

基于图像识别技术的手写数字识别方法

吴忠^{1,2}, 朱国龙^{1,2}, 黄葛峰^{1,2}, 吴建国^{1,2}

(1. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

2. 安徽大学 软件学院, 安徽 合肥 230039)

摘要:随着世界经济越来越发达,各国经济交流日益加深,人们每天要处理大量的票据。手写数字在这个领域是必不可缺少的,例如人们要处理许多支票、发票、货物单等数据,这些都要大量与数据打交道。传统的方法是利用光学的光电变换原理对数字进行识别,该方法对环境要求比较高,成本较高,识别率和实时性比较低。文中提出基于图像识别的对0至9手写数字识别技术。实验表明,对2,3,4,5,8,9字符准确识别,而对1,7,0,6,数字基本识别无误,该方法提高了对数字识别率和实时性。

关键词:数字识别;图像识别;手写数字

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2011)12-0048-04

Handwritten Digit Recognition Based on Image Recognition System

WU Zhong^{1,2}, ZHU Guo-long^{1,2}, HUANG Ge-feng^{1,2}, WU Jian-guo^{1,2}

(1. School of Computer Science and Technology, Anhui University, Hefei 230039, China;

2. School of Software Engineering, Anhui University, Hefei 230039, China)

Abstract: As the world economy is more developed, countries deepening economic exchanges, people has to deal with a lot of notes each day. Handwritten digits in this field is necessary, for example, to deal with a lot of people check, invoice, goods such as single data, which have a lot to deal with data. The traditional method is to use the optical principle of the photoelectric conversion to identify digit, it is relatively high for environmental requirements, high cost, low recognition rate and real-time. In this paper, handwritten numeral recognition technology from 0 to 9 is proposed based on image recognition. Experiments show that accurately identify 2, 3, 4, 5, 8, 9 characters, while the 1, 7, 0, 6, basic identification number is correct, this method improves the recognition rate of digital and real-time.

Key words: digital identification; image recognition; handwritten numerals

0 引言

数字是人们日常生活进行信息交流中不可缺少的重要工具之一,实现计算机的手写数字识别是加快社会信息化进程的关键所在。人们经常运用的数据都是阿拉伯数字,所以手写体数字的识别研究一直以阿拉伯数字为主导。手写数字识别是光学字符识别技术的一个重要分支,其在邮政编码、财务报表、银行票据等方面的广泛应用,一直是图像处理、模式识别领域的研究热点。随着社会的发展,各国经济交流日益加深,人们每天要处理大量的票据,所以手写数字在这个领域是必不可少的,例如人们要处理许多支票、发票、货物单等数据,这些都要大量与数据打交道,手写数字识别

是各国研究热门之一,由于不同的人所写的数字体形态各异,千差万别,有时候书写不规范,所以手写数字识别是极其复杂的。

文中介绍基于图像识别的手写数字识别技术,该方法对手写数字识别率高,实时性高,且要求环境比较低。

1 系统实现的流程图

本系统要对手写数字进行准确地识别,该数字识别系统分为四大模块,具体的模块如图1所示。

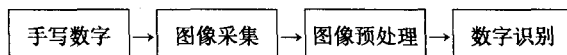


图1 系统模块

其中图像预处理是关键步骤,它的精确度的高低影响数字识别率,下面将分别介绍该系统模块。

1.1 手写体数字来源

如今市场经济日益发达,数据是人们每天不可缺少的处理工作,例如发票、支票、汇款单、成绩单等这些

收稿日期:2011-04-21;修回日期:2011-07-26

基金项目:安徽省高校省级重点自然科学基金(kj2010A023);安徽大学青年科学研究基金(2009QN0019B)

作者简介:吴忠(1986-),男,硕士,研究方向为模式识别;吴建国,博士,教授,研究方向为中文信息处理、智能EDA。

都需要处理数据,所以手写数字识别越来越引起各国重视,这方面目前也是各国研究热点之一。一般情况下,用户书写的数字限定在规定的区域内,这样有利于对手写体数字进行图像采集。

1.2 数字图像采集

用 CCD 摄像头对该区域内的手写体数字进行扫描,采集图像,得到的图像为二维 32 位真彩色图像,以 JPG 格式储存在计算机的硬盘中,下面选择采集到的待识别手写数字部分图像如图 2 所示。



图 2 手写体数字部分图像

由于不同的人所写的数字体形态各异,千差万别,有时候书写不规范等问题,或者是在用 CCD 摄像头进行扫描,采集图像时由于外界影响得到不规则、不清晰手写数字图像。许多因素的影响就为识别增加了难度。为了解决这些问题,需要对采集到的数字图像进行预处理,这样才能提高数字识别率^[1]。

1.3 数字图像的预处理

预处理包括去噪、二值化、细化、图像倾斜纠正、图像归一化等操作将图像带来的外界干扰物排除,得到清晰的图像,这样能够更好地进行手写数字识别操作,所以图像预处理至关重要,直接影响数字识别效果。

1.3.1 图像的二值化

图像的二值化是图像分割方法中的一种,它是通过设定某些阈值来实现,尤其对于背景和目标占据不同灰度级范围的图像,因为在进行数字识别时没有必要对手写数字的颜色、灰度级等其它信息进行识别,为了提高识别效率、删除不必要信息,所以手写数字识别过程中要进行二值化。图像的二值化关键技术就是对阈值的设定。关键要准确设置相应的阈值,然后再设置大于这个阈值的像素为灰度极大值,小于这个阈值的像素称为灰度极小值。目前有许多种二值化算法,主要可分为全局阈值法和局部阈值法^[2]。

1.3.2 图像的细化

图像的细化(Image Thinning)是指二值图像的骨架化(Image Skeletonization)的一种操作运算。一个图像的骨架是由一些线和曲线构成,骨架可以提供一个图像目标的形状和尺寸信息,在数字图像分析中占有重要的地位,图像细化是进行图像识别、线条类图像目标分析的重要手段。细化的算法分为非迭代算法和迭代算法,其中迭代算法根据其检查像素的范围可以再分为串行算法和并行算法^[3]。

1.3.3 图像归一化

手写数字的主要特征就是其书写过程中的不规范

性。例如:同一个数字,不同的人书写,会呈现不同的形状,即使是同一个人,有时候外界和自身的因素影响,也会出现不同的形状。这样使书写出来的数字风格会千变万化,还有采集的数字图像大小不一,为了保证有效地提取手写数字特征,减少这种由大小不一带来的影响,所以要对采集的数字图像进行归一化。图像归一化就是通过一系列变换(即利用图像的不变矩阵寻找一组参数使其能够消除其他变换函数对图像变换的影响),将待处理的原始图像转换成相应的唯一标准形式(该标准形式图像对平移、旋转、缩放等仿射变换具有不变特性)。图像归一化步骤分为四个步骤即坐标中心化、x-shearing 归一化、缩放归一化和旋转归一化^[8]。

1.3.4 图像倾斜校正

由于个人原因、外界影响,有时候得到的图像有一定的倾斜度,这样对于识别具有一定的影响,所以再数字识别之前要进行图像校正,目前有许多方法,如:行间距跨度方法、投影特性法和傅里叶变换法等。

1.4 手写数字识别

前面所做的工作都是为了能够对手写体数字识别服务的,当数字图像预处理后,首先要进行抽取该数字区别其它数字关键的特征,这一步骤是手写数字识别关键部分,做好这一步是非常重要的,然后运用字符识别算法对其识别,最后确认出来。目前数字识别算法有模板匹配、神经网络、BP 神经网络等算法,每种算法都有各自的优点和缺点,要根据实际情况选择这些算法。

2 系统实现过程

传统的方法是利用光学的光电变换原理对数字进行识别,文中基于图像识别技术对手写体数字进行识别,下面将具体介绍其实现过程。

2.1 系统安全设置

为了保护系统安全,所以要对系统设置初始化用户名和密码,用来管理所有用户,包括特别用户和普通用户,其中特别用户即是用户管理员,具有管理普通用户权利,同时还具有删除用户、添加用户,同时给特别用户分配系统使用权利和密码的修改。

2.2 手写体数字实现原理

字符识别方法目前有多种、例如统计识别、句法结构、神经网络、人工智能等,文中采用统计识别算法对手写体数字进行识别,统计识别算法是指通过统计得到待识别模式的一组统计特征,然后按照一定的准则设计决策函数进行分类判别。该方法的流程图如图 3 所示,首先提取手写体数字特征,然后对手写体数字进行分类判别,再经过训练和识别这两个阶段,最终识别

出该数字。数字训练阶段过程,文中所采用的样本数据库是国际上公认的 CEN PA RM I (Centre for Pattern Recognition and Machine Intelligence of Concordia University of Canada) 书写体数据库^[5],以便对待识别数字进行分类。识别阶段,根据计算决策函数,将待识别手写体数字的特征与在训练阶段中所建立的标准样本特征进行比较,最后通过计算数字最大相似度来判别该手写体数字的类别^[6]。

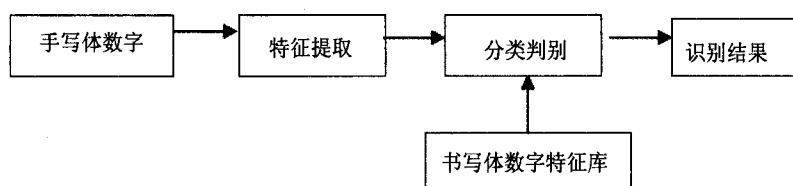


图3 统计方法流程图

2.3 手写数字特征描述

由于不同的人所写的数字体形态各异、千差万别,或者书写不规范,甚至同一个人写出的数字也不一定相同,所以要进行手写体数字识别,关键是要知道该数字与其它数字区别的特征,下面主要介绍一下数字的一些主要特征。

2.3.1 端点特征

数字的点阵特征包含着数字特征大量信息,数字是由许多点进行不同的排列构成的,这些点反应在书写数字时的笔画特征即书写过程中起始位置和结束位置。

2.3.2 数字轮廓特征

数字的轮廓特征是指将数字的轮廓划分为特征片段:凸弧、凹弧、直线段、端点、洞,并由这些特征片段得到特征基元,这样构成了对数字结构完整的描述^[7]。

2.3.3 数字骨架特征

数字的骨架特征是指数字骨架端点和交叉点等结构特征,如端点的个数、交叉点的个数、线条弯曲的方向等,提取这些关键特征,然后再根据这些特征的位置特征进行数字识别^[8]。

2.3.4 数字交叉特征

横向或纵向扫描数字,其中当一行或列的像素由白变黑的变化就是该行或列的横纵向交叉特征。文中主要介绍横竖两个方向的距离不一的六条直线作用于数字,并计算出横或纵向与数字的交叉数。

2.4 数字图像的特征提取

特征提取指通过映射或变换的方法,把模式空间的高维特征变成模式空间的低维特征,提取数字图像中有关拓扑结构、笔划分布等方面的特征信息,这个过程称为特征提取。

特征提取算法可分为粗网格特征提取算法、笔划密度特征提取算法、混合特征提取算法。其中,粗网格

特征提取算法对于高点阵图像数据按比例压缩非常有效,因为这种特征对高密度图按比例压缩后,图像的相对位置和位移变化较少,因此该算法也可用于某些图像细化处理,此外,该算法对抗噪声干扰能力较强,缺点就是对图像位移变化等抗干扰能力较弱。笔划密度特征提取算法对字型畸变和位移变化有了很好的抗干扰能力,该算法对于不同人的书写风格、字型的变化和

位置差异较大非常适合,但该算法有一个弱点就是对于含有噪音的点阵图像不好处理。混合特征提取算法就是综合以上两种算法的优点,减少它们的缺点。选择好特征提取算法对提高数字识别率至关重要^[8]。

2.5 手写数字识别过程

对手写数字识别,要进行两个方面识别:第一,识别在该区域有没有写数字;对于这方面的识别比较容易实现,例如:汇款单上有没有填数字,只要比较一下表格图像与已知填有信息的表格图像均方差大小,就可以识别是否有数字。第二,对填写的手写体数字进行识别,这是文中最主要讨论的部分。

论文在实现对手写体数字的识别用到的算法是采用模板匹配和神经网络相结合来实现的,根据数字的特征将神经网络分为三个子网即数字前部、数字中部和数字后部。这样的划分就可以把数字的特征形象地表述出来,将待识别的数字与国际公认的 CEN PA RM I 手写体数据库中的样本进行匹配,具体的是通过计算数字与模板的相似度来识别手写体数字。

手写体数字识别具体步骤是:

(1)把数字分为前部、中部、后部三个部分,这样更好地提取数字特征;

(2)根据模板匹配算法的判别函数来计算数字与样本库的样本相似度来识别该数字^[5]。

在介绍判别函数之前,首先要知道一些定义:设 S_1, S_2, \dots, S_m 为 m 个数字类别,假定每个类对应一个标准样本,那么一共有 m 个标准样本,用 k_1, k_2, \dots, k_m 来表示。假设任意一数字特征向量 X 和第 i 个 ($i = 1, 2, \dots, m$) 数字样本间的“相似度”为 R_i ,然后计算出待识别手写体数字的特征向量 X 和类标准样本 k_j 之间的“相似度” R_j ^[9],对所有的 j 不等于 i ,若 $R_i > R_j$,则 X 就属于 S_i 类符号,下面介绍一下计算的具体过程。

首先设置数字的特征向量,根据前面的特征提取与分析,该数字的特征向量为一个 19 维向量即 $X = \{x_1, x_2, x_3, \dots, x_{19}\}$,其代表的含义分别为: x_1 :数字端点; x_2 :数字一般连接点; x_3 :数字交叉点; x_4 至 x_{11} 为八条水平线与数字的交叉次数; x_{12} 至 x_{19} 为八条竖直

线与数字的交叉次数; G 为标准样本库的特征向量^[10]。

该模板匹配算法核心思想就是通过度量待识别手写体数字和样本库中标准样本库的相似度,来分类确定待识别数字。设待识别数字特征向量 X 和标准样本库中第 i 个样本的特征向量 G 相似度为 $R(X, G)$ 。 $R(X, G)$ 定义如下:

$$R(X, G) = \frac{\sum_{i=1}^m x_i \times g_{ik}}{\sqrt{\left[\sum_{i=1}^m x_i^2 \times \sum_{k=1}^m g_{ik}^2 \right]}} = \cos \alpha$$

其中在上面的公式里^[11]: x_i 为待识别手写体数字特征向量 X 的第 i 个($i=1, 2, \dots, 19$)分量; g_{ik} 为数字标准样本库中第 i 个样本中的第 k 个分量; m 为数字标准样本类别数。公式中的分子为数字特征向量 X, G 之间的积,分母为数字特征向量 X, G 的模, α 是数字特征向量 X, G 在 m 维空间中的夹角。当 X, G 两个数字特征向量完全相同时,它们之间的夹角大小为0,那么 $R(X, G)=1$,可以得出它们间的距离 $D(X, G)=0$,表明它们的相似度最大,最后计算最大相似度 $R(X, G)$,假如 $R(X, G) \geq$ 给定阈值,这样可以找到与待识别数字最接近的样本类别^[12],最终确认出待识别的手写数字。

3 实验的结论与分析

实验采用 CCD 摄像头采集汇款单的手写体图像,图像预处理、关键特征提取、数字识别等过程,判定汇款单表格中有无数字、该数字是多少,最后识别出这些数字。本实验采用 100 张汇款单图像,对其进行试验,实验结果发现识别错误的手写体数字主要是 1, 7, 0, 6, 4, 原因在于 1 和 7, 0 和 6, 6 和 4 在书写结构方面的相似性非常高,而 2, 3, 5, 8, 9 的识别率达到 99.8%。

(上接第 47 页)

- 略研究[J]. 电子与信息学报, 2006, 28(12): 2382-2385.
- [4] 张登银, 李正, 程春玲. 基于 SLA 的下一代网络计费方法[J]. 计算机应用研究, 2009, 26(7): 2713-2716.
- [5] 肖红叶. 高级微观经济学[M]. 北京: 中国金融出版社, 2003.
- [6] Zhang Yong, Feng Wenying, Hurley R. Integration of QoS Queuing Schedules to QoS Caching Schemes[C]// Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science. [s. l.]: IEEE/ACIS, 2006: 1-6.
- [7] Ibrahim H A, Nossier B M, Darwish M G. Billing System for Internet Service Provider (ISP) [C]//MELECON' 2002. Mediterranean: [s. n.], 2002: 260-268.

4 结束语

文中基于图像识别技术对手写体数字进行识别,相对于传统的光学的光电变换原理对数字进行识别,在识别率和实时性方面得到了有效的提高。但该系统还需要完善的是手写体数字特征提取,这样更进一步提高手写体数字识别率。

参考文献:

- [1] 龚声蓉. 数字图像处理与分析[M]. 北京: 清华大学出版社, 2006.
- [2] 田村秀行. 计算机图像处理(日)[M]. 北京: 科学出版社, 2004.
- [3] 张婷. 基于图像识别技术的光学标记阅读机的研究与应用[D]. 合肥: 安徽大学, 2007.
- [4] 王虎. 基于图像识别的标记阅读机及选举计票系统研究[D]. 合肥: 安徽大学, 2006.
- [5] 刘克. 基于轮廓分段特征的手写阿拉伯数字识别[J]. 计算机学报, 1996(10): 1066-1073.
- [6] 孙家广. 计算机图形学[M]. 北京: 清华大学出版社, 1996.
- [7] 张站. 基于图像识别的阅卷系统的设计与实现[D]. 合肥: 安徽大学, 2008.
- [8] 于振华. 基于骨架结构特征的手写数字识别方法[J]. 模式识别, 2010(6): 95-98.
- [9] 杜敏. 基于混合特征提取的手写体数字识别方法的研究[J]. 西安交通大学学报, 1996(9): 94-98.
- [10] Mahmoud S A. Arabic Character Recognition Using Fourier Descriptors and Character Contour Encoding [J]. Pattern Recognition, 1994, 27: 815-824.
- [11] Pedrycz. Fuzzy Sets in Pattern Recognition: Methodology and Methods[J]. Pattern Recognition, 1990, 23: 121-146.
- [12] Burel D G, Potteier I, Catros D J Y. Recognition of Hand-Written Numerals[J]. IEEE Trans. Syst. Man Cybernet, 1990, 20: 835-847.
- [8] Davoyan R, Altmann J. Real-time market model for pricing differentiated services [C]//Fourth International Conference on Networking and Services. [s. l.]: [s. n.], 2008: 134-140.
- [9] 王平, 李信满, 赵宏, 等. 基于服务质量的网络计费的研究[J]. 小型微型计算机系统, 2002, 23(1): 74-77.
- [10] 李晓花, 李小军. IP 网络 QoS 几种排队算法的对比分析[J]. 四川理工学院院报(自然科学版), 2006, 19(3): 81-84.
- [11] 何芳, 邱雪松, 成璐. 面向业务的下一代互联网计费系统体系模型[J]. 北京邮电大学学报, 2004, 27: 54-58.
- [12] 马莉. MATLAB 语言实用教程[M]. 北京: 清华大学出版社, 2010.