

基于图像处理的盲文自动识别系统研究

Study on image processing based braille automatic identification system

李念峰, 董迎红, 肖志国

LI Nian-feng, DONG Ying-hong, XIAO Zhi-guo

(长春大学 计算机科学技术学院, 长春 130022)

摘要: 未经培训的明眼人无法阅读盲文资料, 从而很多盲儿家长在教育辅导孩子时存在困难, 另外, 从事盲人教育的教师也多数不懂盲文, 在批阅盲生试卷、作业时不得不借助盲文翻译者的翻译, 造成评价过程引入了不客观因素。本文主要研究基于图像处理盲文的识别技术, 把盲文信息扫描为图像, 借助图像处理技术及机器翻译手段, 把盲文翻译成正常汉语文字, 为盲人更好的融入社会服务。实验证明, 本文所涉及的方法能有效、准确的提取并识别出盲文, 并把盲文图像转换成汉字。项目的成果将具有比较好的技术转化和应用前景, 为视障人群学习、生活提供帮助。

关键词: 图像处理; 盲文识别; 机器翻译

中图分类号: TP274

文献标识码: A

文章编号: 1009-0134(2012)02(上)-0063-04

Doi: 10.3969/j.issn.1009-0134.2012.2(上).24

0 引言

使盲人接受良好教育是当今世界各国都必须面临的社会问题之一。全国现共有特殊教育学校1640所, 从事盲人教育的教师, 除专业教授盲文认读的外, 多数不懂盲文。对盲生学习成绩的考核, 目前的一般做法是: 任课教师出汉语试卷→由专职人员翻译成盲文并打印试卷→盲生笔试→专职人员完成盲文试卷的翻译并读给阅卷教师→阅卷教师给出成绩。环节繁琐、耗时费力且易引入不客观因素。针对以上问题, 开展了基于图像处理的盲文识别关键技术及系统的研究。研究的成果可使从事视障教育的教师能够便捷的认读视障学生的作业、试卷等盲文学习材料, 为视障者创造更加舒适的学习环境。另外, 研究成果也可应用于多种便携式电子设备, 在国家倡导的残疾人无障碍公共环境中, 为视障人群中不能够顺利阅读盲文者, 提供帮助; 同时, 研究成果也可以为盲生监护人所用, 促进监护人与盲人儿童之间的交流沟通, 更好的对孩子进行学业上的辅导。

1 研究的主要内容及国内外研究现状

1.1 项目主要研究内容

本项目重点研究盲文图像的检测提取与识别

方法, 主要包括盲文图像的采集及预处理, 盲文图像的点字特征提取与识别、机器翻译算法等。

1.1.1 盲文图像的采集及预处理方法

经过对扫描仪、手机所拍摄的盲文图像的研究发现, 盲文识别对于前期盲文的图像摄取的要求非常高, 而且采集时受到的外界干扰越小, 识别率就越高。采集盲文图像后, 为达到图像识别的要求, 需要对图像进行校正、分割、去噪等相关预处理操作。在校正算法、最佳阈值获取、去噪方面要进行试验对比分析, 结合处理速度要求, 选取最佳算法。

1.1.2 盲方识别

经过预处理的图像, 还需要自动准确定位每个“方”的位置, 自动定位算法是整个系统实现的难点和关键。只有准确定位了每个“方”位置, 才能提取出盲文点, 并进行编码, 并保存编码文件, 再通过查表, 输出每个盲符所代表的意义(拼音字母、声调、标点、数字等)。

1.1.3 机器翻译

用计算机实现盲文到汉语的翻译过程, 实质上是“盲文→拼音→汉字”的过程。由于在汉语中绝大多数音节都对对应许多个候选汉字, 欲在没有人工干预的情况下自动选出正确的候选字是

收稿日期: 2011-08-12

基金项目: 吉林省自然科学基金面上项目(201115135)

作者简介: 李念峰(1974-), 男, 山东德州人, 副教授, 博士, 主要从事计算机应用方面的研究工作。

比较困难的，这是盲文到汉字翻译的难点。将盲文双拼转化成全拼继而转换成汉字，实现盲文到汉字的自动转换是系统开发需要解决的又一个关键技术。

1.2 国内外研究现状

在国内，很多研究机构把盲人人机信息交互作为研究课题，并取得一定研究成果^[1-5]。清华大学智能技术与系统国家重点实验室已初步实现了盲文至汉字的转换和盲文识别理解，正在进行集语音识别、盲文码、扫描盲文等多媒体输入和语音合成、汉字或盲文输出等功能的盲人计算机信息交互系统的研制。由华建电子有限责任公司、中国科学院计算机语言信息工程研究中心、中国盲文出版社共同研制开发完成的“中国盲文计算机系统”，紧紧围绕语音输出、语音识别、键盘操作和盲文输出等方面进行研究，为盲人提供与计算机进行无障碍信息交流的通道。

国外一些发达国家中，较重视残疾人服务领域的研究开发，近年来国外的盲人人机信息交互研究工作已显露出要由实验室转向商品化的明显迹象^[6-8]。在欧美、拉丁等以字母为基础的文字中，盲文的每个“方”对应一个字母，几个方字的组合构成一个单词，这种对应关系比较简单。因此，国外对盲文的翻译研究在90年代曾有一个飞跃期，近年来与中国一样也把重点放在了盲人与计算机的交互技术上。葡萄牙学者就盲文图像识别方法、语义分析等进行了研究，设计了一种自动盲文代码转换系统^[6]；巴基斯坦的学者研究了盲文点字和乌尔都语言的转换问题，开发出一种联系上下文的可识别盲文到乌尔都语的翻译机^[7]。2006年日本东北大学研究生院副教授田中真美开发了一种安装有特殊指套的传感器。

从国内外的研究进展来看，1) 辅助盲人通过计算机与外界进行无障碍交流技术是世界性的研究热点；2) 盲文翻译对于以字母为基础文字的国家研究人员来说是比较简单的问题，因此仅在中国或日本等国在盲文到明文的机器翻译方面需要开展研究；3) 盲文图像信息提取已纳入研究者的视野，是盲文翻译的一个发展方向和趋势，但目前还没有基于图像识别的盲文到明文的自动翻译仪器设备问世；4) 可用于普通教师批阅盲生试卷的软件及仪器装置未见有相关产品。

2 基于图像处理的盲文自动识别系统

2.1 系统的体系结构

系统共分四个模块，结构及流程图如图1所示。

1) 盲文到汉字的机器翻译模块。本模块主要解决图像处理得到的双拼盲文信息到全拼音信息再到汉字语句的翻译问题，其中全拼到汉字语句级的翻译是难点。充分综合文献中给出的各种机器翻译算法及技巧并结合形成试卷过程中汉字转盲文时形成的分词库，在限定知识领域内完成对拼音到汉字的翻译可使问题解决简单化。

2) 基于ARM的盲文高速图像采集装置。目前市场上的高速扫描仪价格较高，采用ARM9以上芯片、嵌入式Linux操作系统自行开发盲卷图像采集装置。

3) 高速高准确率的图像识别模块。包括图像预处理子模块（对扫描图像时产生的噪音进行消除，同时自动辨别与弥补纸张划损、扭曲所带来的随机误差）和盲文信息理解子模块（在传统图像识别方法的基础上引入了基于上下文仲裁的方法以提高识别的容错性和准确率）。本模块的每一环节都经过充分优化，保证对盲文信息而言至关重要的识别准确率。

4) 图像数据的二次压缩模块。经过JPEG压缩技术处理，单个盲文扫描图像数据已然不大，但全部资料数据量之和仍然庞大，对网络传输和存储都有不小的压力。基于图像空间冗余度概念设计的答卷图像数据的二次压缩模块既保留了图像数据的原始性，又减轻了存储负担。

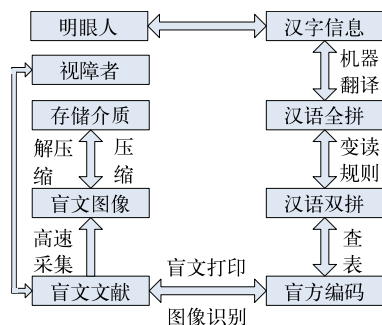


图1 系统结构及流程图

2.2 系统的关键技术及实现

2.2.1 盲文图像处理

1) 图像校正

图像采集的过程中，图像难免会出现失真现象，造成图像失真的原因有很多，如拍摄姿态、

扫描非线性等因素, 成像系统的象差、畸变、带宽有限等因素都可能会造成图像失真。本文在研究过程中所遇到的图像失真, 一方面是由于拍摄姿态造成的几何失真, 另一方面是由于光照不均而造成的图像失真。图像校正是指失真图像的复原性处理。其基本思路就是: 找出造成图像失真的原因, 提取相关信息建立相应的数学模型, 然后使用逆反过程恢复图像的原貌。

(1) 图像倾斜角度的检测

在进行盲文信息扫描前, 预先在每页盲文的空白位置, 以盲点的水平及垂直方向为基准, 标记一条黑色直线, 通过应用霍夫变换来检测直线位置的方法计算图像倾斜角。Hough 变换就是将 x - y 坐标平面的任意一条直线映射为 r - θ 坐标空间的一个点。因此, 对图像上所有的点做霍夫变换, 极坐标平面中直线相交最多的那个点则为所要求的直线在极坐标平面的映射点, 这样就在图像中检测出了直线的参数。具体检测步骤如下:

① 根据预先标记的直线参数信息, 确定目标区域;

② 二值化目标区域, 并在垂直方向对该区域进行边缘提取, 进行霍夫变换, 确定 r, θ 的最大和最小值;

③ 设计一个累加器数组 $A(r, \theta)$, 并设置所有值为零;

④ 对图像中的所有边缘点 (x_i, y_i) , 使 θ 在 $(0, \pi)$ 的范围内取值, 求 r 值, θ 每取一次值, $A(r, \theta)$ 加 1;

⑤ 对于所有的 $A(r, \theta)$, 寻找其最大值, 该最大值所对应的参数即为所求直线, 根据直线参数可求得倾斜角 θ 。

(2) 图像旋转

假设图像的像素点坐标为 (i, j) , 则 (i', j') 为对应点旋转后得到的新图像 $M'(i', j')$ 的像素坐标, 则图像旋转的计算矩阵可描述如下:

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix}$$

具体旋转步骤如下: ① 根据倾斜角及旋转矩阵计算新坐标 (i', j') ; ② 对 (i', j') 进行取整, 再根据取值范围, 给各个坐标值加某一值, 使得所有坐标值都在正整数的范围内, 根据坐标的最大值 N , 调整画布大小为 $N \times N$; ③ 根据对应坐标关系, 填充对应的像素值; ④ 对于空白点, 采用均值插值

法进行添补, 即用空白坐标的四周八连通域内像素值的平均值作为该空白处的像素值。

2) 图像分割与盲点检测

盲点检测提取是指把盲文图像中的凸点字从图像背景中提取出来, 系统采用的提取方法是图像分割法, 利用大津法求得最佳阈值进行合理分割。大津法 (Otsu's) 是一种自适应计算最佳阈值并将灰度图像转换为二值图像的简单有效的图像分割方法。假设图像 M , 记 T 为背景 B 与目标 O 的分割阈值, 图像 M 的灰度级总数为 L , 则在图像的灰度级范围内 $B = \{0, 1, 2, \dots, t\}$, $O = \{t+1, t+2, \dots, L-1\}$ 。假设图像 M 的总像素个数为 n , 则第 i 个灰度级的个数记为 n_i , 那么灰度级 i 的出现概率为:

$$P_i = \frac{n_i}{n} \quad (1)$$

则背景 B 和目标 O 各自包含的灰度级的出现率, 即背景和目标点数的各自的概率 w_0, w_1 为:

$$w_0 = \sum_{i=0}^t P_i \quad (2)$$

$$w_1 = \sum_{i=t+1}^{L-1} P_i \quad (3)$$

由式 (1)、(2)、(3) 可计算出背景和目标各自的均值为:

$$\mu_0 = \sum_{i=0}^t \frac{i P_i}{w_0} \quad (4)$$

$$\mu_1 = \sum_{i=t+1}^{L-1} \frac{i P_i}{w_1} \quad (5)$$

类间方差 σ_0^2 和总方差 σ_1^2 分别为:

$$\sigma_0^2 = w_0 (\mu_0 - \sum_{i=0}^{L-1} i P_i)^2 + w_1 (\mu_1 - \sum_{i=0}^{L-1} i P_i)^2 \quad (6)$$

$$\sigma_1^2 = \sum_{i=0}^{L-1} \left(i - \sum_{i=0}^{L-1} i P_i \right)^2 \quad (7)$$

则由式 (6)、(7) 可得最佳阈值 T :

$$T = \text{Arg} \left\{ \max_{0 \leq t \leq L-1} \left(\frac{\sigma_0^2}{\sigma_1^2} \right) \right\} \quad (8)$$

3) 盲方的检测定位与识别

盲方在分割后的图像中所对应的点阵的排列是有一定规律的, 而且每个盲方间的距离都符合一定标准 (方间盲点与盲点之间的距离即点距为 2.5mm, 方间的距离即方距是 3.5mm, 盲方的行间距离即行距为 5mm), 系统根据这一性质来进行盲方的提取。

通过计算盲点图像的质心, 把同一行、同一

列的盲点，通过其质心坐标连接起来，可据此自动定位每个盲方应存在的位置。对于同行（列）的每个盲点质心在提取时有可能出现横（纵）坐标不在一条直线上的情况，此时，采用统计所有横/纵坐标值，并以出现最多的一个值作为该行/列盲点质心的横坐标值进行替换对齐的办法使所有质心坐标值一致。同时，对于落在两线之间的孤立点（即在相邻两水平或垂直线之间，且与线上的点不接连），视为无用的点，不予考虑，亦可用背景像素替换予以清除。图2即为自动盲方检测定位示意。

定位盲方之后，从左到右、从上到下，在每个盲方内检测其六个交点处是否存在盲点，若有点用1表示，无点则用0标示，最终以一个二进制字符串的形式予以表示并按顺序保存为盲方编码文件，根据编码文件，按照一定规则，兼顾盲文变读规则，可完成二进制编码到双拼并到全拼的转换过程。

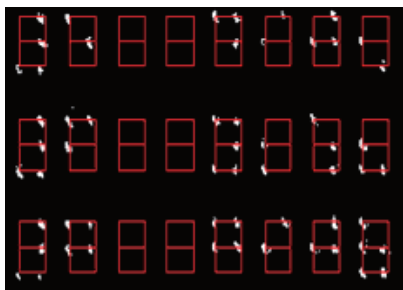


图2 盲方定位示意图

2.2.2 全拼到汉字的翻译

把盲文识别成全拼并不能方便批阅者进行阅读，还必须完成词或着整句的全拼到汉字的翻译。在汉语拼音到汉字的转换中，由于盲文扫描识别不可能达到100%正确，且汉字存在一字多音、一音多字和歧义现象，因此在将扫描盲文转换成汉字时需利用知识库、统计信息库和相应的理解后处理进行多层次的歧义校正和纠错处理。在汉语中由于绝大多数音节都应该有多个候选汉字，如要在没有人工干预下自动选出正确的候选字是比较困难的，这是盲汉翻译的难点。项目利用分词算法，把考试试卷及答案内容进行分词处理，形成自己的翻译对照词库，并结合搜狗拼音输入法的基本词库，对图像识别后得到的全拼文件进行翻译处理，取得了较好的单词及整句翻译效果。

3 实验结果及分析

盲文试卷中的题目类型大致有四种：选择（单选和多选）、填空、判断和简答。其中后三者涉及拼音到汉字的翻译问题，并且仅简答需要整句翻译，判断题及填空题只需完成单字及词的翻译即可。字和词的翻译，通过对自建词库的搜索即能完成95%以上的准确率的翻译。

整句翻译需要解决智能音节分割、智能重码选择、智能多音字识别等问题，本系统中，使用自建小样本词库结合搜狗基础词库的翻译方式，整句翻译效果能够满足试卷批阅的要求。为检验系统所用算法的有效性和合理性，从考生试卷中随机地选取了四份文本进行测试，测试结果如表1所示。

表1 实验数据

项目	文本1	文本2	文本3	文本4
总字数	1040	477	738	420
正确翻译字数	999	445	712	411
正确率	96.06%	94.13%	96.48%	97.86%

4 结论

由于研究主要针对盲人教育过程中教师批阅盲文的环节，经过实验发现扫描仪扫描速度快、受外界干扰相对较小，所以盲文图像的采集使用了自动进纸扫描仪；系统采用平均加权法对盲文图像进行灰度变换；在最佳阈值的选取过程中，针对双峰法、迭代法、大津法进行了算法实现与试验对比，最终使用大津法完成了最佳阈值选取。研究了盲点的检测、定位及盲方的自动定位识别算法。完成了盲文双拼到全拼的实现及全拼文本的单字、词组、句子的翻译，成功的开发了应用软件将图像盲文翻译成汉字。下一步研究工作：1）进一步优化各种算法，提高运行速度和算法效果；2）实现算法向便携式设备的移植。

参考文献：

- [1] 何川.国内信息无障碍的现状 & 展望[J].现代电信科技, 2007(3):4-8.
- [2] 江铭虎,朱小燕.基于多种知识的盲文翻译的研究[J].清华大学学报(自然科学版),2000,(9):60-73.
- [3] 包塔,朱小燕.盲汉转换系统的研究与实现[J].计算机工程, 2004,(10):45-47.
- [4] Jiang Minghu,Zhu Xiaoyan.Segmentation of Mandarin Braille Word and Braille Translation Based on Multi-

knowledge[J].IEEE,Proceedings of ICSP 2000,2000:2070-2074.

[5] 周春耕,张秉权.盲汉机器翻译系统的研究与实现[J].计算机工程与应用,2003,(4).

[6] Hamid Reza Shahbazkia,et al.Automatic Braille Code Translation System[J].CIARP 2005,LNCS 3773,2005:233-241.

[7] Muhammad Abuzar Fahiem.A Deterministic Turing Machine for Context Sensitive Translation of Braille Codes to Urdu Text[J].IWCIA 2008,LNCS 4958,2008:342-351.

[8] Amany Al-Saleh et al.Dot Detection of Optical Braille Images for Braille Cells Recognition[J].ICCHP 2008,LNCS 5105,2008:821-826.

【上接第28页】

简化此系统的数学模型为二阶系统:

$$G(s) = \frac{1}{s^2 + 2s + 1}$$

初选 $kp' = 1$, $ki' = 0$, $kd' = 1$, 其仿真模型

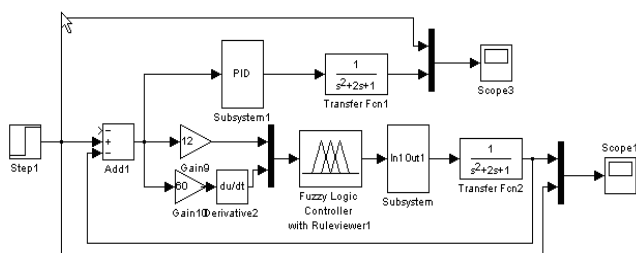


图6 模糊PID自整定仿真模型

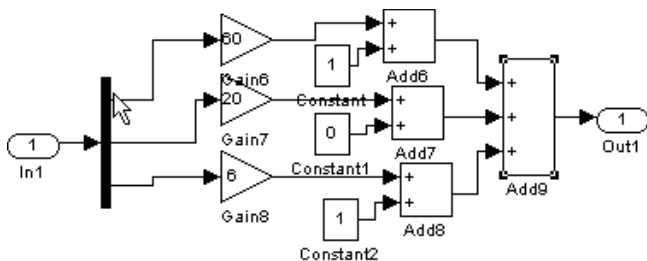


图7 模糊PID自整定子系统模型

如图6, 7所示:

5 仿真分析

针对焊管切割机控制系统的数学模型, 将模糊自适应PID调节器对给定阶跃曲线的跟踪结果进行了仿真。仿真结果表明: 模糊自适应PID控制器阶跃响应的上升时间要小于常规的PID阶跃响应的上升时间; 并且模糊自适应PID控制器的响应曲线的超调量大大减少, 而常规的PID控制器的响应曲线的超调量比较大。可见, 相对于常规PID控制器, 模糊自适应PID控制器具有更快的响应速度,

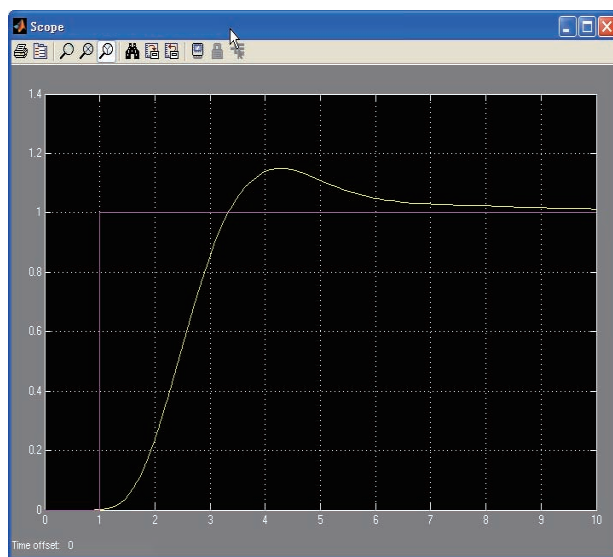


图8 模糊PID自整定仿真结果

更小的响应超调量, 能够很好的满足焊管切割位置控制的要求。其响应曲线如图和图所示:

由此可见, 采用模糊自适应PID控制控制的算法, 系统的响应速度加快、调节精度提高、稳定性变好, 这是单纯的PID控制难以实现的。

参考文献:

- [1] 张铁军.基于模糊控制的剪切线控制系统[D].沈阳:沈阳工业大学,2005.
- [2] 蒋利佳.位置伺服控制在飞剪控制系统中的分析与应用[D].武汉:武汉科技大学,2002.
- [3] 孙增沂,张再兴,邓志东.智能控制理论与技术[M].北京:清华大学出版社,1997,52-54.
- [4] 李军华,安连祥,贾俊林.带有自调整因子模糊控制器的设计[J].河北工业大学学报,2002,4(2):87-90.
- [5] 韩瑞珍,陈国定,杨马英.基于模糊推理的自整定PID控制器[J].基础自动化,2002,9(1):38-40.