

一种适用于移动设备在线阅卷的 答题卡自动识别算法

孙琳, 张琪, 罗念祖, 邓书勤, 陈念年

(西南科技大学 计算机科学与技术学院, 四川 绵阳 621010)

摘要: 基于智能移动设备的阅卷方式无需专用设备即可快速阅卷, 降低成本的同时还能增加阅卷工作的可移动性; 为解决移动设备拍照阅卷时, 出现的阴影、反光、倾斜 (-45° 至 45°) 等情况, 设计并实现了一套答题卡自动识别算法, 算法主要分为三部分: 图像预处理、待识别区域定位与分割、答题卡内容识别; 不同操作系统 (PC、Android、IOS) 下测试结果表明: 该算法正常填涂采集识别率为 100%, 异常填涂采集识别率为 93.6%; 识别速度小于 2 s, 满足实时性要求; 无需修改就能在不同操作系统下编译运行, 提高了程序的通用性和兼容性; 目前该算法已成功应用于某教育企业上线 APP 中。

关键词: 自动阅卷; 跨平台; 答题卡识别; 图像预处理

An Automatic Recognition Algorithm for Online Marking of Mobile Devices

Sun Lin, Zhang Qi, Luo Nianzu, Deng Shuqin, Chen Niannian

(School of Computer Science and Technology, Southwest University of Science and Technology,
Mianyang 621010, China)

Abstract: The marking method based on intelligent mobile device can work fast without special equipment, reduce costs and increase the mobility of marking work. To solve the problem when using a mobile device for marking, such as shadows, reflections, tilted -45° to 45° and so on, this article designed and implemented a set of sheet automatic identification algorithm. The algorithm is mainly divided into three parts: image preprocessing, identify regional orientation and segmentation, sheet content recognition. Experimental results show in different operating systems (PC, Android, IOS): the recognition rate of the algorithm is 100%, and the detection rate of abnormal filling is 93.6%; the recognition speed is less than 2s, satisfying the real-time requirement; it can be compiled and run under different operating systems without modification, thus improving the universality and compatibility of the program. At present, this algorithm has been successfully applied to a education enterprise on-line APP.

Keywords: automatic marking; cross-platform; exam card recognition; image pre-processing

0 引言

答题卡主要用于内容选项较多的考试中。常见的答题卡自动阅卷系统主要有光标阅读机^[1]、扫描仪阅卷系统^[2]和摄像头阅卷系统^[3]。这几种常用的阅卷系统虽已达到了识别速度快、效率高, 但工作时却受设备的硬件条件所限, 主要在 PC 或嵌入式端操作, 使得办公不够灵活, 且存在价格高昂、对纸张质量和印刷的要求高等缺点。因此, 它们主要被应用中高考、研究生考试、公务员考试等一些大型的正式考试场合中。然而对于像学校日常的随堂测验,

英语四六级、会计资格证、银行从业资格证等培训考试, 这些需要用到答题卡的非正式考试场合, 单位或人们更倾向于以一种低廉、简单、便捷的操作方式来进行阅卷。移动互联网的高速发展使得智能手机、智能平板电脑等移动设备越来越普及, 在国家教育信息化的推进行程中, 移动教育方式不断地受到关注。在此背景下, 出现了许多基于移动端的答题卡自动识别系统^[4-5]。

目前基于 PC 端或是基于移动端的答题卡自动识别系统, 大部分是独立开发的, 程序高度定制化, 因而限制了其普遍适用性, 相应地增加了开发周期和成本。在实际使用时, 由于个体间差异较大, 答题卡填涂时可能出现填涂大小和深浅不一致、擦除不干净等情况, 进行采集识别时, 不同的光照、摆放位置会造成采集的图像出现阴影、反光、倾斜等问题, 这些会相应的增大识别难度。

为解决以上问题, 本文设计研究一套识别率高、实时性强、能兼容适应不同采集设备的答题卡识别算法。PC 机可利用外接摄像头, 进行图像采集, 随后通过计算机软件

收稿日期: 2018-04-03; 修回日期: 2018-04-24。

基金项目: 四川省教育厅科技成果转化重大培育项目 (14zd1102); 西南科技大学龙山学术人才科研支持计划 (17LZX425); 西南科技大学研究生创新基金 (17ycx053)。

作者简介: 孙琳 (1995-), 女, 四川广安人, 硕士研究生, 主要从事机器视觉与高精度测量技术研究工作方向的研究。

陈念年 (1977-), 男, 硕士, 副教授, 硕士研究生导师, 主要从事机器视觉高精度测量、图像处理方向的研究。

识别答题卡, 这种方式主要应用在办公室阅卷。移动设备普遍具有较高分辨率的相机, 利用其采集答题卡图像, 然后在 App 上进行识别, 这种方式便捷、高效、成本低, 适合于移动办公, 能够配合基于移动设备的教育软件一起使用, 具有较大的应用价值和前景。

1 答题卡设计

答题卡要求识别正确率达到 100%, 因为一旦出现识别错误, 将直接影响考生本场考试成绩。考虑成本要求, 在答题卡纸质质量和印刷方式方面要求不严格, 支持普通纸 60 g 以上的打印、复印等方式的制作答题卡。针对以上需求, 标准答题卡模板设计如图 1 所示, 四周采用方框做边界标定, 相比传统采用原点作为标记点, 在外观上更简洁大方。

答题卡主要包括两大部分: 个人信息区和答题区。个人信息区内容为身份证号和试卷号(由数字+“X”构成), 在手写体识别中, 由于个体间差异较大, 书写习惯不同, 导致识别率不能得到保证。因此, 为达到识别正确率高的目标, 设计如下数字填涂方格(1... 9)与(X), 以小圆点作为标记, 引导考生做如下书写: 0123456789X。答题区每道题有 A、B、C、D 四个选项, 每个选项之间的间隔相等, 行之间的行间距相同, 每行的行高也完全一致(可增加题量)。

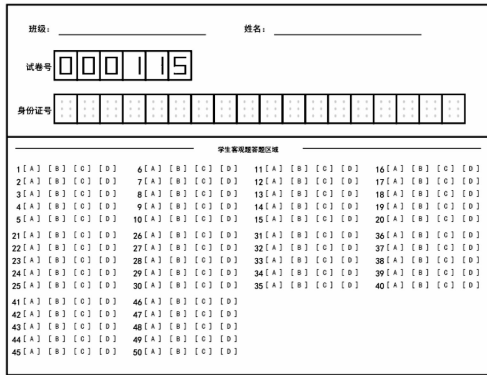


图 1 答题卡模板

2 答题卡图像处理与识别过程

针对答题卡图像特点, 设计算法流程图如图 2 所示。

1) 答题卡预处理: 主要包括对图像降噪、去阴影、灰度拉伸、二值化、形态学等处理, 去除图像采集过程中受到的一些外在因素影响。

2) 待识别区域定位: 根据预处理后的图像特点, 建立几何映射关系, 变换得到模型参数矩阵, 并结合插值算法校正图像。

3) 答题卡内容识别: 对校正后的答题卡图像, ①定位分割填涂项, 结合填涂点灰度、面积等属性设计算法, 进行识别; ②定位分割数字字符, 做归一化等预处理操作后, 采用模板匹配结合结构化特征的方法进行一一识别。

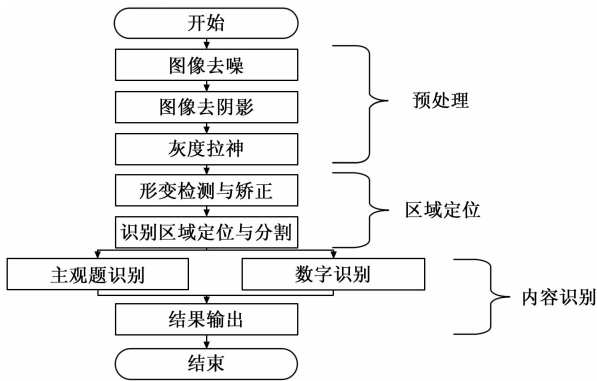


图 2 算法设计流程图

2.1 答题卡图像预处理

在拍照过程中, 受遮挡或光照不均匀等因素影响很容易在拍摄的图像上产生阴影, 若答题卡采用铅笔填涂还容易出现反光现象。如图 3 (a) 所示, 拍摄时造成了图像右区域存在阴影, 答题卡采用铅笔填涂, 在 26~30 题项区域出现反光现象, 若直接采用自适应阈值做二值化处理, 会出现信息区域部分缺失, 如图 3 (b) 所示, 这样会影响识别结果。因此, 识别前应对答题卡做预处理。分析答题卡图像, 发现答题卡内容上表现出许多的细节信息, 出现的阴影区域呈片型且亮度变化缓慢。一般情况下, 线性滤波会给图像的细节带来模糊问题, 这样便可利用梯度滤波器来去除图像中的阴影^[6], 将阴影作为目标, 答题卡内容作为背景, 利用均值滤波器和背景差分法^[7], 即可提取出阴影。设 $W(x, y)$ 为大小为 $M \times N$, 中心点在 (x, y) 处的均值滤波器窗口, $I_{mean}(x, y)$ 为均值滤波后的图像, $I(u, v)$ 为原始图像, 均值滤波器是计算滤波窗口内的像素均值, 并将其赋给窗口中心处的像素:

$$I_{mean}(x, y) = \frac{1}{M \times N} \sum_{(u, v) \in W(x, y)} I(u, v) \quad (1)$$

去除阴影后的图像:

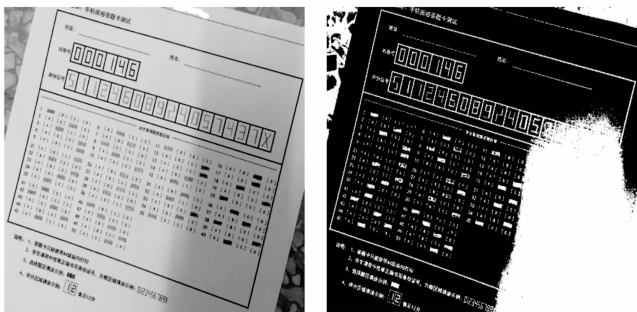
$$I' = 255 + I_{mean} - I \quad (2)$$

去除阴影后的目标图像表现亮度不均, 对比度较差, 利用灰度拉伸延展整个灰度级范围, 使灰度变化具有分段性, 图像边缘更清晰, 从而改善图像质量。预处理效果如图 3 (c)、(d) 所示。

由于目前市面上的移动设备相机普遍具有较高的分辨率, 获得的目标图像数据量大, 为节省设备存储空间, 获得更高效的处理效率, 系统研发时对采集的原始图像按预设比例归一化成标准规格图像后, 再执行答题卡的识别过程。

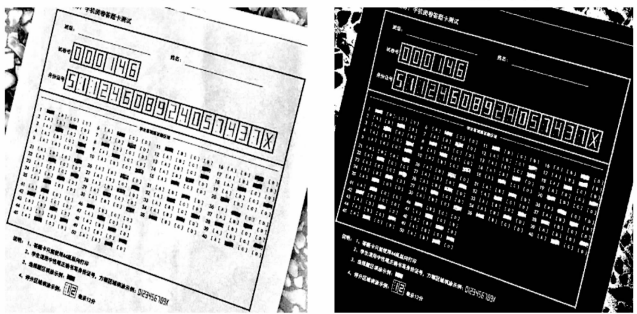
2.2 待识别区域定位分割

相机在拍照过程中, 由于纸张摆放偏差或者相机未平行拍照都有可能使图像产生旋转或者倾斜, 如图 3 (a) 所示。因此, 需要对答题卡图像是否发生形变进行检测, 同时进行相应的校正处理。答题卡是由方框作为边界, 具有



(a) 原图

(b) 二值化(a)图



(c) 去阴影(a)图

(d) 二值化(c)图

图 3 预处理(去阴影、处理反光效果)

突出的 4 个顶点, 对获得的轮廓图像采用多边形逼近方式定位出答题卡的角点坐标位置, 建立相应地映射关系对图像进行倾斜和透视校正。多边形逼近^[8]算法的基本思想是在获得的轮廓上寻找最远的两个点连成线段, 然后在轮廓上查找到该线段最远的点, 并将其添加到新轮廓, 反复迭代, 不断添加新的点到结果中, 最后得到逼近的多边形顶点。定位到角点后, 通过三角函数变换确定目标图像的倾斜角, 采用透视变换的方式对图像进行倾斜校正。透视变换通过将检测得到的目标图像的四个定点坐标 $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ 和基准答题卡的四个顶点坐标 $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4)$ 之间建立线性坐标映射关系, 得到透视变换参数矩阵, 其中 a, b, c, d, e, f, m, l 是透视变换的参数。

$$\begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \\ X_4 \\ Y_4 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -X_3x_3 & -X_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -Y_3x_3 & -Y_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -X_4x_4 & -X_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -Y_4x_4 & -Y_4y_4 \end{bmatrix} \times \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ m \\ l \end{bmatrix} \quad (3)$$

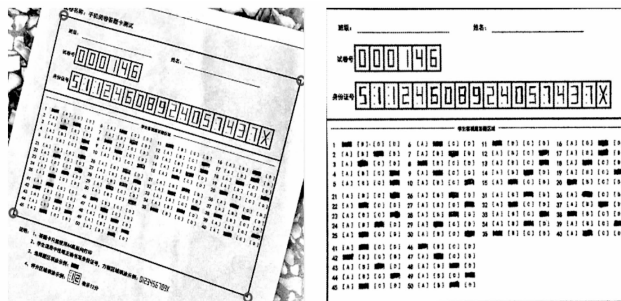
把上式记作:

$$B = A * V \quad (4)$$

由(4)式得:

$$V = A^{-1} * B \quad (5)$$

求解以上线性方程组得到参数值, 将目标图像的每个像素点, 变换到新的坐标点, 完成图像的校正。图像校正过程中, 像素坐标实际为整数, 变换后的图像可能出现空隙, 所以需要变换后的目标图像进行插值运算。从理论上讲, 双线性插值^[9]性能介于最近邻域插值与三次卷积插值之间, 若采用三次卷积插值, 算法时间复杂度较高, 对图像处理速度太慢, 而采用最近邻域插值效果不佳。因此, 针对答题卡特征表现较为单一的特点, 采用双线性插值的方法来使图像信息更接近原图。结果如图 4 所示。



(a) 定位识别区域(红圈表示定位的角点位置)

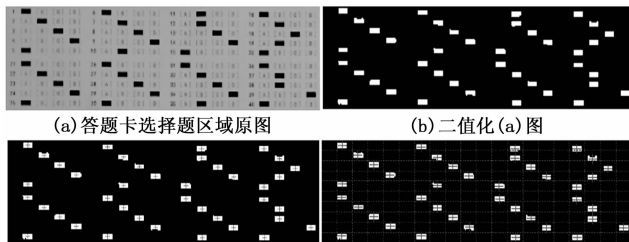
(b) 校正(a)图后二值化

图 4 拍摄角度校正效果

对校正后的答题卡图像利用连通域、正交投影定位分割得到答题区域及个人标识区域。

2.3 答题区识别

实际考试过程中, 由于考生差异化填涂(例如: 用笔不同、填涂大小和深浅不一致等), 这些都有可能影响识别结果。为避免上述情况, 对分割得到的答题区图像, 做形态学处理, 采用自适应阈值进行二值化, 然后根据填涂点灰度、面积及所处位置作为评价指标, 转化得到选项答案。算法使用水平和垂直投影的方式, 确定答题区每组题目之间的距离, 题间距离及选项距离。利用相互之间的距离特性, 对答题卡客观题区域划分网格, 每个网格对应题的选项答案(A, B, C, D)位置或者题号(1, 2, 3等)位置, 存储并记录网格所代表的信息, 及每道选择题所对应的答案位置。然后提取答题卡填涂区域的质心, 按其所处位置转化为选项答案。同时计算该区域像素的平均灰度值及区域面积, 进行有效填涂判定。最后, 对以上结果进行综合判定, 确定选项答案。该方法也适用于多选题。



(a) 答题卡选择题区域原图

(b) 二值化(a)图

(c) 定位出各选项的中心

(d) 结合灰度、面积等属性转化为选项答案

图 5 选择题识别

2.4 个人信息识别

目前字符模式识别方法大致可分为统计模式识别^[10]、结构模式识别和人工神经网络^[11-12]等。模板匹配^[13]是统计模式识别方法中常用方法之一，将待识别数字与预先准备好标准模板进行匹配，根据匹配相似度确定当前识别结果。结构模式识别通过寻找字符本身特征，譬如横线、竖线、连通区域等来判定当前识别内容。神经网络通过对大量数据样本进行训练学习，得到每类样本的特征表示，进而分类识别。

由于神经网络识别方法，参数较多计算复杂度较高，训练过程需要有标签的庞大数据集且会花费大量时间。因此，采用模板匹配结合结构化特征的方法对书写结果进行识别。识别过程如图 6 所示。

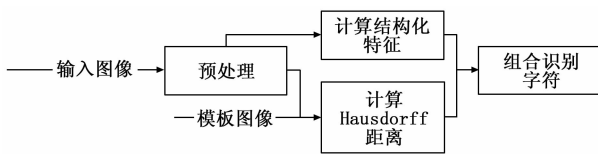


图 6 数字识别原理框图

首先依据设计的数字填涂方格中标记原点的位置关系设计如下模板图像（尺寸大小 22×34）：

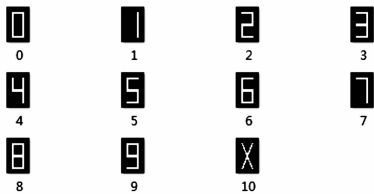


图 7 字符模板图像

归一化输入图像大小，使其与模板图像尺寸相同。然后计算模板与待识别图片的 Hausdorff 距离作为相似性判断，同时计算输入图像的结构化特征。最后组合判定得到识别结果。

2.4.1 Hausdorff 距离

Hausdorff 距离是用来描述两个实体之间相似性的一种度量方式，因平移、旋转、缩放等变换对其影响较小，而受到广泛运用^[14]。假设有两个点集 $A = \{a_1, \dots, a_n\}$ ， $B = \{b_1, \dots, b_n\}$ ，定义 A、B 点集合之间的 hausdorff 距离为：

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (6)$$

其中：

$$h(A, B) = \max(a \in A) \min(b \in B) \|a - b\| \quad (7)$$

$$h(B, A) = \max(b \in B) \min(a \in A) \|b - a\| \quad (8)$$

$\| \cdot \|$ 是点集 A 和 B 间的 Euclidean 距离范式。

2.4.2 结构化特征

分析填涂数字的特征，可通过数字的孔洞质心位置、欧拉数（欧拉数=连通域数-孔数）、横线、竖线（利用水

平投影、垂直投影计算）来区分部分数字，见表 1。

欧拉数：用于将数字分为 3 类： $\{0, 6, 9\}$ ， $\{8\}$ ， $\{1, 2, 3, 4, 5, 7, X\}$ ；

孔洞质心位置：用于区分 $\{0\}$ ， $\{6\}$ ， $\{9\}$ ；

横线、竖线：用于区分 $\{1\}$ ， $\{3\}$ ， $\{4\}$ ， $\{7\}$ ， $\{X\}$ 。

表 1 各数字结构特征

数字字符	欧拉数	横线	竖线	孔洞质心
0	0	2	2	中
1	1	0	1	—
2	1	3	2	—
3	1	3	1	—
4	1	1	2	—
5	1	3	2	—
6	0	3	2	下
7	1	1	1	—
8	-1	3	2	上、下
9	0	3	2	上
X	1	0	0	—

3 答题卡自动识别算法的实现

3.1 答题卡识别算法实现

跨平台（Cross-platform）的应用程序是指其在多种平台下不改动或稍改动即可编译运行^[15-16]。本文跨平台系统主要分为两部分：答题卡自动识别算法库和跨平台实现。总体架构如图 8 所示。

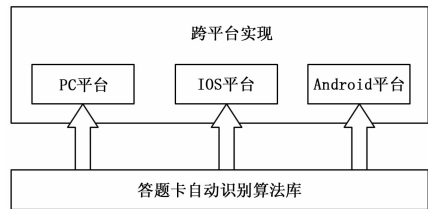


图 8 跨平台系统总体框架图

答题卡自动识别算法的开发环境是 VS2013 + OpenCV3.0，利用 C++ 语言编写实现。C++ 语言具有很好的跨平台性，面向对象编程适合处理像图像这类具有复杂数据结构的问题。OpenCV 是一个跨平台且开源的计算机视觉图像处理算法库，常用于商业和科研用途。在算法库的基础上实现跨平台性，共包括 PC 平台、Android 平台以及 IOS 平台。Android 平台下，将原始算法库封装成符合 JNI 命名规范的 so 库，并再次封装为能直接使用的 jar 包，做成发布版 Android SDK。IOS 平台下，由于 Xcode 支持 Object C 与 C++ 的混编，因此可直接将算法库编译成静态类库供开发人员调用。

3.2 测试结果与分析

测试使用的答题卡模板均由 6 位试卷号、11 位身份证号、50 个选项项组成。为验证算法的跨平台性，测试平台选择了基于 IOS、Android 系统的多种采集设备，分别采集了

表 2 不同机型正常填涂测试结果

机型	系统版本	运行内存	处理器	总识别数/张	识别率%	平均阅卷速度/ms
PC	Windows7	4GB	Intel Core i3-2130 3.4GHz	80	100	1934
iphone 6	IOS11.2	1GB	苹果 A8+M8 协处理器 1.38GHz	80	100	939
oppo A57	Android6.0.1	3GB	高通骁龙 435 1.4GHz	80	100	1626
oneplusA3010	Android 7.1.1	6GB	高通骁龙 821 2.35GHz	80	100	564
iphone7plus	IOS 11.2	3GB	苹果 A10+M10 协处理器 2.23GHz	80	100	245
vivo X9s	Android 7.1.1	4GB	高通骁龙 625 2.0GHz	80	100	630

110 张答题卡图像。

正常填涂采集 (正确填涂, 拍摄时答题卡方框区域拍摄完全, 成像区域不能过小) 图像 80 张, 测试结果如表 2 所示。正常填涂识别正确率为 100%。不同操作系统下算法运行时间在 2 s 内, 运行时间主要与设备处理器和运行内存有关, PC 采用的是双核处理器, 运行时间较长。

异常填涂采集 (答题卡的填涂大小、深浅、擦除不干净等特殊状况, 采集时不同的光照、摆放位置等情况) 图像 30 张, 如图 9 所示, 这些会相应的增大识别难度。经测试, 算法能有效处理拍照时出现的阴影、反光、倾斜 ($-45^{\circ} \sim 45^{\circ}$) 情况, 异常填涂识别正确率为 93.6%。不能识别 \checkmark 、 \times 、O 类填涂方式, 及答题卡旋转倾斜角度较大等情况。



图 9 不同条件下拍摄的图片

4 总结

本文针对答题卡图像采集时存在阴影、反光、倾斜等问题进行了研究, 采用梯度滤波结合背景差分法的方式去除阴影。对预处理后的图像, 采用多边形轮廓逼近的方式定位角点, 进而进行倾斜校正等操作, 为进一步识别做铺垫。采用网格定位法对填涂区域信息进行定位, 并结合填涂区域位置、面积及灰度等信息, 形成评判体系, 确定选项答案。对个人信息区利用模版匹配结合结构化特征的方式, 进行数字的提取识别。最后, 根据跨平台编程的思想, 设计实现了一套能兼容不同系统下不同的采集设备, 实际运用时识别速度快, 准确率高的答题卡自动识别算法。下一步将考虑复杂的拍摄背景, 以及答题卡出现折皱做进一步研究。

参考文献:

[1] 张凯兵, 黄襄念, 秦岸, 等. OMR 图像倾斜矫正与分割

[J]. 计算机应用, 2005 (03): 586-588.

- [2] 孙建芳, 李万秋, 张国辉, 等. 扫描阅卷系统中模板定制和图像聚类方法的研究 [J]. 计算技术与自动化, 2017, 36 (03): 105-108.
- [3] 宋锦秀. 基于图像识别的答题卡自动阅卷研究 [J]. 中小学电教 (下半月), 2015 (08): 69.
- [4] 陈浩鹏. 基于图像识别的移动端阅卷系统的研究与实现 [D]. 中山大学, 2014.
- [5] 郝平. 图像处理耦合模板定位的答题卡识别研究与应用 [J]. 计算技术与自动化, 2015, 34 (04): 105-109.
- [6] 秦秀丽. 基于 YUV 颜色空间和图论切割的阴影去除算法 [D]. 武汉: 武汉理工大学, 2010.
- [7] 朱明早, 罗大庸, 曹倩霞. 帧间差分与背景差分相融合的运动目标检测算法 [J]. 计算机测量与控制, 2005, 13 (3): 215-217.
- [8] 吴东东, 周东翔, 关涛, 等. 一种基于轮廓多边形逼近的可见光与红外图像配准方法 [J]. 计算机应用与软件, 2012, 29 (10): 28-30+122.
- [9] 宋湘芬, 林仕令, 宋磊, 等. 基于 B 样条的医学图像弹性配准 [J]. 生物医学工程研究, 2017, 36 (03): 213-217.
- [10] 陈嵩祥. 用于字符和数字识别的若干分类方法的比较研究: 实验结果 [J]. 计算机科学, 2015, 42 (S1): 102-106, 121.
- [11] 刘欢, 邵蔚元, 郭跃飞. 卷积神经网络在验证码识别上的应用与研究 [J]. 计算机工程与应用, 2016, 52 (18): 1-7.
- [12] Lu P, Shan L, Li J, et al. A new segmentation method for connected characters in CAPTCHA [A]. International Conference on Control, Automation and Information Sciences [C]. IEEE, 2015: 128-131.
- [13] Roh G P, Roh J W, Hwang S W, et al. Supporting Pattern-Matching Queries over Trajectories on Road Networks [J]. IEEE Transactions on Knowledge & Data Engineering, 2011, 23 (11): 1753-1758.
- [14] 洪洋, 乔晓君, 白晓东. 基于模板匹配的移动设备手写汉字笔画识别 [J]. 计算机工程与应用, 2017, 53 (20): 105-110.
- [15] McDuff D, Mahmoud A, Mavadati M, et al. AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit [J]. 2016: 3723-3726.
- [16] 潘永之, 李玉和, 刘东岳. 跨平台 BMP 图像处理程序的实现 [J]. 计算机工程与应用, 2007, 43 (14): 84-86.