

工程制图习题答题区域分割算法研究

董志宏, 罗立宏

广东工业大学, 广东 广州

Email: auto_dzh@163.com

收稿日期: 2021年3月23日; 录用日期: 2021年4月18日; 发布日期: 2021年4月25日

摘要

针对工程制图习题册数字化阅卷时不能快速从中分割出学生答题区域的问题, 提出了一种基于投影法的答题区域分割算法。首先对习题图像进行灰度转换、图像二值化等预处理操作, 然后使用Canny算子得到边缘图像, 进行Hough直线变换, 筛选出符合条件的直线, 计算出倾斜角完成图像的倾斜校正; 之后找出边框的矩形轮廓, 结合水平垂直投影法得到的投影特征定位出各小题的图形位置, 实现对答题区域的分割。实验结果表明, 该算法很好地实现了对工程制图习题的答题区域分割, 具有效率高, 准确率高的特点。

关键词

工程制图习题, 图像处理, 图像分割, Hough变换, 投影法

Research on Region Segmentation Algorithm of Engineering Drawing Exercises Based on Hough Transform

Zhihong Dong, Lihong Luo

Guangdong University of Technology, Guangzhou Guangdong

Email: auto_dzh@163.com

Received: Mar. 23rd, 2021; accepted: Apr. 18th, 2021; published: Apr. 25th, 2021

Abstract

In order to solve the problem that the students' question area cannot be quickly segmented from the digital marking of engineering drawing exercise book, an algorithm of question area segmen-

tation based on projection method is proposed. Firstly, preprocess the exercise image such as gray conversion and image binarization, then use Canny operator to get the edge image, carry out Hough line transformation, select the line that meets the conditions, calculate the tilt angle, and complete the tilt correction of the image; then find out the rectangular outline of the frame, and locate the graphic position of each sub problem combined with the projection features obtained by the horizontal and vertical projection method, realize the segmentation of the answer area. The experimental results show that the algorithm can segment the answer area of engineering drawing exercises well, and has the characteristics of high efficiency and high accuracy.

Keywords

Engineering Drawing Exercises, Image Processing, Image Segmentation, Hough Transform, Projection Method

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

工程制图是一门实践性很强的基础课,在理工类院校的教学体系中有着重要的作用[1]。随着信息化时代的发展,很多科目的习题都进行了数字化,其中数字化的方式一般是使用扫描技术将纸质习题试卷转化成图像,然后利用计算机提取出学生答题区域图像,最后交由计算机自动批阅或者教师主动批阅[2]。

对于答题区域的提取,一般采用基于模板的方式分割答题区域。要曙丽[3]等人提出了一种以描述文档为模板的答题区域定位方法;张菲菲[4]等人采用了以图像块作为模板,根据模板图案与答题区域相似性分割出答题区域。采用模板匹配的方法定位答案区域优点在于精度高,图像预处理简单,但缺点是不够灵活,识别精度完全依赖于模板。在更换样式时,不但需要重设模板信息,而且如果模板信息设置不当,那么分割出的区域可能会出现包含题号等干扰信息或者没有涵盖所有的作答信息等情况。针对以上问题,罗朝阳[5]等人采用基于直线的边界检测和区域分割的方法,根据直线边界位置判定答题区域,但该方法仅适用于只包含少量直线的情况;方慧琴[6]等人采用水平投影结合竖直膨胀的方法对答题区域进行智能分割,但该方法仅适用于题目自上向下排列的样式。

对于一般的工程制图习题册,厂家不一定提供专门的模板,并且其中的题型含有练习图形干扰,针对上述情况,本文提出了一种基于投影法的工程制图习题答题区域分割算法,可以在不依赖于模板的同时对不同样式的工程制图习题的答题区域进行智能分割。算法首先对习题图像进行预处理,之后利用Hough变换确定图像的倾斜角度并校正图像,然后提取矩形轮廓结合投影特征完成初分割,定位出答题区域的边界框,最后对初分割图像使用两次投影法进行再分割,提取出最终的答案图像。整体流程如图1所示。

2. 基于 Hough 变换的倾斜校正算法

2.1. Hough 变换

Hough变换由Paul Hough在1962年提出,是如今最常见的直线检测方法之一。算法的基本思想是利用图像空间和参数空间中的点与线的对偶关系,把图像空间中的直线转换到参数空间的点,从而把检测图像空间中直线的问题转化成检测参数空间中交点数量极值的问题[7]。

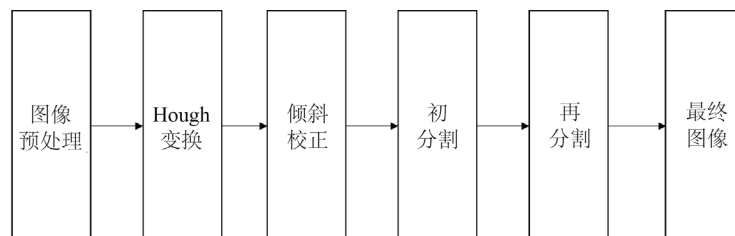


Figure 1. Flow chart of the division of answer area

图 1. 答题区域分割流程图

具体来说, 在 x - y 坐标系中, 一条过点 (x_0, y_0) 直线方程一般记为:

$$y_0 = kx_0 + b \quad (1)$$

其中, 参数 k 表示直线的斜率, 参数 b 表示直线的截距。如果将 x_0 和 y_0 视作常数, 而将参数 k 和 b 视为变量, 则式(1)改写为:

$$b = -kx_0 + y_0 \quad (2)$$

这就将 x - y 坐标系的直线转换成 k - b 坐标系的点。同理, x - y 坐标系的点也可以转换成 k - b 坐标系的直线。如图 2 所示。

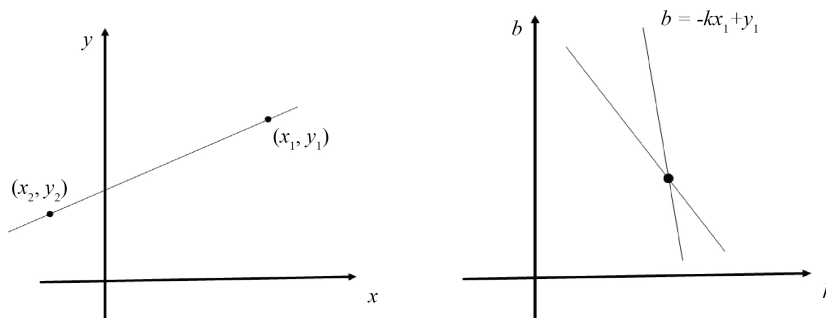


Figure 2. Point line duality relationship between x - y space and k - b space

图 2. x - y 空间和 k - b 参数空间的点线对偶关系

但转换成 k - b 坐标系也存在一些问题, 比如 (k, b) 的取值范围过大以及无法表示垂直直线等[8]。针对这种情况, 可以将 x - y 坐标系对偶到 θ - ρ 参数坐标系, 利用极坐标系代替直角坐标系, 则可将式(1)改写为:

$$\rho = x_0 \cos \theta + y_0 \sin \theta \quad (3)$$

其中, 参数 ρ 表示坐标原点到给定直线的距离, θ 表示直线过原点的垂线与 x 轴的夹角, 取值范围为 $\pm 90^\circ$ 。这样就将 x - y 坐标系上的点与极坐标空间中以 ρ 为波幅的正弦曲线相关联, 并且同一条直线上的点都有相同的 (ρ, θ) 值, 即图像中的每一个像素坐标点对应 θ - ρ 参数空间的一条曲线, 图像中的一条直线对应极坐标系中多条曲线的交点。如图 3 所示。

标准 Hough 变换的流程一般是[9]:

- 1) 创建一个二维矩阵 $M(\rho, \theta)$ 并把其初始值设为 0, 选择一个适当的 θ 值作为量化步长, 然后对每组 (ρ, θ) 值进行累加统计。
- 2) 对图像的每一个像素点进行一次 Hough 变换, 把图像中的点转换成 θ - ρ 参数空间中的正弦曲线集。
- 3) 如果像素点在同一条直线上, 那么这些像素点转换成极坐标系的曲线会相交于一点, 且每一个交点的坐标对应着一组 (ρ, θ) 值, 然后统计经过同一个交点的曲线的数目, 并在矩阵 $M(\rho, \theta)$ 相应位置保存下

来。

4) 设置一个阈值 T , 遍历矩阵 $M(\rho, \theta)$ 中的值, 当矩阵 $M(\rho, \theta)$ 中的值大于阈值 T 时, 就认为匹配到一条有意义的直线, 且该数值所对应的一组 (ρ, θ) 值即为对应的图像中直线的参数。

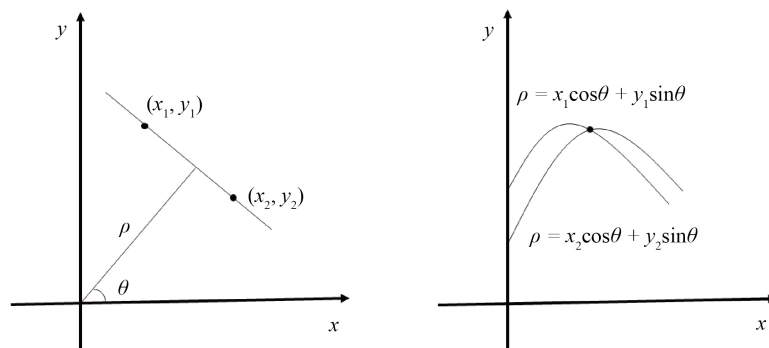


Figure 3. Point line duality relationship between x - y space and θ - ρ space
图 3. x - y 空间和 θ - ρ 参数空间的点线对偶关系

标准的 Hough 变换方法是基于图像的边缘图像实现的, 优点在于算法对图像的噪声、断裂的容错率很高, 但也比较依赖边缘检测算法的准确性。如果图像存在较高的边缘密度, 往往会得到一些异常方向的直线段[10]。同时, 标准 Hough 变换由于需要对图像的每个边缘像素点进行计算, 算法的计算量较大。因此在实际使用时往往需要采取一定的措施减少 Hough 变换的次数[11], 提高算法的效率。

2.2. 图像预处理

在图像采集的过程中, 受到环境和设备的影响, 图像往往携带着很多噪声, 如果不对图像进行一定的处理, 将会影响到后续的分割处理[12]。因此为了减少噪声对结果的干扰, 可以对获取到的习题图像采取预处理操作, 包括灰度变换、图像二值化、高斯滤波等。为了降低算法的计算量, 加快算法检测速度[13], 可以使用高斯金字塔对图像进行一次向下采样处理以减少 Hough 变换检测的像素点, 最后对下采样处理后的图像使用 Canny 算子, 得到习题的边缘图像(图 4)。

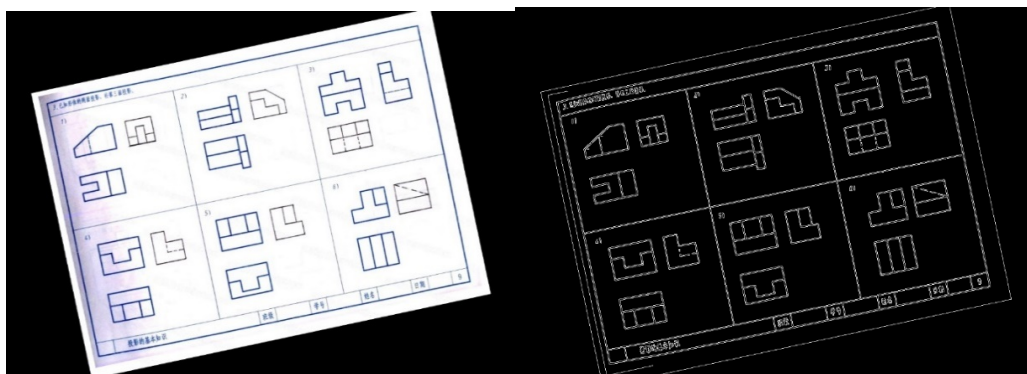


Figure 4. Original image and edge image
图 4. 原图和边缘图像

2.3. 边界直线点筛选

为了进一步提高算法的检测速度, 避免算法对图像中的每一个像素点进行 Hough 变换, 在不影响算

法精度的同时,可以对边缘图像的像素点进行适当的筛选[14],找出可能的边界点,只针对这些边界点进行 Hough 变换。而习题册中每一道习题的答题区域都用一个矩形边界划分出来,这个矩形边界的倾斜角度即是答题区域的倾斜角度。这个矩形边界的直线的像素点一般满足一下两个条件:① 自身为黑色像素;② 如果直线是上边界或者下边界,它同一列的上面一行和下面一行的像素点为白色;如果直线是左边界和右边界,它左边一列和右边一列的像素点为白色(图 5)。

具体实现过程如下:

- 1) 对预处理后的边缘图像从第 i 行开始自左向右对每一个点进行扫描。
- 2) 判断当前像素点 (x_i, y_j) 的值是否为 0;若是,则执行步骤 3);若否,令 $j = j + 1$,继续判断像素点 (x_i, y_j) 的值是否为 0;
- 3) 判断与该像素点位于同一行右一列的像素点 (x_i, y_{j+1}) 的值是否为 1;若像素点 (x_i, y_{j+1}) 为 1,记录像素点的坐标值,然后令 $j = j + 1$,重复步骤 3),直到 j 等于图像的列数;若像素点 (x_i, y_{j+1}) 为 0,则并执行步骤 4);
- 4) 令 $k = i$,判断该像素点位于同一列的上一行的像素点 (x_{k-1}, y_j) 或下一行的像素点 (x_{k+1}, y_j) 的值是否为 1;若是,记录该像素点;之后令 $j = j + 1$ 并执行步骤 2);
- 5) 令 $i = i + 1$,重复步骤 1),直到遍历完图像的行数。

经过上述流程可以得到疑似矩形边界行边缘和列边缘的像素点,然后对这些像素点进行快速 Hough 变换,即可得到直线集。



Figure 5. The straight line after Hough transform
图 5. Hough 变换的直线

2.4. 倾斜校正

经 Hough 变换后可以得到边缘图像中直线集,其中包含多条直线的坐标。由于工程制图习题中存在各类图形,检测到的直线不一定是边界的直线,也可能是习题中图形的直线段,因此需要对检测出的直线按照一定的条件进行过滤。

答题区域边界的直线一定是相互垂直的,而误检直线不一定能找到与其互相垂直的直线。利用这一特征,对所有检测出的直线进行滤波,首先从中筛选出所有符合互相垂直条件的直线簇,并记录这些直线的横坐标和纵坐标。然后对这些直线按照线段中点坐标 (x, y) 的值升序降序排列的规则,提取出最靠近图像边界的直线,利用垂直直线簇的倾斜角度得到整个图像的倾斜角度。最后根据倾斜角度计算出旋转矩阵,并与原图进行矩阵运算即可得到校正图像。经过筛选后的直线效果图以及最终校正效果如图 6。

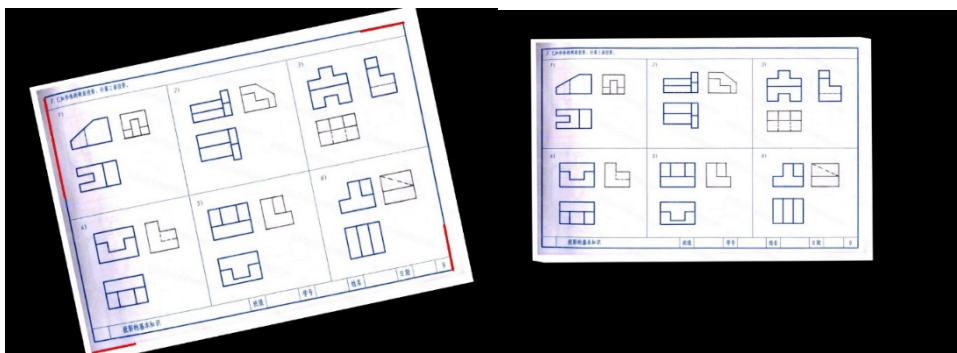


Figure 6. Lines after filtering and Corrected image
图 6. 筛选后的直线和校正图像

3. 基于矩形轮廓和投影法的分割算法

工程制图习题的样式一般有两类, 一类是每道题之间用明确的直线区分, 另一类是每道题之间以大片的空白区域区分(图 7)。

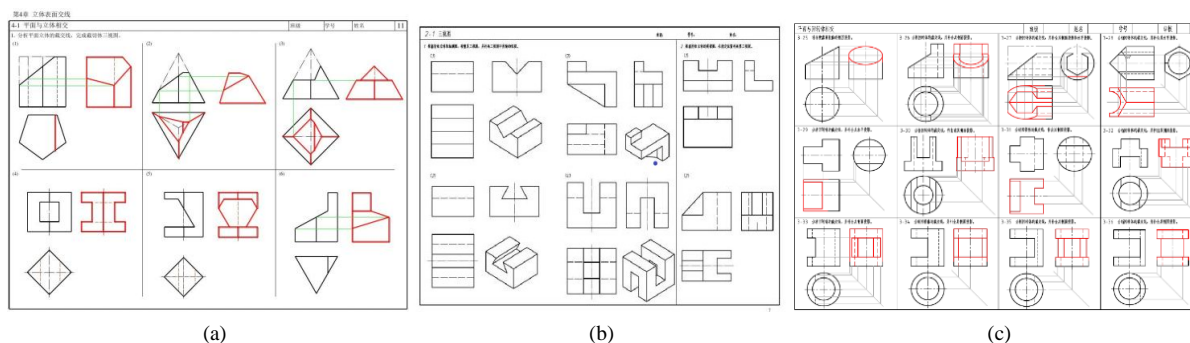


Figure 7. Three exercise styles
图 7. 习题样式(a), (b), (c)

3.1. 矩形轮廓和投影特征的提取

矩形轮廓提取的基本思想是利用 Canny 算子或者 Sobel 算子等边缘增强算子, 突出图像中的局部边缘, 之后通过设置阈值的方法提取边缘点集作为轮廓点集并对各个轮廓拟合最小外接矩形, 最后剔除其中过小和过大的矩形, 留下候选矩形区域。设 $f(i, j)$ 为图像的像素值, 轮廓提取的流程[15]为:

- 1) 对校正后的图像使用 Sobel 算子得到边缘图像;
- 2) 自上到下, 自左到右遍历边缘图像的像素点 $f(i, j)$, 当遇到 $f(i, j) = 1$ 时, 记为起始点;
- 3) 判断 $f(i, j)$ 右边点, 右下点, 下边点, 左下点四个邻点的值是否为 1, 如果其中一个像素点为 1, 则执行步骤 4);
- 4) 判断该邻点的 4 邻域内的点是否不全都是黑点, 若是, 则记录该点位置, 按照逆时针顺序查找该点相邻点, 执行步骤 5), 若否, 该点为内部点, 剔除该点;
- 5) 判断找到的点是否是初始点, 若是, 找到轮廓并保存轮廓; 若否, 执行步骤 2), 直到遍历完整个图像。

在获得图像的所有轮廓后, 可以利用旋转卡尺[16]算法寻找轮廓最小外接矩形。算法原理如下: 首先计算轮廓区域的外接矩形, 以轮廓 X 方向的最大差值作为宽, 轮廓 Y 方向的最大差值作为高, 计算得到

该矩形的面积; 然后对轮廓区域逆时针不断旋转一个角度 θ , 同时记录每一个旋转后的外接矩形的顶点坐标、宽、高和面积; 之后比较所有外接矩形的面积, 获得面积最小的外接矩形以及此时轮廓的旋转角度 a ; 最后将该矩形反方向旋转 a 角度, 即得到最小外接矩形, 记录反向旋转后的矩形顶点坐标。其中, 轮廓点 (x_1, y_1) 绕固定点 (x_0, y_0) 旋转 θ 角度后的位置 (x_2, y_2) 如式(4)所示:

$$\begin{cases} x_2 = (x_1 - x_0)\cos\theta - (y_1 - y_0)\sin\theta + x_0 \\ y_2 = (x_1 - x_0)\sin\theta + (y_1 - y_0)\cos\theta + y_0 \end{cases} \quad (4)$$

下面以样式(b)为例子, 习题答题区域图像、矩形轮廓图像如图 8 所示。

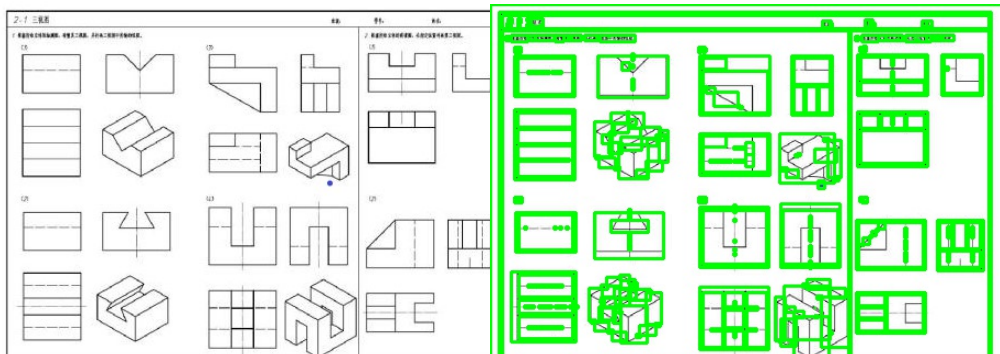


Figure 8. Image of rectangular outline
图 8. 矩形轮廓图

投影法分割是一种基于投影图的方法, 需要先对图像中某一特定方向的像素点个数进行统计, 得到图像的投影图, 之后根据投影特征对图像进行分割[17]。水平投影法指在图像按行向 y 轴方向投影, 统计该方向的黑色像素点的数量, 累加求和后作为 y 轴该行的值, 当遍历完所有行后得到的一维数组, 得到图像的水平投影特征[18]。垂直投影法则是统计按列向 x 轴方向的像素点数量, 获得投影特征。下图 9 为样式(b)水平投影图和垂直投影图。

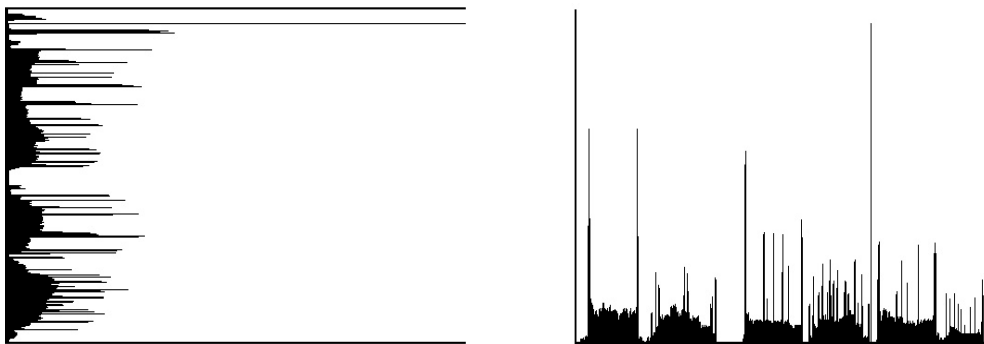


Figure 9. Horizontal and vertical projection
图 9. 水平投影图和垂直投影图

3.2. 矩形轮廓结合投影特征的初分割

由投影图可以看出, 习题框中边界直线位置投影值会比其它位置的投影值大, 可以利用投影特征对矩形轮廓进行筛选, 定位出图形区域的大致位置(图 10)。

- 1) 首先, 设置遍历水平投影得到的一维数组 $X[i]$, 比较数组的值与阈值的大小, 如果 $X[i]$ 大于阈值, 则记录 i 值, 作为分割的纵坐标。
- 2) 遍历垂直投影得到的数组 $Y[j]$, 比较 $Y[j]$ 的值与阈值的大小, 如果 $Y[j]$ 大于阈值, 则记录 j 值, 作为分割的横坐标。
- 3) 纵坐标和横坐标两两组合, 保存为一个点。
- 4) 遍历矩形轮廓, 比较矩形轮廓的左上坐标值 (x, y) 与点 (i, j) 的欧氏距离是否小于预设值, 如果是, 则该轮廓是初分割的区域; 如果否, 则删除该轮廓。

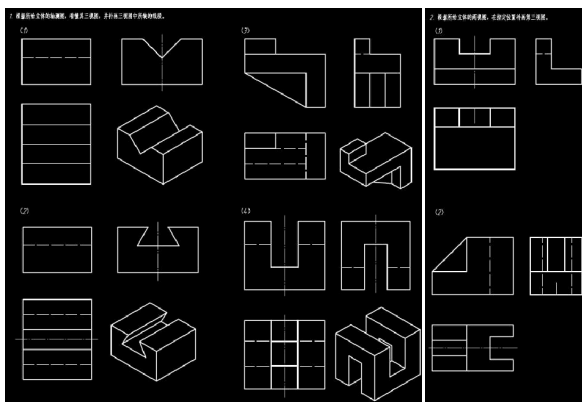


Figure 10. The inverted image of left split image and right split image

图 10. 左边分割图反相图和右边分割图反相图

3.3. 再分割

初次分割后的习题图像仍然存在尚未分割的其它习题和题号, 为了使得最后的分割图像只保留图形区域, 需要进行二次分割。再次使用投影法可以确定两道题之间以及题号和图形间的空白区域。对初分割图像先进行一次水平投影分割, 再进行一次垂直投影分割即可得到只剩下图形的分割图像。水平投影分割流程如下:

- 1) 在初分割的答题区域图像内进行扫描, 当出现投影值 $x[i]$ 不为零时, 判断是否存在起始点 $startindex$, 若 $startindex$ 不存在, 则 $startindex = i$; 若存在, 执行步骤 3);
- 2) 判断当前像素点 $X[i]$ 和上一个像素点 $X[i - 1]$ 是否为零, 若 $X[i]$ 为零且 $X[i - 1]$ 不为零, 则执行步骤 4);
- 3) 判断 $i - endindex$ 是否大于空白区域的大小, 如果是, 则把 $startindex$ 和 $endindex$ 保存。
- 4) 判断 $i - startindex$ 是否小于连块区域的值, 如果是则重置 $startindex$; 否则, $endindex = i$;
- 5) 令 $i = i + 1$, 继续遍历下一个像素点的值 $X[i]$, 直到 i 等于图像的 $rows$;
- 6) 以每一组 $(startindex, endindex)$ 中的 $startindex$ 作为分割上边界, $endindex$ 作为分割下边界, 分割后保存图像。

左边分割图经水平分割后的图像效果如图 11 所示。

垂直投影分割流程与水平分割流程相似, 以经过水平分割后的图像作为输入图像, 最终以左右边界点划分图像。最终分割结果如图 12 所示。

4. 实验对比

本文比较了 50 张不同样式的习题图像, 习题图像分辨率为 500×375 的 JPG 图片, 测试环境为 64

位 Windows10 系统, 处理器型号为 Intel Core i5-8300H, 2.30 Ghz, 内存 4 GB。在 Visual Studio 2017 环境下配置 OpenCV 3.41 运行测试程序。图 13 和图 14 为测试的部分分割结果, 其中图 13 的第一列为正常模板参数, 第二列为模板参数设置过大, 第三列为模板参数过小, 第四列为本文算法; 图 14 第一列为样式(c)正常模板参数, 第二列为复用了样式(b)模板参数, 第三列为本文算法。

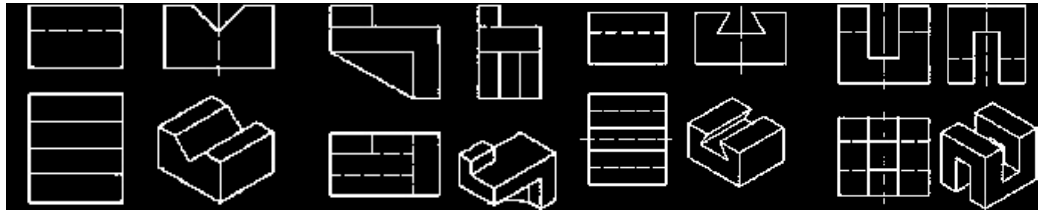


Figure 11. The inverted image of split image by Horizontal projection

图 11. 水平投影分割结果反相图

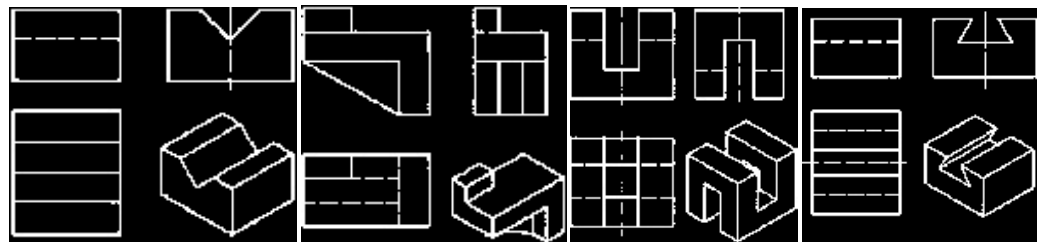


Figure 12. The inverted image of split image by Vertical projection

图 12. 垂直投影分割结果反相图

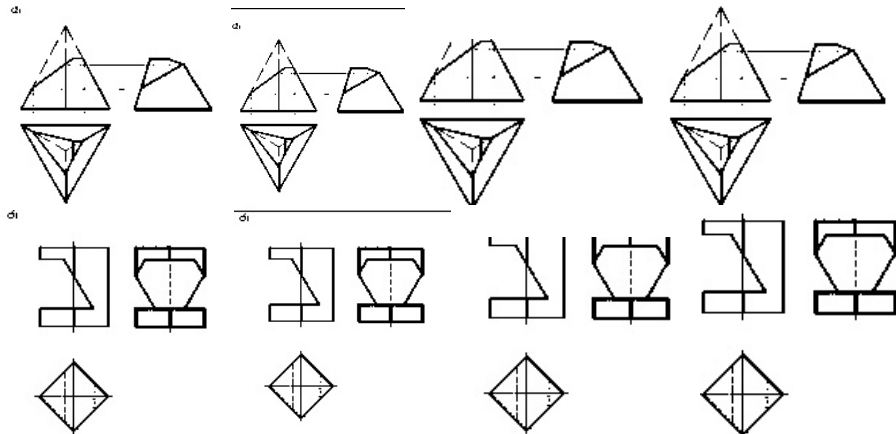
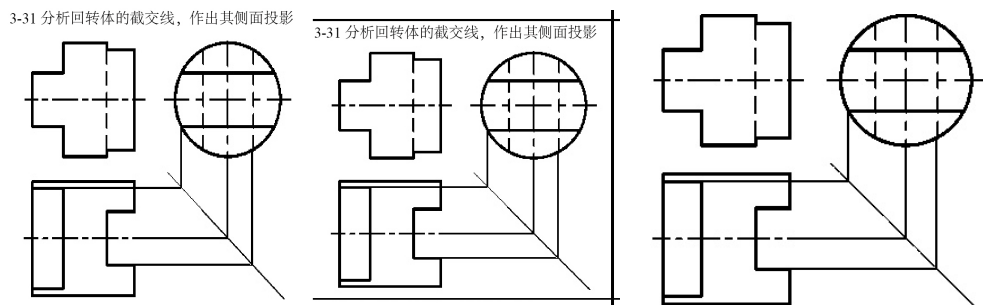


Figure 13. Style (b) segmentation effect comparison

图 13. 样式(b)不同模板参数分割效果比较



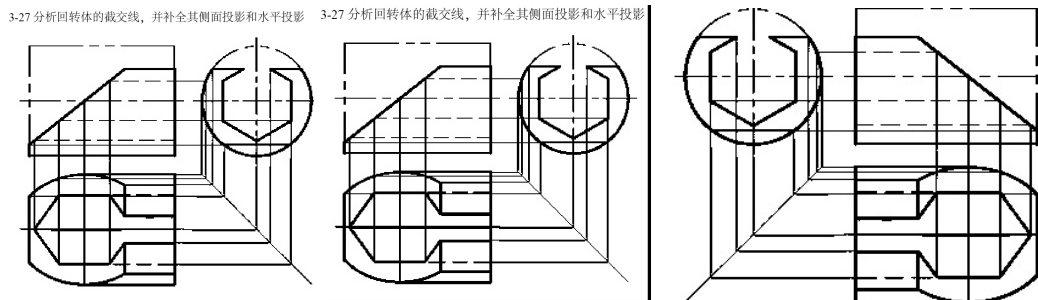


Figure 14. Comparison of splitting effect of style (c) parameter multiplexing
图 14. 样式(c)模板参数复用的分割效果比较

50 份习题图像的整体测试结果如表 1 所示。

Table 1. Test overall result data

表 1. 整体测试结果数据

	文献[3]	文献[6]	本文算法
运行时间	24.695	18.046	20.622
答题区域准确率	100%	96%	100%
信息区域准确率	100%	/	100%
更换时重设参数	是	否	否

由图 13 和图 14 的分割结果可以看出, 由于每道题的图形大小不一致, 基于模板的文献[3]算法在分割答题区域时需要把模板参数尽量大来包含全部的图形, 这样会把题目信息也包含在里面; 其次, 分割需要设置较为准确的参数, 因此同一个模板参数不能适用于其它样式, 普适性较差。本文算法不但可以分割出题号等信息, 把习题中每道题中作答区域的图形准确地分割出来, 而且适用于不同尺寸的样式, 达到了智能分割的效果。同时, 由表 1 可以看出, 基于水平投影的文献[6]分割算法虽然速度较快, 但准确率不够高, 不能正确区分学生信息区域和答题区域。而本文算法在准确率高的同时, 更换样式后也不用重设参数, 速度也比一般的基于模板的文献[3]算法快, 具有高效性。

5. 总结

本文从目前工程制图习题的实际情况出发, 针对工程制图习题册特有的样式提出了基于 Hough 变换的改进的校正算法。该算法可以快速有效地检测出互相垂直的边界线, 然后计算出倾角, 完成对图像的校正。之后, 提出了一种结合矩形轮廓和水平投影的分割算法, 对校正后的习题图像使用投影法得到投影特征, 结合投影特征和矩形轮廓定位出分割的位置信息, 有效地解决了不同样式的习题区域的分割问题, 具有一定的实用价值。

基金项目

广东省科技计划项目(2016A040403110); 教育部人文社科交叉项目(20YJAZH073)。

参考文献

- [1] 赵凯莉, 李鹏, 元立. “工程制图”课程智慧课堂建设与教学模式的探索[J]. 现代信息科技, 2020, 4(21): 195-198.
- [2] 刘淇, 李瑞阳, 李雨盈, 刘佳文. 基于文字识别技术的作业自动批改系统[J]. 电子技术与软件工程, 2020(19):

31-33.

- [3] 要曙丽, 王少荣, 盖孟, 王震. 一种答题卡客观题识别算法[J]. 图学学报, 2019, 40(5): 900-907.
- [4] 张菲菲, 种兰祥. 基于计算机视觉的答题卡识别系统设计与实现[J]. 中国教育信息化, 2019(6): 51-54.
- [5] 罗朝阳, 张鹏超, 姚晋晋, 王彦, 赵瑞杰. 基于 Hough 变换的答题卡识别[J]. 计算机应用与软件, 2020, 37(3): 251-256.
- [6] 方慧琴, 程永强. 阅卷系统中的答题区域快速智能分割算法研究[J]. 现代电子技术, 2018, 41(20): 33-36, 41.
- [7] 郝泽兴, 郭改枝. 基于改进 Canny 算子和 Hough 变换的表格图像校正方法[J]. 内蒙古师范大学学报(自然科学汉文版), 2020, 49(5): 433-441.
- [8] Anagha, P.H. and Baskar, A. (2020) An Automatic Histogram Detection and Information Extraction from Document Images. *International Journal of Speech Technology*, **24**, 77-85. <https://doi.org/10.1007/s10772-020-09756-1>
- [9] Bezmaternykh, P. and Nikolaev, D.P. (2020) A Document Skew Detection Method Using Fast Hough Transform. *12th International Conference on Machine Vision*, Amsterdam, 16-18 November 2019, Article ID: 114330J. <https://doi.org/10.1117/12.2559069>
- [10] 郑行家, 钟宝江. 图像直线段检测算法综述与测评[J]. 计算机工程与应用, 2019, 55(17): 9-19.
- [11] 王琦, 宋伟东, 王竞雪. 基于 Hough 变换的直线提取方法及改进[J]. 测绘与空间地理信息, 2019, 42(6): 214-217+221.
- [12] Chettat, M., Gaceb, D. and Belhadi, S. (2020) A Fast High Precision Skew Angle Estimation of Digitized Documents. *International Arab Journal of Information Technology*, **17**, 825-831. <https://doi.org/10.34028/iajit/17/5/16>
- [13] Oba, B., Wkh, A. and Dm, B. (2020) Using Skeleton and Hough Transform Variant to Correct Skew in Historical Documents. *Mathematics and Computers in Simulation*, **167**, 389-403. <https://doi.org/10.1016/j.matcom.2019.05.009>
- [14] Ramanan, M. (2019) A Hybrid Approach for Skew Detection and Correction in the Multi-Script Scanned Document. *Asian Journal of Research in Computer Science*, **4**, 1-8. <https://doi.org/10.9734/ajrcos/2019/v4i230112>
- [15] Martorell, O., Buades, A. and Lisani, J.L. (2021) Multiscale Detection of Circles, Ellipses and Line Segments, Robust to Noise and Blur. *IEEE Access*, **9**, 25554-25578. <https://doi.org/10.1109/ACCESS.2021.3056795>
- [16] 周再文, 王建, 朱恰, 刘星雨, 马紫雯, 高贤君. 基于最小外接矩形的遥感影像建筑物轮廓优化方法[J]. 北京测绘, 2021, 35(1): 1-6.
- [17] Babczyński, T. and Ptak, R. (2020) Line Segmentation of Handwritten Text Using Histograms and Tensor Voting. *International Journal of Applied Mathematics and Computer Science*, **27**, 195-206.
- [18] 王莉丽, 陈晔, 刘玲. 基于投影轮廓分析的文本图像版面分割算法研究[J]. 数字技术与应用, 2017(3): 164-165.