识别为 Y, 否则为 M. 若 $\text{comp}=2$, 则为 0, 若 $\text{comp}>2$ 且 $\text{left}(2)='V'$, 则应识别成 Q, 否则为 8.

否则, 也可能是 2, 9, C, G, S 中的一个. 若 $\text{XB}>2$ 则应为 9. 否则, $\text{XL}=2$ 的字符可能是 C, G, 而 $\text{right}(\text{XB})='V'$ 的为 G, 否则为 C. $\text{XL}>2$ 的可能是字符 2 和 S, 他们的区别是 $\text{bottom}(2)$ 是 'V' 还是 'C'.

3) $\text{XT}=2$ 且 $\text{top}(2)='V'$ 的字符有: 3, 5, 7, B, D, E, F, P, R, T, Z

若 $\text{XR}>2$ 则有可能是 E, F, T, Z, 5, P. 其中, 若 $\text{XL}=2$ 且 $\text{left}(2)='V'$ 的有 E, F, P, 若此时的 $\text{right}(2)='C'$, 则字符应为 P, E 和 F 的区别在于 XB 是否大于 2 且底部是否有突变. 否则, 若 $\text{XB}=2$, 则有可能是 5 或 2, 两者的区别是底部特征是 'C' 还是 'V'. 不然, 就是 T 了.

若 $\text{XR}=2$, 则有可能是 B, D, R, 3, 7. 其中, 若 $\text{XL}>2$, 可能是 3 或 7, 其区别在于 comp 大于 2 还是等于 2. 否则 $\text{XL}=2$, 若 $\text{XB}>2$, 字符为 R, 否则, 若 $\text{comp}>2$, 就为 B, $\text{comp}=2$, 字符就为 D.

4) $\text{XT}=2$ 且 $\text{top}(2)='L'$ 的字符为 4.

5) 若 $W<H/2$, 字符为 1.

4 实验总结

神经网络识别方法^[3]因为其高适应性成为模式识别较通用的方法, 但其需要大量训练样本, 并且训练时间较长. 模板匹配方法^[4]是先从图像中检测出已知形状的目标物, 然后用这个目标物的形状模板与图像匹配, 在约定的某种准则下检测出目标物图像. 因为要取得最佳匹配结果, 也需要较长的匹配时间. 而本文中的方法利用字符的先验知识, 不需要大量样本, 采用判决树结构, 节省了识别时间. 在 matlab^[5]环境下, 对从车牌中分割出来的字符进行识别测试, 图像大小归一化为 24×18 , 识别率可达到 96%, 误识的字符集中为 0, 8, B, D 等, 还需在算法中加入其他特征加以改善.

以上只是给出简单的识别算法, 在实际情况中, 由于字体的不同或图像预处理的效果影响, 字符的实际轮廓结构特征会有所不同.

(注: 本文受到德国罗德与施瓦茨公司资助)

参考文献:

- [1] 四维科技, 胡小锋, 赵 辉. 图像处理与识别应用案例精选[M]. 北京: 人民邮电出版社, 2004. 196-230.
- [2] 荣 江, 王文杰, 陈建华. 智能车牌识别系统设计与实现[J]. 微型机与应用, 2005, 24(1): 50-51, 60.
- [3] KOVAL V, TURCHENKO V, KOCHAN V, et al. Smart License Plate Recognition System Based on Image Processing Using Neural Network[J]. IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems, 2003, (10): 123-127.
- [4] 路小波, 凌小静, 黄 卫. 一种基于模板匹配的汽车牌照识别方法[J]. 交通与计算机, 2004, 22(6): 79-82.
- [5] 孙兆林. MATLAB6.x 图像处理[M]. 北京: 清华大学出版社, 2002. 29-105.

Character Recognition Based on Structural and Statistical Features

JIA Jing, GE Wan-cheng, CHEN Kang-li

(College of Electronic and Information Engineering, Tongji University, Shanghai 200092, China)

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 character. This paper introduces a character recognition algorithm based on structural and statistical features of character's figures. According to the features on all sides of top, bottom, left and right, as well as some statistic features, such as the width of characters, the count of strokes across the vertical direction and so on, we have designed a tree structure to recognize the numbers and letters from the license plate. This algorithm can recognize the numbers and letters successfully in a shorter time. In addition, it can be used to recognize other kinds of characters.

Key words: structural features; statistical features; character recognition