

一种基于水族濒危文字的图像增强及识别方法

杨秀璋¹ 夏 换² 于小民²

(贵州财经大学信息学院 贵阳 550025)¹

(贵州财经大学贵州省经济系统仿真重点实验室 贵阳 550025)²

摘 要 随着图形图像处理技术的迅速发展,图像增强及识别方法已广泛应用于各行各业。在此基础上,文字识别技术也取得了极大的进步。针对水族文字笔触随意、字形多变、噪声较多等问题,文中提出了一种改进的图像增强及识别方法。通过中值滤波算法降低图像噪声,利用直方图均衡化方法增强图像对比度,再经过二值化处理提取图像中的目标轮廓,通过腐蚀膨胀处理细化和扩张背景,最后采用改进的文字提取算法凸显水族文字,采用 Sobel 算子提取水族文字边缘,并对其进行仿真对比实验。实验结果表明,该方法有效地降低了图像噪声,准确地提取出了水族文字轮廓,可以应用于民族文字提取及识别、文物修复、图像增强等领域,对保护民族文物遗产、弘扬少数民族传统文化具有重要意义。

关键词 图像增强,图像识别,水族文字识别,直方图均衡化,文字提取

中图分类号 TP391 文献标识码 A

Image Enhancement and Recognition Method Based on Shui-characters

YANG Xiu-zhang¹ XIA Huan² YU Xiao-min²

(School of Information, Guizhou University of Finance and Economics, Guiyang 550025, China)¹

(Guizhou Key Laboratory of Economics System Simulation, Guizhou University of Finance and Economics, Guiyang 550025, China)²

Abstract With the rapid development of graphic image processing technology, image enhancement and recognition methods have been widely used in various industries. On this basis, text recognition technology has also made great progress. Aiming at the problems of shui text random brush strokes, variable fonts and more noise, this paper proposed an improved image enhancement and recognition method. The median filtering algorithm is used to reduce image noise, and the histogram equalization method is used to enhance image contrast. The binarization process is executed to extract the target text in the image, and the corrosion expansion process is executed to refine and expand the background. Finally, the improved text extraction algorithm is used to highlight the outline of the shui text, and the Sobel operator is used to extract the edge of the shui text. The simulation contrast experiment was carried out. The experimental results show that the method effectively reduces image noise, and accurately extracts shui characters. The method can be used in the fields of national character extraction and recognition, cultural relics restoration, image enhancement, etc. It is of great significance for protecting the heritage of ethnic cultural relics and carrying forward the traditional culture of ethnic minorities.

Keywords Image enhancement, Image recognition, Shui-character recognition, Histogram equalization, Text extraction

1 引言

水族是一个具有悠久历史和古朴文化的民族,繁衍生息于贵州、云南、广西等省份,拥有独特的民族文化、风土人情和语言文字,具有重要的社会历史文化价值^[1]。水书是水族的文字,水族语言称其为“泐睢”,由汉字改编而成,其形状类似甲骨文和金文,主要用来记载水族的天文、地理、宗教、民俗、哲学等文化信息^[2]。水族文字图像从各层面反映了各个历史

时期的水族群体的社会活动、家族关系、意识形态以及演化趋势,是人类宝贵的历史文化遗产。近年来,学术界和社会媒体掀起了研究水族文化和水书的热潮,但针对水族文字识别的研究还很少,水族文字图像受自然侵蚀和人为破坏,文物遗留的信息已经有所缺失。为了更好地推进水族文字研究,修复水书文物图像,传承和保护水族文化遗产,本文提出了一种改进的图像增强及识别方法,用于提取濒危的水族文字^[3]。

随着数字图像处理技术的迅速发展,图像增强和图像识

本文受贵州省教育厅青年科技人才成长项目(黔教合 KY 字[2016]172,黔教合 KY 字[2016]178),贵州省普通高等学校科技拔尖人才支持计划项目(黔教合 KY 字[2016]068)资助。

杨秀璋(1991—),男,硕士,助教,主要研究方向为 Web 数据挖掘、图像识别、知识图谱,E-mail:1455136241@qq.com;夏 换(1981—),男,博士,教授,主要研究方向为计算机仿真、大数据分析,E-mail:66374769@qq.com(通信作者)。

别方法已广泛应用于各行各业,比如车牌识别、人脸识别、文字提取、图像修复等。图像识别(Image Recognition)是通过算法和函数提取像素中的某些特征,并对图像进行识别和分类的过程。当前国内外学者对其的研究取得了一定进展,在方法和实践上都有不少的成果。王积分等^[4]详细概述了计算机图像识别的各种技术,包括图像变换、阈值处理、边缘检测、图像分割、图像识别等;王月新等^[5]分析研究了 Sobel 算子和 Prewitt 算子;李晓飞等^[6]对图像腐蚀和图像膨胀算法进行了对比研究;李振宏等^[7]提出了蒙古文文字特征提取及模式分类的识别方法;王琪等^[8]提出了一种改进的小波阈值图像去噪方法,有效地去除了噪声干扰;刘军等^[9]提出了一种基于高斯金字塔和梯度方向直方图的车牌模糊汉字识别方法;李雄等^[10]结合 Roberts 梯度算子和几何特征形态学识别灰度车牌;刘磊^[11]针对秦汉文字瓦当特有的结构,提出了一种适用于秦汉文字瓦当的区域消除算法和文字分割算法;黄攀^[12]提出了基于深度学习的自然场景文字识别算法。

然而实际应用中,自然侵蚀、人为破坏、恶劣天气、拍摄角度、光照不均匀、背景复杂、文字污损等因素,严重降低了文字识别的准确率。针对这些问题,本文基于水族文字笔触随意、字形多变、噪声较多等特点,并结合实际应用场景,提出了一种改进的图像增强及识别方法,用于提取水族文字。该方法通过中值滤波算法降低图像噪声,利用直方图均衡化方法增强图像对比度,再经过二值化处理提取图像轮廓,腐蚀膨胀处理细化和扩张背景,最后采用改进的文字提取算法凸显水族文字轮廓,利用 Sobel 算子提取水族文字边缘。本文的研究成果具有重要的理论研究意义和良好的实际应用价值,可以广泛应用于民族文字提取、文物修复、图像识别等领域,对传承、弘扬和保护水族传统文化具有重要意义,为水族文化的研究和进一步发展提供相关启示,为后续的水族文字识别、水族迁徙研究和水族文化传承提供有效支撑。

2 算法的整体结构设计

一个完整的水族文字增强识别系统是一个复杂的系统,主要包括图像灰度处理、图像平滑、图像增强、阈值分割、形态学处理、文字提取、边缘检测、锐化处理等步骤。本文提出了一种改进的图像增强及识别方法,其框架如图 1 所示。

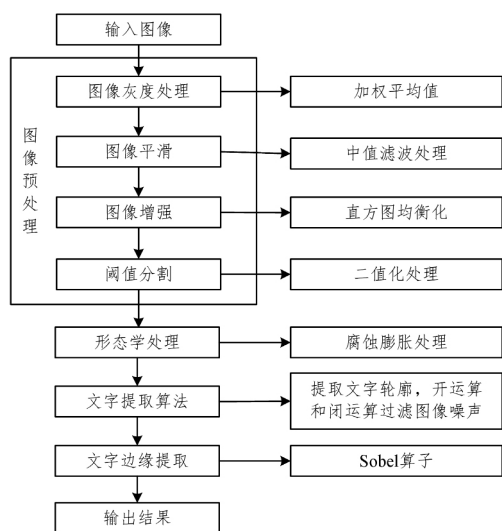


图 1 水族文字识别框架图

1) 读取原始数据,经过图像预处理操作,使用加权平均法

将彩色图像转换为灰度图像,利用中值滤波算法降低图像噪声,通过直方图均衡化方法扩展灰度值的范围,再经过二值化处理提取轮廓。

2) 通过形态学腐蚀处理和膨胀处理细化和扩张水族文字图像背景。

3) 最后采用改进的文字提取算法凸显水族文字轮廓,利用 Sobel 算子锐化突出图像的边缘细节,改善图像的对比度,使模糊的图像变得更清晰,并输出识别的结果。

3 水族文字识别方法的具体实现

3.1 灰度处理

图像灰度转化是将彩色图像转换为灰度化图像的过程。灰度图像中每个像素仅具有一种样本颜色,其灰度是位于黑色与白色之间的多级色彩深度。式(1)是本文所采用的灰度处理公式。

$$Gray = R * 0.299 + G * 0.587 + B * 0.114 \quad (1)$$

其中, R, G, B 分别为图像中每个像素点的三原色即红、绿、蓝分量; $Gray$ 为灰度处理后该像素点的灰度值,它将原始 $RGB(R, G, B)$ 颜色均匀地替换成新颜色 $RGB(Gray, Gray, Gray)$,从而将彩色图像转化为灰度图像。由于人类的眼睛感官蓝色的敏感度最低,感官绿色的敏感度最高,因此将 RGB 三分量按照 $0.299, 0.587, 0.114$ 比例加权求平均,从而得到较合理的灰度图。

3.2 中值滤波处理

中值滤波(Median Blur)是一种非线性平滑方法,能够在消除噪声的同时保留边界信息。该方法用每一个像素点某邻域范围内所有像素点灰度值的中值来替换该像素点的灰度值,从而让周围的像素值更接近真实情况,消除孤立的噪声。

中值滤波算法的计算过程如图 2 所示。选择含有 5 个点的窗口,依次扫描该窗口中的像素,每个像素点所对应的灰度值按照升序或降序排列,然后获取最中间的值来替换该点的灰度值。

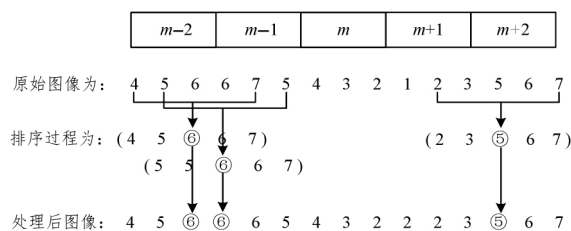


图 2 中值滤波算法的计算过程

图 2 展示的是矩形窗口,常用的窗口还包括正方形、十字形、环形和圆形等,不同形状的窗口会带来不同的过滤效果,其中正方形和圆形窗口适合于外轮廓边缘较长的图像,十字形窗口适合于带尖角形状的图像。

处理后的结果如图 3 所示,其中图 3(a)表示带噪声的原始图像,图 3(b)表示中值滤波处理后的图像,它有效地过滤掉了图像的噪声。

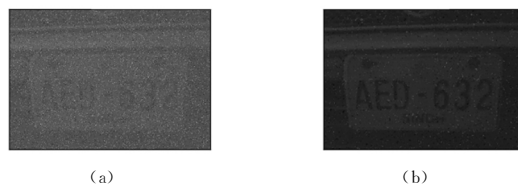


图 3 中值滤波处理

3.3 直方图均衡化

直方图均衡化是图像灰度变化的一种重要处理方式,被广泛应用于图像增强领域。它是指通过某种灰度映射将原始图像的像素点均匀地分布在每一个灰度级上,其结果将产生一幅灰度级分布概率均衡的图像。直方图均衡化的中心思想是把原始图像的灰度直方图从比较集中的某个灰度区间转变为全范围均匀分布的灰度区间,通过该处理,增加了像素灰度值的动态范围,从而达到增强图像整体对比度的效果^[13]。该算法的处理过程如下。

首先,计算原始图像直方图的概率密度,如式(2)所示。

$$P(r_k) = \frac{n_k}{N} \quad (2)$$

其中, r_k 表示第 k 个灰度级, $k=0,1,2,\dots,L-1$; L 最大值为 256; n_k 表示图像中灰度级为 r_k 的像素个数; N 表示图像中像素的总个数; $P(r_k)$ 为图像中第 k 个灰度级占总像素数的比例。

其次,通过灰度变换函数 T 计算新图像灰度级的概率密度。新图像灰度级的概率密度是原始图像灰度级概率密度的累积,如式(3)所示。

$$s_k = T(r_k) = \sum_{j=0}^k P(r_j) \quad (3)$$

其中, k 是新图像的灰度级, $k=0,1,2,\dots,L-1$; r_k 表示原始图像的第 k 个灰度级; s_k 为直方图均衡化处理后的第 k 个灰度级。

最后,计算新图像的灰度值。由于式(3)计算所得 s_k 的取值位于 0 至 1 之间,需要乘以图像的最大灰度级 L ,将其转换为最终的灰度值 res ,如式(4)所示。

$$res = s_k \times L \quad (4)$$

图 4 表示原始图像及其直方图,图 5 表示直方图均衡化后的图像及其直方图。从效果图可以看出,经过直方图均衡化处理,图像变得更加清晰,图像的灰度级分布也更加均匀。

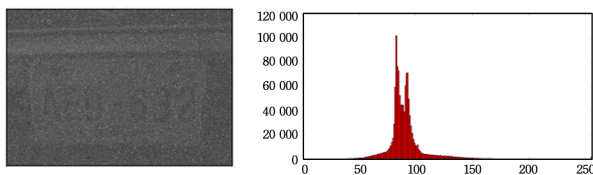


图 4 原始图像及其直方图

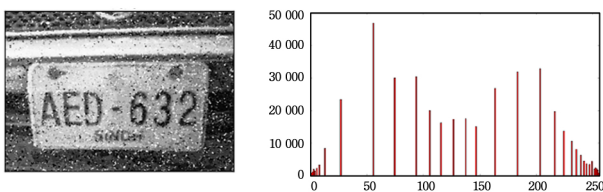


图 5 直方图均衡化后的图像及其直方图

3.4 二值化处理

图像二值化(Binarization)旨在提取图像中的目标物体,将背景以及噪声区分开来。通常会设置一个临界值,将图像的像素划分为两类,常见的二值化算法如式(5)所示。

$$Gray(i,j) = \begin{cases} 255, & Gray(i,j) \geq T \\ 0, & Gray(i,j) < T \end{cases} \quad (5)$$

当某个像素点的灰度 $Gray(i,j)$ 小于阈值 T 时,其像素设置为 0,表示黑色;当灰度 $Gray(i,j)$ 大于或等于阈值 T 时,其像素值为 255,表示白色。

Python OpenCV 中提供了阈值函数 `threshold()` 来实现二值化处理,当像素大于或等于 127 时,灰度值设置为最大值 255,小于 127 的像素点灰度值设置为 0。二值化处理为后续的水族文字识别提供了有效支撑。图 6 表示图像二值化处理后的效果图。



图 6 二值化处理结果

3.5 腐蚀处理和膨胀处理

图像的腐蚀(Erosion)和膨胀(Dilation)是两种基本的形态学运算,主要用来寻找图像中的极小区域和极大区域。其中,腐蚀类似于“领域被蚕食”,它将图像中的高亮区域或白色部分进行缩减细化,其运行结果比原图的高亮区域更小;膨胀类似于“领域扩张”,它将图像中的高亮区域或白色部分进行扩张,其运行结果比原图的高亮区域大^[14]。

腐蚀的运算符是“ $-$ ”,其定义如下:

$$A - B = \{x | B_x \subseteq A\} \quad (6)$$

该公式表示图像 A 用卷积模板 B 来进行腐蚀处理,通过模板 B 与图像 A 进行卷积计算,得出 B 覆盖区域的像素点最小值,并用这个最小值来替代参考点的像素值。如图 7 所示,将左边的原始图像 A 腐蚀处理为右边的效果图 $A - B$ 。

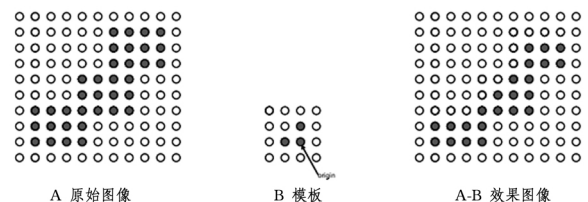


图 7 腐蚀处理的原理图

膨胀的运算符是“ $+$ ”,其定义如下:

$$A + B = \{x | B_x \cap A \neq \emptyset\} \quad (7)$$

该公式表示用 B 来对图像 A 进行膨胀处理,其中 B 是一个卷积模板,其形状可以为正方形或圆形。通过模板 B 与图像 A 进行卷积计算,扫描图像中的每一个像素点,用模板元素与二值图像元素做“与”运算,如果都为 0,那么目标像素点为 0,否则为 1。依此计算 B 覆盖区域的像素点最大值,并用该值替换参考点的像素值,实现图像膨胀。图 8 是将左边的原始图像 A 膨胀处理为右边的效果图 $A + B$ 。

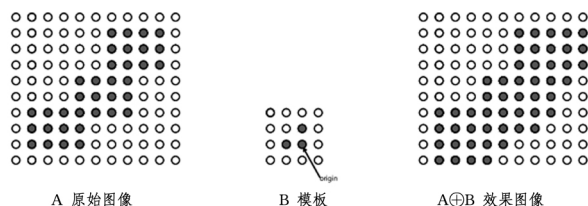


图 8 膨胀处理原理图

图 9 和图 10 是经过膨胀和腐蚀处理后的效果图,图像经过腐蚀处理过滤掉文字部分,再通过膨胀处理提取图像的背景轮廓。

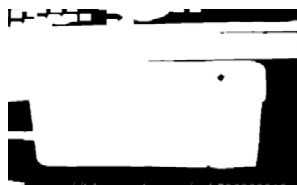


图 9 图像腐蚀处理



图 10 图像膨胀处理

3.6 文字提取算法

经过图像腐蚀和膨胀处理后,得到如图 10 所示的效果图,它将图像的背景轮廓提取出来,接着利用文字提取算法识别目标文字。算法流程如下:

- 1) 用腐蚀膨胀处理后的图像减去二值化处理的图像,提取图像中的文字轮廓;
- 2) 通过开运算处理去除图像的噪声并保留文字内容;
- 3) 通过闭运算处理过滤掉图像文字中的黑点噪声,提取出更清晰的文字轮廓。

图 11 是图像文字提取处理过程中每个步骤对应的效果图,其中(a)是原始图像;(b)是中值滤波处理后的图像;(c)是直方图均衡化处理后的图像;(d)是二值化处理后的效果图;(e)表示提取图像的文字内容;(f)表示经过开运算和闭运算过滤掉噪声后的效果图,可以清晰地看到文字“AED-632”。



图 11 图像文字提取处理

3.7 Sobel 算子

图像数据在收集或传输过程中受一些外界因素影响,会使得图像存在模糊和有噪声的情况,从而影响图像识别工作。此时,需要通过图像锐化处理,加强原图像的高频部分,锐化突出图像的边缘细节,从而改善图像的对比度,使模糊的图像变得更加清晰。

Sobel 算子是一种计算图像明暗程度近似值的差分算子,它根据图像边缘附近的明暗程度把该区域内超过某个值的特定点记为边缘。Sobel 算子根据相邻点的距离远近为当前像素点赋予不同的权重,距离越近的像素点对当前像素的影响

越大,反之越小,从而实现图像锐化并突出边缘轮廓。Sobel 算子的边缘定位更加准确,常用于噪声较多、灰度渐变的图像^[14]。其模板如式(8)所示:

$$d_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, d_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (8)$$

其中, d_x 表示水平方向, d_y 表示垂直方向,对应的像素计算公式如下:

$$\begin{aligned} d_x(i, j) &= [f(i-1, j-1) + 2f(i-1, j) + f(i-1, j+1)] - \\ &\quad [f(i+1, j-1) + 2f(i+1, j) + f(i+1, j+1)] \\ d_y(i, j) &= [f(i-1, j+1) + 2f(i, j+1) + f(i+1, j+1)] - \\ &\quad [f(i-1, j-1) + 2f(i, j-1) + f(i+1, j-1)] \end{aligned} \quad (9)$$

Sobel 算子的最终计算如式(10)所示:

$$S = \sqrt{d_x(i, j)^2 + d_y(i, j)^2} \quad (10)$$

图 12 为 Sobel 算子的处理结果,其考虑了各方面因素,对噪声较多的图像处理效果更好。



图 12 Sobel 算子提取边缘轮廓

4 实验及分析

本文提出了一种改进的图像增强及图像识别方法并将其用于濒危水族文字的抢救中,该方法经过灰度转换、中值滤波去噪、直方图均衡化、二值化处理、腐蚀膨胀处理、文字提取算法、Sobel 算子识别图像内容。为了客观地评价本文提出的方法,实验采集了具有代表性的水书文物图像,并应用该方法进行水族文字图像识别实验。

图 13 是对具有一定年份的水书刺绣进行图像识别实验,其中图(a)是水书刺绣原始图像,图(b)是灰度图像,图(c)表示中值滤波去除噪声后的图像,图(d)是直方图均衡化处理,图(e)是二值化处理提取轮廓,图(f)文字提取算法处理后的图像,图(g)表示腐蚀和膨胀处理,图(h)是去除背景噪声后的图像,图(i)是经过 Sobel 算子锐化提取水族文字轮廓。

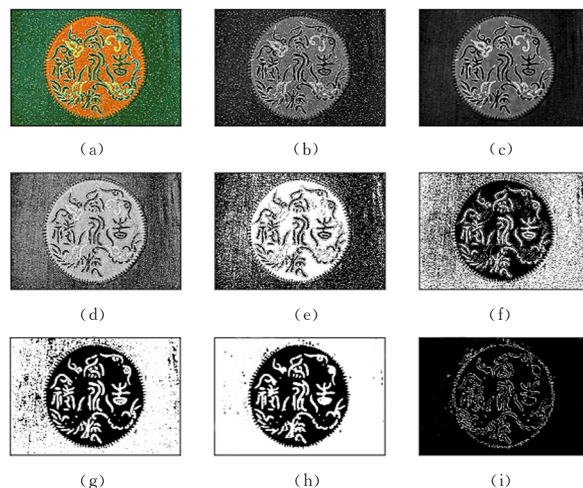


图 13 水族刺绣中的文字识别

图 14 是一张包含水族文字的图像,照片比较陈旧、噪声较多;图 15 是采用本文提出的算法进行图像识别后的结果,它清晰地识别出了水族文字,过滤了图片中的噪声。

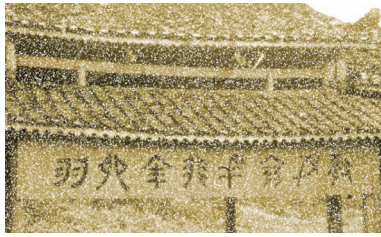


图 14 识别前的水族文字建筑图像



图 15 识别后的水族文字建筑图像

接着对水书文字图像进行实验,实验结果如图 16 所示,其中图(a)是原始图像,图(b)是图像增强识别后的效果图。通过实验处理,可以有效地提取出水族文字,消除噪声,凸显轮廓。



(a)



(b)

图 16 水书文字的图像增强识别前后对比

结束语 本文的研究成果主要应用于民族文字图像识别、文物修复以及图像增强领域,对传承、弘扬和保护民族文化具有重要意义。本文采用灰度转化将彩色图像转换为灰度图,通过中值滤波算法和直方图均衡化算法降低图像噪声,增强图像对比度,再经过二值化处理提取图像轮廓,腐蚀膨胀处理细化和扩张背景,最后采用改进的文字提取算法凸

显水族文字轮廓,并利用 Sobel 算子提取水族文字边缘。

作者采用 Python 语言和 OpenCV 技术对实际场景中的水族文字图像进行仿真实验,证明了该算法的有效性和较强的环境适应性,它能降低噪声,有效提高水族文字识别的准确率。该技术可以广泛应用于民族文字提取、文物修复、图像识别等领域,为实际的民族文字识别提供了依据,具有一定的应用前景和实用价值。本文为水族文化的研究和进一步发展提供了相关启示,为后续的水族文字识别、水族迁徙研究和水族文化传承提供了有效支撑。下一步计划结合深度学习技术实现水族文字与汉字的自动翻译,同时采用并行计算方法优化该算法的运行效率。

参考文献

- [1] 饶文谊,梁光华.关于水族水书起源时代的学术思考[J].原生态民族文化学刊,2009(4):90-93.
- [2] 廖崇虹.水书文化开发性保护的产业化发展战略研究[D].天津:天津大学,2010.
- [3] 王展,王慧琴,吴萌,等.数字图像修复技术在文物领域的应用展望[J].文博,2012(3):79-82.
- [4] 王积分,张新荣.计算机图像识别[M].北京:中国铁道出版社,1988.
- [5] 王月新,刘明君.sobel 算子与 prewitt 算子分析与研究[J].计算机与数字工程,2016(10).
- [6] 李晓飞,马大玮,粘永健,等.图像腐蚀和膨胀的算法研究[J].影像技术,2005(1):37-39.
- [7] 李振宏,高光来,侯宏旭,等.印刷体蒙古文文字识别的研究[J].内蒙古大学学报(自然科学版),2003,34(4):454-457.
- [8] 王琪,程彬,杜娟,等.一种改进的小波阈值图像去噪方法[J].计算机与现代化,2015(4):65-69.
- [9] 刘军,白雪.基于梯度方向直方图与高斯金字塔的车牌模糊汉字识别方法[J].计算机应用,2016,36(2):586-590.
- [10] 李雄,裴承鸣,郑华.几何特征形态学车牌识别系统研究[J].计算机仿真,2012,29(4):353-356.
- [11] 刘磊.基于内容的秦汉瓦当小篆文字识别方法研究[D].西安:西北大学,2015.
- [12] 黄攀.基于深度学习的自然场景文字识别[D].杭州:浙江大学,2016.
- [13] 图像处理-python 实现图像处理(消噪,直方图均衡化,二值化,形态学)[EB/OL]. <https://blog.csdn.net/wsh596823919/article/details/79982485>.
- [14] Python 图像处理及图像识别[EB/OL]. <https://blog.csdn.net/Eastmount/column/info/26274>.