

基于图像识别的阅卷系统的设计与实现^{*}

张 站^{1,2}, 刘政怡^{1,2}, 吴建国^{2,3}, 李 炜^{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 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039;

3. 安徽大学 软件学院, 安徽 合肥 230039)

摘 要: 设计了一种阅卷系统, 允许使用者使用任何颜色的圆珠笔、钢笔或铅笔在一般纸张的固定位置上涂写如“√”、“×”、“○”等手写符号完成答题, 由阅卷系统采用图像识别技术对答题卡图像进行预处理、符号识别、统计判分, 实现阅卷自动化。实验表明, 由于结构方面的相似性, 本阅卷系统识别错误集中在符号“√”、“×”上, 而对符号“○”的识别基本无误。

关键词: 图像识别; 字符识别; 阅卷系统

中图分类号: TP39

文献标识码: A

文章编号: 1674-7720(2011)04-0044-04

Design and implementation of scoring system based on image recognition system

Zhang Zhan^{1,2}, Liu Zhengyi^{1,2}, Wu Jianguo^{2,3}, Li Wei^{1,2}

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

2. Key Laboratory of Intelligent Computing & Signal Processing of Ministry of Education, Anhui University, Hefei 230039, China;

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

Abstract: This paper designs a scoring system that allows users to use any color pen, paper, pen or pencil in the general location of the graffiti, such as a fixed "√", "×", "○" and other symbols to complete hand-written answer. By the scoring system uses image recognition technology to answer card image preprocessing, character recognition, statistical sub-divided to achieve grading automation. Experiments show that, due to structural similarity, the scoring system focused on the symbol recognition errors "√", "×", while the symbol "○" Identification of the basic correct.

Key words: image recognition; character recognition; scoring system

高校学生的期末考试是普通高等学校管理的一项重要工作, 它是根据国家对高等学校学生德智体全面发展的要求, 按照统一的原则、方法和程序, 对学生学习和行为的表现进行阶段和全程的质量考核、记载、评价和处理。目前客观题型较多的英语考试或计算机基础考试均实现了阅卷自动化, 无需教师干预, 既能加快试卷反馈的速度, 又能保证公平公正。传统的阅卷系统利用光学标记阅读机的光电变换原理, 对填涂在答题卡上的内容进行高速采集, 然后进入计算机处理^[1]。但是它对于答题卡纸张和印刷质量要求太高, 如套印误差和剪切误差必须分别控制在 0.1 mm 和 0.2 mm 以内; 对使用者涂写要求太高, 限制太多, 如需要用专用铅笔涂满长方条, 不

得涂出长方条外, 整张答题卡涂写要深浅一致等; 机械传动机构复杂、使用寿命短、维护量大、设备一致性差^[2]。

本文提出一种基于图像识别的阅卷系统, 允许使用者使用任何颜色的圆珠笔、钢笔或铅笔在一般纸张的固定位置上涂写如“√”、“×”、“○”等手写符号完成答题, 由阅卷系统采用图像识别技术自动识别答题卡信息, 实现阅卷自动化。本系统与传统的光学标记阅读机阅卷系统的不同之处在于, 本阅卷系统对答题卡纸张和填涂符号均无特殊要求, 无需特制答题卡, 无特殊涂写要求^[3]。

1 系统实现流程及模块组成

(1) 设计和定义答题卡, 将允许用户填写的手写符号限定在一定的矩形区域内;

(2) 通过 CCD 采像设备采集答题卡图像, 经扫描、采样、量化等过程, 同时将答题卡图像以 JPG 格式保存在

《微型机与应用》2011 年 第 30 卷 第 4 期

^{*} 基金项目: 安徽大学人才队伍建设经费资助(02203105), 安徽大学青年科学研究基金(2009QN0019B), 安徽大学质量工程教改项目(XJ200921)

图形、图像与多媒体

主机内存和硬盘里;

(3)对答题卡图像进行预处理,包括粘连字符分割、灰度变换、二值化、图像偏斜纠正、平滑和细化等过程,以将答题卡固有的纸张问题、书写不规范、答题卡传动机械定位精度所带来的干扰因素排除掉,并为后续的识别工作做好准备^[4];

(4)对答题卡信息进行识别。通过图像分析抽取图像,并经过综合特征提取,结合答题卡表格的逻辑结构和几何结构,准确识别填涂在矩形块位置上的“√”、“×”、“○”等手写符号得到答题信息。答题卡识别算法的优劣决定了整个系统的性能(如识别的精度、可靠性等),是整个软件系统中最重要的部分;

(5)对识别的答题信息进行加工、整理、分析和统计,结合软件的设置对答题卡信息进行判分。

综上所述,系统实现流程如图1所示。

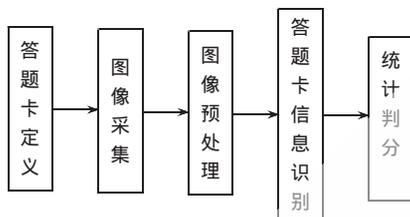


图1 系统实现流程

为了便于学生和教师查询和分析阅卷结果,本阅卷系统还设置了用户登录模块、学生信息管理模块、课程基本信息管理模块、试卷分析模块、成绩查询模块、答题卡图像查询模块、报表打印模块和补考管理模块。其中试卷分析模块对试卷的各项数据进行分析,包括最高分、最低分、平均数、及格率方差、标准差等,完成试卷分析功能;答题卡图像查询模块允许学生通过姓名查询存储在硬盘中的答题卡图片;补考管理模块统计不及格、缺考或舞弊学生信息,生成补考表;报表打印模块完成答题卡图片、成绩报告单、试卷分析报告和补考表等的打印功能。系统模块组成如图2所示。

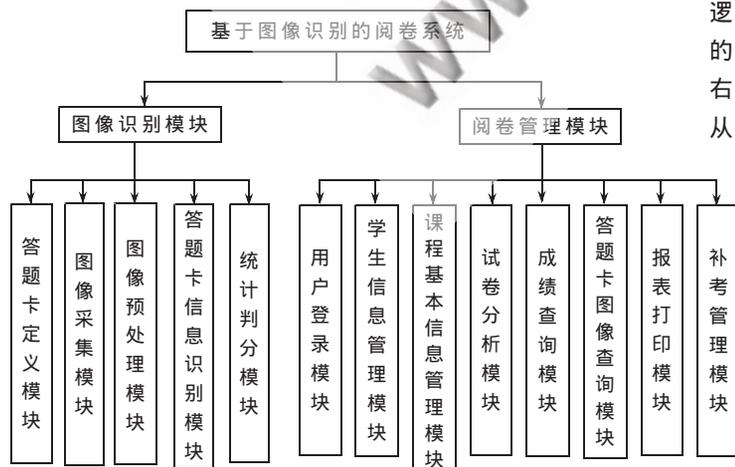


图2 系统模块组成

2 系统主要实现技术

2.1 答题卡设计

答题卡设计成表格的形式,由试卷答题卡基本信息、考生基本信息、答题卡填写说明及题目选项列表组成,考生用“√”、“×”、“○”等手写符号在相应的矩形框中选择答案。答题卡示例如图3所示。

试卷答题卡基本信息: × × × ×		考生基本信息: × × × ×			
答题卡填写说明: × × × ×					
选项	A	B	C	D	
题号					
1	√				
2			√		
3		○			
⋮	⋮	⋮	⋮	⋮	⋮

图3 答题卡定义示例

答题卡版面与普通表格一样具有几何结构和逻辑结构,几何结构反映了所填写的信息区域的位置和大小,逻辑结构则表示答题卡中所填写信息的实际意义以及填写信息与填写项之间的对应关系。

在传统的光学标记阅读机阅卷系统中,答题卡的几何结构大部分都是采用定位标记块进行描述,这种描述方法具有很多缺点,如浪费版面空间、缺乏灵活性、版面设计比较复杂、对印刷质量要求高、不便于修改等,而且定位标记块看起来也不美观。在本系统答题卡版面中,填写信息在几何结构上可以看成由若干个互不相交的矩形块组成,它们组成了答题卡版面的最小单位^[5],可以以答题卡的边框线建立二维坐标系,以矩形的对角顶点坐标来描述矩形块的位置和大小,完成对答题卡的几何结构描述,该方法简洁、灵活,便于识别。

答题卡的逻辑结构描述是定义矩形块的属性。矩形块的属性包括对填写项以“√”、“×”、“○”表示的选择或者不选择。

本系统采用文档结构描述语言同时对几何结构和逻辑结构进行描述。设一张答题卡包含 n 个填写有信息的矩形块 B_1, B_2, \dots, B_n , 矩形块之间存在着上下结构和左右结构的几何位置关系,其逻辑顺序一般是从上到下、从左到右,在描述文档结构时也采用这种顺序,则文档结构描述语言 DDL 表示如下:

$$DDL(i, n, B_1((x_1, y_1), (l_1, w_1), attr_1), B_2((x_2, y_2), (l_2, w_2), attr_2), \dots, B_n((x_n, y_n), (l_n, w_n), attr_n))$$

其中 i 是答题卡的序号; n 是矩形块的总数; x_i 为矩形块在水平方向的位置; y_i 为矩形块在垂直方向的位置; l_i 为矩形块的长度; w_i 为矩形块的宽度; $attr_i$ 表示矩形块的属性,当 $attr_i$ 为 0 时,表示该矩形块的填写内容为待识别的字符, $attr_i$ 为 1 时,矩形块作为图像保存。

将答题卡设计成常见的表格形式,一方面它和欢迎网上投稿 www.pcachina.com 49

图形、图像与多媒体

Image Processing and Multimedia Technology

一般考试用到的答题卡的形式类似,符合人们的使用习惯和书写习惯;另一方面可以利用文档描述语言对它的几何结构和逻辑结构进行描述,以便更好地对矩形框内的字符特征进行提取与识别,辨别矩形框中的字符,与标准答案比对,对考生客观题进行判分。

2.2 答题卡信息识别

本系统采用汉字识别方法中的统计决策方法对手写字符进行识别,如图4所示,首先提取字符特征,对字符进行分类判别,经过训练和识别两个阶段,最后识别出字符^[6]。训练阶段,收集日常人们习惯的手写符号为样本,经过筛选分类建立样本库,以便对待识别字符进行分类和识别。识别阶段,将待识别符号的特征与训练阶段中所建立的标准样本特征比较,计算最大相似度以判别该手写符号所属的类别^[7]。

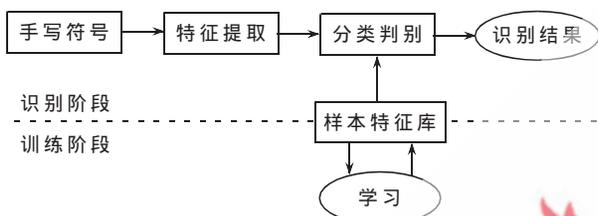


图4 答题卡信息识别框架图

2.2.1 字符特征提取

学生在答题过程中由于紧张和个人原因,写出的√、×、○等答题符号往往千变万化。为准确识别这些手写符号,需要进行字符结构特征提取,将最能体现这个字符特点和字符间差异的结构特征提取出来。本系统提取的特征是以下特征的组合。

(1)点特征

点特征是一种重要的结构特征,是指字符笔画中的端点。端点反映了字符中笔画的起点和终点信息,与该点相连的点数^[1]。

(2)笔画密度特征

笔画密度特征是取得符号水平方向笔画密度函数 $d(x)$ 和垂直方向的笔画密度函数 $d(y)$,然后进行相同项合并^[1]。如图5所示,符号“○”水平方向笔划密度函数 $d(x)$ 和垂直方向的笔划密度函数 $d(y)$ 分别是: $d(x)=(1, \dots, 1, 2, \dots, 2, 1, \dots, 1)$, $d(y)=(1, \dots, 1, 2, \dots, 2, 1, \dots, 1)$,则合并后笔画密度可表示为 $d(x)=d(y)=(1, 2, 1)$ 。



图5 笔画密度特征

(3)基于链码方法的结构特征

从曲线起点开始与其相连的像素点有8种可能的方向; $k \times 45^\circ (k=0, 1, \dots, 7)$,如图6所示,如果两个像素点间的连线方向为 $k \times 45^\circ$,就用“ k ”作为这条连线的代码,则一条曲线最终可近似地用下式表示:

$$A_n = a_1 a_2 \dots a_n, a_i \in \{0, 1, 2, \dots, 7\}, i = 1, 2, \dots, n$$

(4)空洞特征

在二值图像中,被目标像素1包围的背景像素0(的集合)称为空洞(hole)。在字符的骨架线的链码形成过程中,若搜索到的下一点就是该骨架线的搜索起始点,同时已形成的骨架链码码长超过了一定的阈值,则认为搜索到一个空洞^[5]。

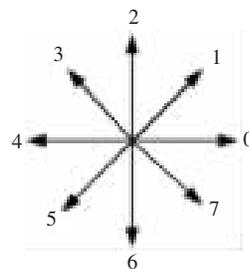


图6 方向码

(5)横竖方向交叉特征

横向或纵向扫描字符,某一行或列的像素由白变黑的次数就是该行或列的横或纵向交叉特征。本系统将横竖两个方向距离不等的7条线作用于字符,计算水平和垂直方向与字符的交叉数,如图7所示。

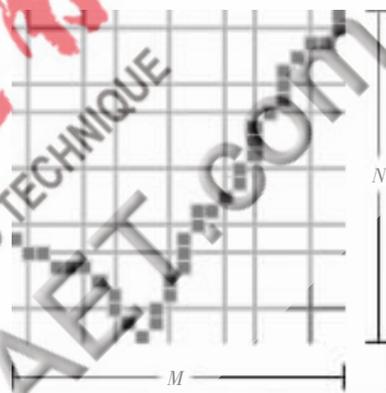


图7 横竖方向交叉特征

2.2.2 符号模型库建立

答题卡信息识别的训练阶段需要建立符号模型库,以便对待识别手写符号进行分类和识别。符号模型库建立的好坏直接影响分类器的应用,从而影响手写符号识别效果^[5]。

由于手写符号的多样性,需要选择某一类手写符号中具有代表性的多个样本来构造标准样本,本系统采用手写字符样本特征向量的均值来描述类目标。设有 n 个符号类,每个符号类中有 a 个训练样本,每个样本有 b 个符号特征,每个符号类中样本的特征记为 f_{kj} , k 为样本特征序号, j 为各个手写符号的样本序号,则第 i 个目标类特征的均值为 $P(i)$,即:

$$P(i) = [p_{i1}, p_{i2}, \dots, p_{ik}, \dots, p_{ib}]$$

$$p_{ik} = \sum_{j=1}^a \frac{f_{kj}}{a} \quad (1 \leq k \leq b)$$

每次计算 P_{ik} 时, k 为大于等于1小于等于 b 的固定值。 P_{ik} 为对于第 i 个目标类中 a 个样本中各个样本对应第 k 个特征值的均值。

2.2.3 手写符号识别

对答题卡中矩形框信息识别包括两个方面,一是识别矩形框中是否有字符,二是识别具体是哪种字符。其

《微型机与应用》2011年第30卷第4期

图形、图像与多媒体

Image Processing and Multimedia Technology

中识别是否书写了字符比较简单,只要比较增强对比度后的矩形框图像与已知填有信息的矩形框的均方差大小,即可识别是否有字符,因为空白的矩形框和被书写的矩形框均方差差别很大。下面主要介绍怎样识别矩形框中的具体字符。

手写符号识别就是在提取到符号的特征向量之后,依据一定的判别函数来判定出某一图形点阵具体代表的是哪一个手写符号。

判别函数可以先简单地作如下定义:考虑有 P_1, P_2, \dots, P_m 个符号类别,假使每类有一个标准样本,则共有 m 个标准样本,分别表示为 k_1, k_2, \dots, k_m 。任意一符号特征向量 X 和第 i 个 ($i=1, 2, \dots, m$) 标准样本间的“相似度”为 R_i 。计算待识别的符号特征向量 X 与每类标准样本之间的“相似度”^[7],并将 X 分到与它“相似度”最大的类别,即对所有的 j 不等于 i ,若 $D_i > D_j$,则 X 就属于 P_i 类符号。

系统采用基于最邻近域分类器的模板匹配算法来对手写符号进行识别。

首先定义字符特征向量,经过前面的特征提取分析,该特征向量为一个 16 维向量, $X = \{x_1, x_2, \dots, x_{16}\}$, 具体定义为:

x_1 : 孔洞数;

x_2 : 端点数;

$x_3 \sim x_9$: 7 条水平线与字符的交叉次数;

$x_{10} \sim x_{16}$: 7 条竖直线与字符的交叉次数。

通过度量待识别字符和样本库中样本字符的接近程度,确立最近分类的一个准则。在最邻近分类中,经常使用的是相似度。如图 8 所示,在提取了待识别字符的特征向量并建立了字符库后,将待识别字符和样本库中第 i 个样本的特征向量之间求近似度 $R(X, G)$ 。 $R(X, G)$ 定义如下:

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

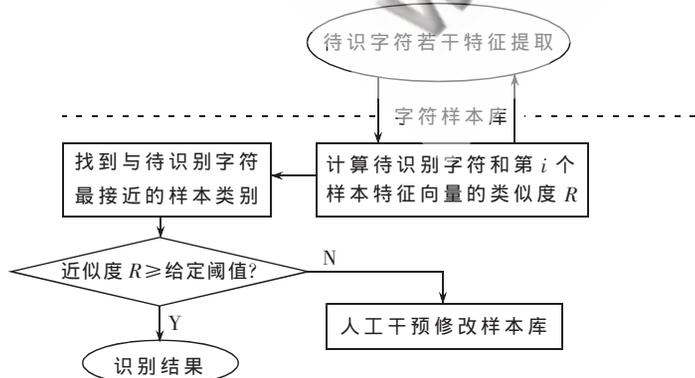


图 8 基于最邻近域分类器的模板匹配算法的实现框图

式中, x_i 为待识别符号特征向量的第 i 个分量, g_{ik} 为样本库中第 i 个标准样本的第 k 个分量, m 为样本类别数。分子为向量 X, G 之间的内积, 分母分别为向量 X, G 的模。 α 是向量 X, G 在 m 维空间的夹角。显然, 当 X, G 两个向量完全相同时, 其夹角为 $0, R(X, G) = 1$, 它们的距离 $D(X, G) = 0$, 即相似度最大。求出最大 $R_i(X, G)$, 若 $R_i(X, G) \geq$ 给定阈值, 即可找到与待识别字符最接近的样本类别, 否则人工干预并修改样本库^[7]。

3 实验结果与分析

实验采用 CCD 摄像头采集答题卡图像, 经图像预处理、若干特征提取、信息识别等过程, 判定矩形框中是否有字符、是什么字符, 最后对答题卡信息分析和统计。实验采用 100 份试卷作为样本, 对 20 份试卷进行测试, 结果发现识别错误的手写符号主要是“√”和“×”, 原因在于这两者在结构方面相似, 而符号“○”的识别率达到 100%。

本系统将图像预处理、字符特征提取与图像识别等技术应用于阅卷系统的开发, 实现了阅卷自动化, 加快了成绩考核的速度, 改善了教学管理环境。相比于传统的基于光学标记阅读机的阅卷系统, 本系统利用图像识别技术实现阅卷自动化, 不需要特殊的答题卡, 考生也可以随意使用各种“√”、“×”、“○”等手写符号进行答题, 不必用指定的 2B 铅笔填涂矩形框, 更符合人们的习惯。

参考文献

- [1] 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.
- [2] 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.
- [3] 吴元君, 张婷, 雷惊鹏. 一种改进的 OMR 技术在标准化考试中的应用[J]. 计算机教育, 2007(13): 250-272.
- [4] 丁慧东. 脱机手写体汉字识别研究[D]. 长春: 东北师范大学, 2006.
- [5] 庞东虎, 金伟杰. 英文字符特征提取系统[J]. 计算机仿真, 2007, 24(12): 208-210.
- [6] 杨玲, 毛以芳, 吴天爱. 基于多特征多分类器的脱机手写汉字识别研究[J]. 计算机与网络, 2008(01): 217-217.
- [7] 覃胜, 刘晓明. 基于图像的 OMR 技术的实现[J]. 电子技术应用, 2003, 29(10): 17-19.
- [8] 翁功平. 光标阅读机 OMR 原理的设计与实现[J]. 工业控制计算机, 2010, 23(04): 61-62.

(收稿日期: 2010-12-29)

作者简介:

张站, 男, 1986 年生, 硕士, 主要研究方向: 智能 EDA。

刘政怡, 女, 1978 年生, 博士, 副教授, 主要研究方向: 图像识别, 自然语言处理。

吴建国, 男, 1954 年生, 博士, 教授, 主要研究方向: 中文信息处理, 智能 EDA。