

基于轮廓检测的第二代身份证字符识别

陈 伟,王小鹏,成 亮

CHEN Wei,WANG Xiao-peng,CHENG Liang

兰州交通大学 电子与信息工程学院,兰州 730070

School of Electronic & Information Engineering,Lanzhou Jiaotong University,Lanzhou 730070,China

E-mail:dwwei6617@sohu.com

CHEN Wei,WANG Xiao-peng,CHENG Liang,Second generation identification card's character identification using profile detection.Computer Engineering and Applications 2009 45(22) 147-149.

Abstract: Aiming at the second generation identification card's character identification, this paper puts forward a method that based on character string's profile detection. This method firstly adopts texture segmentation of the area of image character after gaining the area of character, then adopts fluctuation profile concavo-convex character to approximately detect single character's width, and accordingly choose the steady local character and use the method about frame sentence identification to identify the character. Through the experimentation of ten pieces of the second generation identification card's character identification which is screened in the nature scene, this method is proved validity.

Key words: digital binary, threshold, gray scale, texture, profile detection, character identification

摘 要: 针对第二代身份证图像字符识别,提出了一种基于字符串轮廓检测的方法。该方法采用对图像字符区域先进行纹理分割,获取字符区域后对其上下轮廓凹凸特征近似检测单个字符的宽度,从中选出稳定的局部特征,利用结构语句识别的方法进行字符识别。通过对自然场景下拍摄的10幅第二代身份证图像的字符识别实验,证明了该方法的有效性。

关键词: 二值化,阈值,灰度值,纹理特征,轮廓检测,字符识别

DOI:10.3778/j.issn.1002-8331.2009.22.048 文章编号:1002-8331(2009)22-0147-03 文献标识码:A 中图分类号:TP391

1 引言

目前OCR(Optical Character Reader)文字识别技术已相对成熟,但是对粘连字符识别错误率还比较高。一般认为主要原因是粘连字符的错误切分而导致字符的严重失真变形,无法正确地识别,因此粘连字符的切分成为提高识别率的关键技术^[1]。现有的切分方法主要有:(1)基于图像分析的直接切分法,通过图像分析寻找字符之间较为合理的切分点,但切分错误率比较高;(2)基于识别的切分法,先通过图像分析,确定几个可能的切分点,借助识别结果,选择合理的切分点^[2]。后者切分方法的识别比较高,但是多次识别,步骤繁琐,比较耗时。

为了提高粘连字符识别率与识别速度,提出了一种先对图像进行字符区域纹理分割,进而对分割出的字符区域采用字符轮廓结构特征和统计特征相结合的方法。该方法从中选出稳定的局部特征,利用结构语句识别的方法进行字符的识别,能够有效地提高识别正确率的同时简化步骤降低运算量。

2 原始图像的预处理

身份证图像样本如图1,是通过扫描仪等设备获取的彩色图像。由于彩色图像包含大量的颜色信息,在计算机处理时也

会占用大量资源,降低了运行速度,而且实际采集的图像由于光照影响,使得整幅图偏暗或偏亮,没有灰度层次感。因此在对图像进行识别时常将彩色图像转变为灰度图像,以加快处理速度,然后对其进行图像预处理。在预处理过程中主要是针对图像的偏暗或偏亮问题,对图像进行灰度拉伸,将灰度范围拉伸^[3]。在灰度拉伸的基础上对图像采用空间域法进行图像的滤波,达到图像平滑的目的,能够有效地滤除高频噪声。经过灰度拉伸和图像平滑过程后,然后对图像进行二值化处理可得到的二值图像如图2。



图1 原始图像



图2 预处理后的二值图像

3 身份证号码与汉字字符的提取

图像字符分割有多种方法,可以充分利用字符固定的高宽比和间距作为先验知识对目标图像进行精确的字符定位切分;

基金项目:甘肃省自然科学基金(the Natural Science Foundation of Gansu Province of China under Grant No.0710RJ2A046)。

作者简介:陈伟(1980-)男,硕士研究生,研究方向为图像分析、图像识别;王小鹏(1969-)男,博士,副教授,硕士生导师,研究方向为图像分析与识别、多媒体信息处理等;成亮(1982-)男,硕士研究生,研究方向为图像分析、电子稳像。

收稿日期:2008-04-28 修回日期:2008-07-28

可以根据图像垂直投影直方图的期望与方差确定一个闭值^[4], 以此来判定某字符的起始位置和结束位置, 从而分割出该字符; 也可以利用字符投影图的波峰、波谷特性先进行字符的水平分割, 确定每个字符的上下边界, 然后进行字符的垂直分割, 确定每个字符的左右边界, 从而实现字符的分割。由于图像中字符与背景的灰度特征不同^[5], 可以根据图像灰度分割出图像中的字符串, 并通过二值化处理使其中的字符呈黑色, 背景呈白色。为此, 需找到一个合适的灰度分割阈值。传统的直接采用灰度直方图的方法不仅无法准确获得图像中字符灰度相对于背景灰度的高低, 而且难以取得理想的分割阈值。采用基于纹理的分割阈值选取方法。该方法首先对图像作纹理分析, 确定图像中字符像点的灰度相对背景图像底色灰度的高或低, 同时取得字符灰度与背景图像底色灰度的近似分布, 最后采用模式识别技术中的最大最小准则获取灰度分割阈值^[6]。

这里只对姓名与号码区域做详细介绍, 其他文字区域与姓名文字区域类似。图3是分割后的文字区域, 图4为号码分割后的区域。

成 亮

图3 分割后姓名文字区域图像

271791009587019251

图4 分割后号码字符区域图像

4 字符识别

字符经分割后需考虑以下4种情况: (1) 字符可能存在倾斜角度; (2) 字符图像中可能含有随机黑色区域干扰; (3) 字符图像中可能出现笔画融合、断裂、部分缺失和分割引入的误差, 因此对字符区域中的文字与数字区域分别进行识别, 以提高识别率。

4.1 文字识别

由于受噪声和随机污点干扰, 以及二值化和粘连字符处理会引起字符的变形, 为了尽量减少这种变形对信息特征的干扰, 从变形的字符中提取可靠的特征信息, 将文字的整体轮廓分解为顶部、底部、左侧和右侧4个方向的轮廓特征^[7]来描述, 使得当其中某部位的笔画发生变形时, 不会改变或减少其对其他部位特征的影响。

4个轮廓分别为左侧轮廓 LP 、右侧轮廓 RP 、顶部轮廓 TP 和底部轮廓 BP 。左侧轮廓($LP(k)$ $k=1, 2, \dots, M$)定义为字符最左侧边界像素点的水平方向坐标值。

$$LP(i) = \min\{x | P(x, y) \in C, y=i\} \quad i=1, 2, \dots, M \quad (1)$$

其中 $P(x, y)$ 表示图像中坐标为 (x, y) 的像素点, C 表示字符像素的集合。同理, 右侧轮廓($RP(k)$ $k=1, 2, \dots, M$)定义为字符最右侧边界像素点的水平方向坐标值。

$$RP(i) = \max\{x | P(x, y) \in C, y=i\} \quad i=1, 2, \dots, M \quad (2)$$

相应地, 顶部轮廓($TP(k)$ $k=1, 2, \dots, N$)定义为字符最高边界像素点的垂直方向坐标值。底部轮廓($BP(k)$ $k=1, 2, \dots, N$)定义为字符最低边界像素点的垂直方向坐标值。

$$TP(i) = \min\{y | P(x, y) \in C, x=i\} \quad i=1, 2, \dots, N \quad (3)$$

$$BP(i) = \max\{y | P(x, y) \in C, x=i\} \quad i=1, 2, \dots, N \quad (4)$$

为了描述轮廓的变化特性, 定义4个方向轮廓的一阶微分:

$$LPD = LP(i+1) - LP(i) \quad RPD = RP(i+1) - RP(i)$$

$$TPD = TP(i+1) - TP(i) \quad BPD = BP(i+1) - BP(i) \quad (5)$$

其中 $i=1, 2, \dots, M-1$ $j=1, 2, \dots, N-1$ 。

其中轮廓检测判别规则的结构基元如下: 利用式(5)的轮廓一阶微分变化趋势, 可定义构成字符轮廓的基本基元。基本基元共5个, 分别为左斜(L)、右斜(R)、竖直(V)、圆弧(C)和突变(P)^[7]。以左侧轮廓为例, 假设左侧轮廓一阶微分值不小于0, 则为结构 V 。假设左侧轮廓一阶微分值不等于0且小于0的个数大于其左侧阈值, 则为结构 L 。假设左侧轮廓一阶微分值不等于0且小于0的个数大于其右侧阈值, 则为结构 R 。假设左侧轮廓一阶微分值等于0的个数大于其右侧阈值且小于0的个数大于其左侧阈值, 则为结构 C 。对于连续的字符轮廓, 其一阶微分值的变化量比较小, 而当字符轮廓不连续时, 其一阶微分值相对较大。因此, 当轮廓的一阶微分值超过阈值时则字符轮廓有变化, 即为结构 P 。

由于字符轮廓突变处, 表示字符的轮廓不连续, 因而突变前后的轮廓特征必须分别进行检测。即若 k_1 处检测到 P , 则在 $[1, k_1-1]$ 的字符轮廓范围内统计其轮廓一阶微分值大于0、小于0和等于0的个数, 对其独立进行结构基元检测。若在 k_2 处又检测到 P , 则在 $[k_1+1, k_2-1]$ 范围内进行基元检测, 依此类推。检测到突变结构 P 的有效范围在 $x \in [ST, N-ST+1]$ $y \in [ST, M-ST+1]$, 其中 ST 表示字符笔画的宽度。这主要是为了避免干扰严重情况下, 轮廓边缘光滑处理不够理想时, 可能检测到的假突变基元。

4.2 数字识别

对于字符中的数字, 将数字字符的顶部、左右两侧的局部轮廓结构特征和轮廓统计特征组合成特征向量, 用以描述10个数字。根据特征向量, 采用轮廓检测算法识别底部残缺的和完整的数字字符。由于底部特征丢失, 会改变左右两侧的部分结构特征, 但不会影响顶部特征, 因此特征描述和机构匹配识别都从顶部轮廓特征开始。

采用上述结构特征还不足以准确识别数字“1”、“0”和“8”, 因此引入轮廓的统计特征。当最大轮廓宽度小于字符间距的一半时, 则认为是数字“1”。由于“0”和“8”的轮廓结构特征及其相似, 所以借助于垂直方向的笔划数加以区分。受数字底部残缺的影响, “8”在垂直方向的最大笔划数也可能是2。此时采用检测到笔划数为2时垂直方向的最小值来代替。假设 j 列上像素点 $p(j, i-1)$ 检测到垂直方向的笔划数为1, 在 $p(j, i)$ 检测到了第二个笔划, 则记下第二个像素点的位置 i , 并判断第二个像素的位置, 当位置 i 小于字符高度一半时, 检测到的字符为“8”, 否则为“0”。

5 仿真实验

为了验证方法的有效性, 在仿真平台为matlab 7.0下, 对经过纹理分割后的图3与图4采用轮廓识别。图5与图6是对名字区域单个字符识别与其轮廓曲线识别后的结果。图7为对号码区域单个数字识别的结果。

图6所示的轮廓曲线, 姓名区域“成”字上轮廓为双峰值, 下轮廓为4个峰值, “亮”字上轮廓为一个峰值, 下轮廓为两个峰值, 轮廓结构符合汉字的特征, 可识别为图5所示的汉字字

成 亮

图5 名字区域字符识别后结果

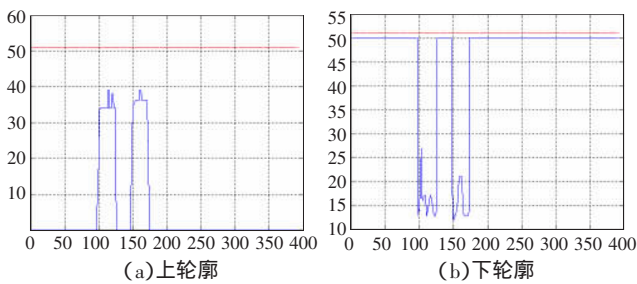


图6 名字区域字符串的上与下轮廓曲线

271791009587019251

图7 号码字符区域识别结果

符。在数字号码区域能够有效地识别出单个数字字符,并很好地区分了字符“0”与“8”。

字符区域识别过程中,必须设置行方向的间隔阈值 T_y 。如果间隔阈值 T_y 设置太小,即间隔阈值 $T_y \leq D_y$ (D_y 为两可能文字图像行的间距),则会出现对左右偏旁汉字识别为两个汉字,没有达到识别的预期效果。例如,在文字区域里的住址区域,出现了“坝”与“坪”字,此时识别出的效果如图8与图9所示。

土

贝

土

平

图8 为“坝”字误识别后的结果 图9 为“坪”字误识别后的结果

这样会把每个左右偏旁的汉字识别为独立的两个甚至多个文字。所以必须事先设置间隔阈值 $T_y \geq D_y$,则左右偏旁的文字进行合并,可识别出理想的效果。

经过对10张不同身份证,在不同光照和角度图像的字符识别,并采用文字分割识别的正确率(D-rate)指标^[8]来评价算法的性能。10幅图像中共包括字符区域420个,正确识别出文字区域408个,遗漏文字区域12个,误识别非字符区域7个,即D-rate为97.14%。以上数据表明此方法可以在图像中准确地识别字符区域。

(上接122页)

和Adaboost比较可知,KCNE方法在部分数据集上体现出了较好的性能。在以后的工作中,要进一步研究神经网络、决策树和支持向量机等模型的聚类算法,以及不同融合方法对集成学习性能的影响。

参考文献:

- [1] Opitz D, Shavlik J. Generating accurate and diverse members of a neural-network ensemble[C]//Advances in Neural Information Processing Systems. Denver, CO: MIT Press, 1996: 535-543.
- [2] Rosen B E. Ensemble learning using decorrelated neural networks[J]. Connection Science, 1996, 8: 373-383.
- [3] Zhou Z H, Wu J, Tang W. Ensembling neural networks: Many could be better than all[J]. Artificial Intelligence, 2002, 137(1/2): 239-263.
- [4] Altincay H. Decision trees using model ensemble-based nodes[J]. Pattern Recognition, 2007, 40(12): 3540-3551.

6 结束语

针对第二代身份证图像字符识别,提出了一种基于字符串轮廓检测的方法。其优势在于对二值图像中对字符区域先采用基于文字边缘纹理进行粗分割,有效避免了纹理特征的复杂计算,同时尽可能多地检测出可能的字符区域。然后对粗分割区域采用上下轮廓凹凸特征近似检测单个字符的宽度。由于文字区域字符间隔的不统一性,需对分割出的字符区域事先设置间隔阈值,有效避免了对左右偏旁识别过程中的错误。并利用切分处笔划宽度特征,对切分处的变形笔划进行了恢复,有效抑制噪声干扰而产生的字符笔划变形,实现了对身份证单个字符的切分与识别。在自然场景中拍摄的第二代身份证图像的字符识别实验中,识别正确率为97.14%,证明了方法的有效性。

参考文献:

- [1] 阮秋琦. 数字图像处理[M]. 2版. 北京: 电子工业出版社, 2007: 412-420.
- [2] Fumera G, Roli F. A theoretical and experimental analysis of linear combiners for multiple classifier systems[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(6): 942-956.
- [3] Lienhart R, Wernicke A. Localizing and segmenting text in images and videos[J]. IEEE Trans Circuits and Systems for Video Technology, 2002, 12(4): 256-268.
- [4] 胡小锋, 周勇, 叶庆泰. 复杂背景彩色图像中的文字分割[J]. 上海交通大学学报, 2006: 141-147.
- [5] 叶晨洲, 杨杰, 宣国荣. 车辆牌照字符识别[J]. 上海交通大学学报, 2000: 672-675.
- [6] 王学文, 丁晓青, 刘长松. 基于Gabor变换的高鲁棒汉字识别新方法[J]. 电子学报, 2002, 9: 1317-1321.
- [7] 胡小锋, 赵辉. Visual c++/MATLAB图像处理与识别实用案例精选[M]. 北京: 人民邮电出版社, 2004: 194-253.
- [8] Li C, Ding X Q, Wu Y S. Automatic text segmentation and text recognition for video indexing[J]. Multimedia Systems, 2000, 8(1): 69-81.
- [5] Ko A H R, Sabourin R, de Souza Britto Jr A et al. Pairwise fusion matrix for combining classifiers[J]. Pattern Recognition, 2007, 40(8): 2198-2210.
- [6] Sirlantzis K, Hoque S, Fairhurst M C. Diversity in multiple classifier ensembles based on binary feature quantisation with application to face recognition[J]. Applied Soft Computing, 2008, 8: 437-445.
- [7] Giacinto G, Roli F. Design of effective neural network ensembles for image classification purposes[J]. Image and Vision Computing, 2001, 19(9/10): 699-707.
- [8] Yang M S, Wu K L. A similarity-based robust clustering method[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(4): 434-448.
- [9] Yu J, Huang H. A new weighting fuzzy C-means algorithms[C]//The 12th IEEE International Conference on Fuzzy Systems, St Louis, Missouri, USA, 2003: 896-901.
- [10] Blake C L, Merz C J. UCI repository of machine learning database [EB/OL]. (1998). <http://www.ics.uci.edu/~mlearn/MLRepository.html>.