

采用多级分类器的手写数字识别技术研究

赵 锋¹ 薛惠锋¹ 王 伟² 吴慧欣¹

(1 西北工业大学 自动化学院, 陕西 西安 710072)

(2 西北工业大学 航空学院, 陕西 西安 710072)

摘 要: 在分析目前手写数字识别策略基础上, 提出了采用多级分类器的手写数字识别技术, 并给出了多级分类器识别系统体系结构。首先对手写数字样本进行预处理, 然后针对不同的子分类器, 提取不同的分类特征。系统中上一级的拒识样本作为下一级的输入样本, 因此系统识别率逐级增加。最后以面向对象语言 Visual C++和 MatLab 为工具, 开发了多级分类器实验原形系统, 结果表明, 该识别技术不仅操作简单, 而且具有较好的识别性能。

关键词: 分类器, 特征抽取, 二值化, 数字识别

中图分类号: TP182

文献标识码: A

文章编号: 1000-7180(2006)12-0030-03

Handwritten Digit Recognition Technology Based on Multistage Classifying Device

ZHAO Feng¹, XUE Hui-feng¹, WANG Wei², WU Hui-xin¹

(1 College of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

(2 College of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: On the basis of analyzing present Handwritten Numeral Recognition, a multistage classifying device is proposed and a multistage Classifying device system structure is given. Two procedure are included: Preconditioning on Handwritten Digital samples; Drawing different classify characters in light of different subclassifying device. The identifying ratio increase by level number as the refuse identify samples of the last level are the input of the next level. Finally, a multistage classifying device experimental system is developed based on Visual C++ and Matlab. The result demonstrates that the identifying technology. Operates simply and has better performance as well.

Key words: Classifying device, Feature choice, Two values, Handwritten numeral recognition

1 引言

手写数字识别 (Handwritten Numeral Recognition) 是光学字符识别技术 OCR (Optical Character Recognition) 的一个分支, 它研究的对象是: 如何利用计算机自动辨认写在纸张上的阿拉伯数字。手写数字识别涉及到模式识别、图像处理和人工智能等技术, 它在邮政、银行、金融以及模式识别理论研究等众多领域有着广泛的应用前景。

目前, 传统的手写数字识别一般分为联机 (在线) 识别和脱机 (离线) 识别。联机识别是从坐标点列中抽取笔划等字符特征进行识别, 该方法虽然易于抽取字符的笔划特征, 容易识别书写的顺序和笔尖的走向, 但与书写顺序起点和终点有关, 难以识别有“补笔划”或有“自由始终点”的字符。脱机识别

将书写字符转化为点阵或二值图像再抽取特征进行识别, 该方法与书写顺序无关, 较易识别具有“补笔划”或有“自由始终点”的字符, 但不易抽取象笔划特征这样分类能力强的特征。本文在深入研究上述识别技术的基础上, 提出了基于多级分类器的手写数字识别技术, 并以 Visual C++和 MatLab 为工具开发了一套四级分类器手写数字识别试验系统。最后结合给定的样本, 进行了试验分析。

2 手写数字识别预处理

预处理是手写数字识别的重要一环, 它可以把原始图像转换成识别器所能接受的形式 (二值化), 消除一些与类别无关的因素 (尺寸和位置的归一化)。手写数字识别的预处理通常包括数字图像的提取, 数字图像的二值化处理, 细化处理, 特征值提取等, 在不同的识别方法中, 对预处理的项目和要

收稿日期: 2005-11-21

基金项目: 霍英东教育基金项目 (81044)

求也各不相同。

2.1 数字图像的提取

数字图像的提取是从图像采集中所得到的整幅图像中分出需要识别的数字图像区和无用的背景图像区。如邮政编码中方框之内的区域(不包括方框)和方框及方框之外的区域。数字图像的提取主要是根据应用系统的应用环境所具有的先验知识进行识别处理。

2.2 数字图像二值化

把数字灰度图像处理成二值(0, 1)图像的过程,称为对数字图像的二值化(Binarization)。对灰度图像二值化能显著的减小数据存储的容量,降低后续处理的复杂性。

设扫描、A/D 转换后的数字图像点阵为:

$$C=(f(i, j)) \quad i=1, 2, \dots, p, j=1, 2, \dots, q \quad (1)$$

式(1)中 $f(i, j)$ 是像素 (i, j) 的灰度值,有白笔画的部分, $f(i, j)$ 大,而黑背景部分, $f(i, j)$ 则小。最简单的二值化通过设定固定灰度值阈值 T 完成,即当 $f(i, j)$ 值比预定阈值 T 小,看成背景,以“0”标志;否则看成数字笔画,以“1”标志,如式(2)所示。

$$g(i, j)=\begin{cases} 1 & f(i, j) > T \\ 0 & f(i, j) \leq T \end{cases} \quad i=1, 2, \dots, p, j=1, 2, \dots, q \quad (2)$$

由式(2), $B=g(i, j)$ 即为字符二值化后的图像点阵。

对数字图像二值化,要求二值化后的图像能够忠实地再现原数字。基本要求为:(1)笔画中不出现空白点;(2)二值化的笔画基本保持原来文字的结构特征。针对数字图像较为简单、背景和图像有较明显的区别的特点,本文采用了最大方差设定法进行二值化,这样既考虑速度又兼顾二值化的效果。

2.3 细化处理

将二值化的字符图像抽取出字符框架的处理过程称之为细化处理。细化算法基于对每个像素点是字符边界点框架点的判定。无论对字符每一圈扫描后的边界点去除,还是全部标记后的一次性边界点去除,其主要差别在于框架点配模式的选择。

2.4 特征值提取

特征值提取的目的是从分析数字的拓扑结构入手,把它的某些结构特征提取出来,使数字的位移、大小变化、字形畸变等干扰相对减小,而把那些反映数字特征的关键信息提取出来。细化后的数字字符基本上保持了手写数字字符的形状,由于同一数字字符的形状数不胜数,因此并不能直接根据数字字符形状进行识别,特征值提取就是将数字字符形

状转换为的一组特征值的过程,为了有效地进行数字识别,特征值提取是必要的。

综上所述,在预处理过程中,除数字图像提取这一步骤丢失的是无用信息外,后面三步处理中丢失的信息都可能是后继识别中的有用信息。

3 多级分类器识别原理与系统设计

3.1 多级分类器识别原理

联机识别易于抽取字符笔划特征,对书写的顺序和笔尖的走向容易判别,但与书写顺序起点和终点有关,难于识别“补比划”的情况。脱机识别与书写顺序无关较易识别“补笔划”字符,但是不易抽象笔划特征这样分类能力强的特征。可见,不论是联机识别方法还是脱机识别方法都有其固有的困难,使得各自的识别效果都很难达到实用的要求。鉴于此,本文将联机识别技术和脱机识别技术有机结合起来,设计出的多级分类器识别系统如图1所示。

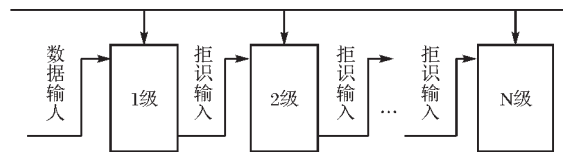


图1 多级识别系统结构

该多级识别系统将不同的识别技术运用于不同特点的字符,充分利用不同识别方法的优点,从而在性能上优于单级识别系统。系统中下一级的输入样本是上一级的拒识样本,这样正确识别率将从第一级开始逐渐增加,如果各级子系统相互独立,则整个系统的识别率将显著增加。

假设一个 N 级系统中第 i 级的识别率为 r_i ,误识率为 m_i ,拒识率为 α_i ,则整个系统的识别率如下:

$$\begin{aligned} R_n &= r_1 + r_2(1 - r_1 - m_1) + r_3(1 - r_1 - m_1)(1 - r_2 - m_2) + \dots \\ &\quad + r_n(1 - r_1 - m_1) \dots (1 - r_{n-1} - m_{n-1}) \\ &= \sum_{i=1}^{n-1} \left(r_i \prod_{j=1}^{i-1} \alpha_j \right) \end{aligned} \quad (3)$$

其中, $\alpha_0=1, \alpha_j=1 - r_j - m_j$ 。用数学归纳法不难证明,误识率和拒识率分别如下。

$$M_n = \sum_{i=1}^n m_i \prod_{j=1}^{i-1} \alpha_j \quad (4)$$

$$\alpha_n = \prod_{j=1}^{n-1} \alpha_j \quad (5)$$

在式(3) R_n 计算式中,如果 m_i 均为0,则:

$$R_n = \prod_{j=1}^n r_j \tag{6}$$

如果各级系统相互独立, 则 $r_n \rightarrow 0 (n \rightarrow \infty)$, 从而 $R_n \rightarrow 1 (n \rightarrow \infty)$ 。可见, 提高 N 级识别系统的识别率的关键是: (1) 在各级子系统中选取相互独立的特征, 使各子系统相互独立, 以减少拒识率; (2) 在各子系统内部进行各种处理时, 以降低误识率为首要目标。

因此, 在设计多级识别系统时, 重点是特征的提取。在靠前的子系统中要尽量减少误识率, 因此选取的特征要更注重其分类能力, 而不是稳定性 (即当同一字符发生微小变形时特征保持不变或变化很小); 但在靠后的子系统特别是最后一级子系统中则要提高识别率, 减少误识率和拒识率, 因此其特征选取要更注重稳定性。而各级的特征选取应尽量使得对字符的描述相互补充。

3.2 多级分类器系统设计

根据上述多级分类器识别原理, 设计手写数字多级 (四级) 识别系统如图 2 所示。系统由四个子系统组成: (1) 联机识别模块 I: 选取分类能力较强的字符结构特征用神经网络分类器对样本进行一级识别; (2) 脱机识别模块 I: 通过字符转换, 将坐标字符转化为二值图象, 选取分类能力较强的结构特征用神经网络分类器对样本进行二级识别; (3) 联机识别模块 II: 选取稳定性较大的字符结构特征对样本进行三级识别; (4) 脱机识别模块 II: 对字符二值图象, 选取稳定性较大的字符结构特征对样本进行四级识别。

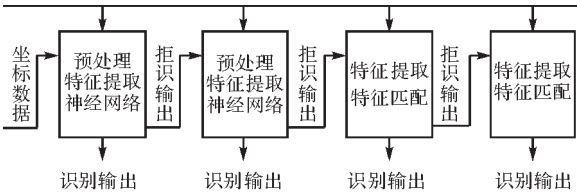


图2 手写数字四级识别系统

联机识别模块 I 采用带有一个隐层的三层神经网络分类器进行识别。网络之间连接隐元的传输函数为高斯函数 (即正态分布函数), 网络输出节点的传输函数为 Sigmoid 函数, 选取笔划特征和笔划间的位置关系特征作为网络的输入特征。脱机识别模块 I 采用与联机识别模块相同的神经网络分类器, 以样本的轮廓特征和点特征为网络输入。联机识别模块 II 主要功能用于对输入样本进行粗分类, 选择笔划方向码、相邻笔划位置关系码、旋转方向码等

为分类特征。脱机识别模块 II 同样用于粗分类, 采用传统的脱机识别方法, 一般选择字符的凹特征和中线特征为进行识别匹配。

上述四级分类器手写数字识别系统将不同的识别技术运用于不同特点输入样本, 充分利用不同识别方法的优点。系统中的每一级分类器具有三种输入输出接口 (样本输入、识别输出、拒识输出)。上一级分类器的拒识输出为下一级分类器的样本输入, 系统的总输出为每级分类器识别输出的加和。

4 试验分析

本文设计的四级分类器手写数字识别系统可进行分级独立工作。为了进行试验对比, 这里采用自建的数据库样本进行实验。首先运用在线识别模块, 在手写数字输入框内手写输入 1000 个数字样本 (为进行脱机识别, 把手写输入的数字样本保存为 BMP 文件) 进行实验; 然后, 读入手写数字样本 BMP 文件, 对同一样本进行脱机模块识别; 最后对同一样本进行多级分类器识别。试验结果如表 1 所示。

表 1 试验结果分析

分类器模块	识别率	误识率	拒识率
多级分类识别	97.5%	1.5%	1%
在线模块识别	86.4%	11.2%	2.4%
脱机模块识别	88.7%	10.1%	1.2%

表 1 中对比试验数据表明, 不论是联机识别技术还是脱机识别技术都有其固有的难处, 识别效果都很难达到实用的要求。而针对于二者相结合的多级分类器识别系统则具有较好的识别效果。

5 结束语

运用基于多级分类器的手写数字识别原理构建多级分类器识别系统体系结构, 充分利用了在线识别和离线识别各自的技术优点, 解决了单级识别器不能有效地提取分类特征和识别率较低的问题, 而且基于该方法建立的识别原形系统具有以下特征: (1) 数字图像预处理过程中, 在进行二值化之前采用一定的事前处理算法, 提高了输入样本的质量; (2) 整个识别系统由多个单级系统依据一定的耦合关系级连而成, 下一级的输入样本是上一级的拒识样本, 大大提高了系统识别率。实践表明, 该方法能够有效地克服不同手写风格所带来的干扰, 具

(下转第 37 页)

at most 0.03ns. When the cell works in a 500MHz register file, the ratio is only 1.5% of the cycle time and is accepted for a 500MHz register file. In this case, the cutoff capacity effect to bit-line voltage is at most 0.04V, which is small enough to be ignored.

When bit-lines output high voltage, the difference changes a lot. The worst case happens in the case of one port reading. Using method 1, bit-line in high voltage is discharged less than 1.4V. There will be static current in the drivers next to bit-lines or a wrong content will be output. The result from reference [4] has discharged about 0.02V. And method 2 doesn't discharge. Method 2 has better reliability and avoids the unnecessary power consumption in the drivers. When bit-lines are discharged to low voltage, the two methods are both faster than result of reference [4]. They speed up 0.05ns, which is about 12.5% of discharge time.

In a word, method 2 is a good choice for a 16-port write-through memory cell in a high performance register file. Besides the high access speed and access reliability, it will save 7 transistors area and power consumption.

5 Conclusion

A new 16-port write-through cell is introduced in this article, every port of which can be accessed individually. The proposed method cuts down some transistors in the cell and achieve the like performance. In addition, the long rail plan will reduce the

crosstalk noise in layout. Customized in 1.8V 0.18 μm CMOS technology, it is designed to achieve high performance with consideration of reliability, compactness and power consumption. When it works in 500MHz, the register file has a data bandwidth of 32GB/s. In a word, the memory cell is a useful and effective method for a custom register file working in high-performance micro-processors.

Reference:

- [1] John L Hennessy, David A Patterson. Computer architecture a quantitative approach. Third Edition, China Machine Press, 2003, 390~392
- [2] Jan M Rabaey. Digital integrated circuits a design perspective. Prentice-Hall International, Inc. 1999: 578~590
- [3] Feipei Lai, Ying Lang Chuang, Shyh Jong Chen. A new design methodology for multiport SRAM cell. IEEE Trans. on Circuits and Systems, 1994, 41(11): 677~685
- [4] Nestoras Tzartzanis, William W Walker, Hoa Nguyen. A 34word \times 64b write-through self-timed dual-supply-voltage register file. ISSCC, 2002: 416~417, 478, 409

作者简介:

于倩女, (1978-), 博士研究生。研究方向为 VLSI 集成电路设计。

王东辉 男, (1973-), 副研究员。研究方向为 VLSI 集成电路设计。

张铁军 男, (1974-), 副研究员。研究方向为 VLSI 集成电路设计。

侯朝焕 男, (1936-), 中国科学院院士。研究方向为 VLSI 集成电路设计。

(上接第 32 页)

有较好的应用前景。

参考文献:

- [1] 曾卫明, 等. 基于改进 BP 算法的数字字符识别[J]. 微电子学与计算机, 2002, 20(4): 4~6
- [2] 陆从德, 张太镒, 等. 基于支持向量域描述的学习分类器[J]. 微电子学与计算机, 2005, 11(11): 75~78
- [3] Nadal C, Suen C Y. Applying human knowledge to improve machine recognition of confusing handwritten numerals. Pattern Recognit, 1993, 26(3): 376~388
- [4] Gader P, Foreater B, Ganzberger M, et al. Recognition of handwritten digits using template and model matching.

pattern recognit, 1991, 24(5): 418~430

- [5] Fujisawa H, Naknamo Y, Kurina K. Segmentation of automatically located handwritten numeric strings. Frontiers in Handwriting recognition, 1992: 47~59
- [6] Richard G Casey, Eric Lecolinet. A Survey of methods and strategies in character segmentation. IEEE Trans on Pattern Analysis and Machine Intelligence, 1998, 18(7)

赵锋 女, (1981-), 硕士研究生。研究方向为模式识别、复杂系统建模与仿真、数据挖掘。

薛惠锋 男, (1964-), 教授, 博士生导师。研究方向为模式识别、复杂系统建模与仿真、数据挖掘、信息化系统工程。