

基于多维侧窗聚类分块的退化书法文档二值化^①



徐占洋¹, 张家瑞², 侍虹言¹, 秦飞扬¹, 林巍³

¹(南京信息工程大学 软件学院, 南京 210044)

²(南京信息工程大学 计算机学院, 南京 210044)

³(江苏省少儿春互联教育科技有限公司 南京研发中心, 南京 210031)

通信作者: 张家瑞, E-mail: zjr_1368@163.com

摘要: 书法字文档图像在不良光照条件下的灰度值分布差异较大, 低光照区域图像对比度较低、笔画形态纹理特征出现退化, 传统方法通常仅考虑了局部信息的均值、平方差、熵等因素, 在形态纹理方面考虑较少, 从而对低对比度区域的特征信息不敏感. 针对此类问题, 本文提出了一种多维侧窗聚类分块的退化书法文档的二值化方法 CS-SWF (clustering segmentation based SWF), 该方法首先利用 SWF 卷积核描述具有相似形态学特征的像素块, 之后提出多种修正规则利用下采样提取低纬度信息去修正特征区域. 最后, 对特征图中聚类块进行前后景分离, 得到二值化结果图. 本文使用 FM 、 $PSNR$ 和 DRD 为指标, 将现有方法和本文方法进行对比, 实验结果表明, 在自建的 100 张手写退化文档图像数据集下, 本文方法在低对比度暗部区域的二值化效果较为稳定, 在精准度和鲁棒性上优于对比算法.

关键词: 自适应二值化; 不均匀照明; 侧窗滤波器; 退化文档

引用格式: 徐占洋, 张家瑞, 侍虹言, 秦飞扬, 林巍. 基于多维侧窗聚类分块的退化书法文档二值化. 计算机系统应用, 2024, 33(6): 223-231. <http://www.c-s-a.org.cn/1003-3254/9526.html>

Degraded Calligraphic Document Binarization Based on Multidimensional Side Window Clustering Segmentation

XU Zhan-Yang¹, ZHANG Jia-Rui², SHI Hong-Yan¹, QIN Fei-Yang¹, LIN Wei³

¹(School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China)

²(School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China)

³(Nanjing Technology R & D Center, Jiangsu Shao Er Chun Internet Education Technology Co. Ltd., Nanjing 210031, China)

Abstract: The distribution of grayscale values in calligraphic character document images exhibits significant variations under poor lighting conditions, resulting in lower image contrast in low-light areas and degradation of morphological texture features of the strokes. Traditional methods typically focus on local information such as mean, squared deviation, and entropy, while giving less consideration to morphological texture, rendering them insensitive to the features of low-contrast areas. To address these issues, this study proposes a binarization method called clustering segmentation-based side-window filter (CS-SWF) specifically designed for degraded calligraphic documents. Firstly, this method utilizes multi-dimensional SWF to describe pixel chunks with similar morphological features. Then, with multiple correction rules, it utilizes downsampling to extract low-latitude information and correct feature regions. Finally, the clustered blocks in the feature map are classified to obtain the binarization results. To evaluate the performance of the proposed method, it is compared with existing methods using F -measure (FM), peak signal-to-noise ratio ($PSNR$), and distance reciprocal distortion (DRD) as indicators. Experimental results on a self-constructed dataset consisting of 100 handwritten degraded

① 基金项目: 中小学书法教学智能评价平台项目 (SRC202201)

收稿时间: 2023-12-20; 修改时间: 2024-01-17; 采用时间: 2024-01-29; csa 在线出版时间: 2024-04-19

CNKI 网络首发时间: 2024-04-23

document images demonstrate that the proposed binarization method exhibits greater stability in low-contrast dark regions and outperforms the comparison algorithm in terms of accuracy and robustness.

Key words: adaptive binarization; uneven illumination; side window filtering (SWF); degraded document

随着中国书法日益受到社会各界和教育部门的重视,为弘扬和传承我国书法文化,教育部门积极颁布了《关于中小学开展书法教育的意见》^[1]和《关于实施中华优秀传统文化传承发展工程的意见》^[2]等文件,书法教育也逐步深入中小学教学体系.然而,受书法教育特殊性影响,使得书法教育资源明显短缺与分配不均.为应对现状,硬笔临帖书法字的智能评价已成为重要的辅助手段,具有广泛的应用前景.

因此,本文联合书法教学公司对硬笔临帖书法字的智能评价系统进行了深入研究,其目的在于开发一款可以高效、客观、稳定的智能评价系统,快速识别和分析学生的书法作品,并给出具体的改进意见,使学生能够有针对性地进行练习和提高.

首先本课题针对中学生日常书法练习的方式进行调研,初学者大都会使用常见的辅助习字格练习纸来书写汉字进行练习,而后续处理中,可以通过拍照设备将练习纸上传到智能书法评价系统,系统通过对练习纸的分割、二值化等处理得到用户在练习纸上的单个书法字切片图,从而利用智能评价模块进行客观评价并给出书写建议从而高效的帮助用户提升书法水平.而在系统的图像预处理步骤中,书法字的提取、二值化在后续的汉字骨架细化、笔画拆分、笔画评价中占有决定性地位.为了防止用户的拍摄图片中,因设备或光源等问题导致的退化现象影响到后续的步骤,因此,退化文档的二值化是本文研究的重点.

在退化文档图像处理的研究中,光线分布不均是最常出现的一种退化类型,是由光源条件以及环境因素等多种原因而导致的.近年来也有学者不断致力于研究此类问题,一般最常见的方法分为全局和局部方法.全局方法的做法是找到一个全局阈值,依据图片灰度值大小分类前景与背景. Otsu^[3]所提出的经典算法在光线均匀情况下,具有不错的表现.由于退化文档的特殊性,全局法往往无法处理图像中的部分区域,因此,局部阈值被提出,其做法是在图像的局部区域选取一个阈值,依次参照图像的局部区域信息对中心像素进行分类,取得了较好的准确性与鲁棒性.当前有一些基

于 Niblack 或 Sauvola 的改进算法^[4-6]相继提出,无论是利用神经网络获取最佳参数的思想,还是动态计算窗口大小的方法,都取得了一定的效果,文献^[7-11]也采用了类似的方法,但是此类方法依旧无法处理图像中某些对比度较低的局部区域,且往往还需要去不断手动调整参数来获得较好的效果.

同时,为了综合全局和局部方法的优点,一种基于混合阈值的方法也被相继提出,刘朋远等人^[12]通过改进的局部阈值算法结合全局阈值进行加权,最终得到一个自适应的混合阈值进行二值化.这种方法虽然结合了2种方法的优点,但无法同时很好的处理伪影的产生,笔画边缘不够清晰.

Abdullah 等人^[13]介绍了有一种基于对比度增强的技术来处理退化文档,通过一种均值直方图拉伸方法来抑制背景噪声,同时增加边缘或近边缘的像素对比度,从而增强图像,最终在利用局部阈值法进行二值化. Lin 等人^[14]同样利用对比度增强来进行二值化.然而这种方法依赖于局部信息标准差的影响,在暗度低对比度区域效果不明显.

何皇兴等人^[15]提出了一种背景估计和二值化集成自适应的方法,用局部方法得到高召回率的图片,在进行背景估计得到高精确率的图片.但是这种基于背景估计^[15,16]的方法往往由于退化文档的多样性很难保证最终的二值化结果.

还有学者提出了基于聚类分割法,通过聚类算法将图像亮度值进行分类.李志杰等人^[17]提出一种改进樽海鞘群优化 K-means 算法的图像分割,但这种二值化方式存在无法处理光线突变明显的图片或者计算量较大等缺点.

一些最新的研究是基于偏微分方程进行文档的二值化,文献^[18]采用偏微分方程框架,将退化文档图像分解为背景和前景分量的乘积,提出一种弱耦合非线性扩散系统,用于同时恢复和二值化退化的文档图像.文献^[19]也是类似的方法,不过这种方法往往需要手动调整参数,结果不够稳定.

基于深度学习的方法是当下的一个热门研究点,

尤其是卷积神经网络, 一个重要的里程碑是 Long 等人^[20]在 2015 年提出的端到端全卷积神经网络 (fully convolutional network, FCN). 文献[21–23]分别采用 U-net 网络、GAN 网络等网络结构来分割前景背景, 但这种基于神经网络的方法难免会因为数据集的差异而取得不同的二值化结果图, 同时计算量大、样本集难以创建也是这种方法不可避免的缺陷.

上述的二值化方法虽然对于一般的退化类型效果较好, 但是都具有一定的局限性, 通常仅考虑了局部信息的均值、平方差、熵等因素, 在形态纹理方面考量较为欠缺, 从而导致针对低对比度的暗部区域很难进行处理. 为了能更好地解决在单字图片中的此类问题, 如图 1 所示, 本文提出一种基于多维侧窗聚类分块 CS-SWF (clustering segmentation based SWF) 的二值化方法. SWF (side window filtering) 是由 Yin 等人在 2019 年提出的思想^[24], 可以充分地保留图像纹理和边缘信息, 这种方法后来也被广泛地运用到其他研究方面. Ren 等人利用 SWF 算子提取边缘位置信息, 从而获取到更多的纹理细节进行二值化^[25]. Lu 等人还提出了一种改进的 SWF 算法对采集的图片进行去噪^[26]. 上述研究表明 SWF 算子具有良好的形态学聚类特性, 可以保证边缘及纹理的连续性, 故而本文提出 CS-SWF 算法可以完整的体现出书法字字体的边缘信息, 平滑笔画段, 在暗部低对比度区域表现出不错的特征提取性能. 实验结果表明, 在自制的 100 张手写退化文档图像数据集下, 本文方法对于低对比度暗部区域的图片二值化效果较为稳定.



图 1 非均匀光照退化文档单字图

1 算法流程

本文使用提出的 CS-SWF 方法对含有暗部低对比度区域的书法字图像进行处理, 如图 1 所示. CS-SWF 算法共分为预处理、基于 SWF 的特征聚类、相似块特征图修正、区域分类二值化以及图像去噪这 5 大步骤.

1.1 预处理

(1) 图像灰度化及去噪^[27]. 这里为了降低彩色辅助虚线部分的影响, 设 (i, j) 表示二维图片像素点的坐标, $g_1(i, j)$ 表示像素点坐标 (i, j) 经过灰度化后的像素值, 则灰度化公式为:

$$g_1(i, j) = \max(R(i, j), G(i, j), B(i, j)) \quad (1)$$

由于 SWF 特征对噪声可能较为敏感, 故采用高斯模糊进行去噪, 设 $s(i+u, j+v)$ 表示图片 g_1 在该坐标的像素窗口, $f(u, v)$ 表示高斯滤波函数. 则去噪后图片 g_2 公式如下:

$$g_2(i, j) = \sum_{u=-r}^r \sum_{v=-r}^r s(i+u, j+v) f(u, v) \quad (2)$$

(2) 基于细化算法的笔画宽度 (TH-SW) 计算方法. 在后续二值化的步骤中, 笔画的宽度会影响最终的结果, 因此本文提出一种基于汉字细化的笔画宽度计算方法 (TH-SW), 可以较为准确地确定笔画宽度, 算法如算法 1.

算法 1. TH-SW

- 1) 利用局部的二值化算法^[5]粗略的提取字体前景, 同时进行形态学闭操作平滑汉字图像.
- 2) 设置初始权重图 map_0 , 初始化时, 令前景部分权值 $map_0(f-g)=1$, 背景部分权值 $map_0(b-g)=1$.
- 3) 采用索引表细化算法进行细化减少运行时间, 在每次细化迭代过程中, 第 i 次迭代产生的权值图 map_i 由第 $i-1$ 次的权重值 map_{i-1} 得到. 设 (x, y) 为当前坐标, (u, v) 是相邻的距离 (x, y) 最近的且在第 $i-1$ 次迭代过程中消去的坐标点, 则 $map_i(x, y)$ 定义如下:

$$map_i(x, y) = map_{i-1}(x, y) + last(map_{i-1}(u, v)) \quad (3)$$

其中, $last(map_{i-1}(u, v))$ 表示离坐标 (x, y) 最近的非零权重坐标 (u, v) 在第 $i-1$ 次消除的像素点权重值. 在选取相邻坐标合并时, 共有 4 种情况, 图 2 中深灰色方框代表第 $i-1$ 次迭代后会消除的权重像素点, 本文列出了在这 4 种情况下, 每次消除该权重像素点时, 其权重的变化.

- 4) 重复步骤 3), 直到细化完成;
- 5) 统计最终的权重图 map_i , 同时剔除其中骨架长度较短的 map_i 权值以及骨架 2 端的相邻 2 个像素的权值, 防止造成干扰. 根据统计结果, 寻找一个值 l , 使得函数 f 值最小, 此时 $l = \min(f)$ 被定为平均笔画宽度.

$$f = \frac{\sum_{x=0}^n \sum_{y=0}^m |l - map_i(x, y)|}{count(map_i)} \quad (4)$$

其中, $count(map_i)$ 表示对 $map_i(x, y) \neq 0$ 的坐标点进行计数.

图 3 揭示了一个简单笔画段“点”和“横”在上述步骤中细化过程中的权重图值 map_i 的变化, 其中 map_0 是笔画段的原图, 此时所有前景权值被初始化为 1, map_1 则是经过第 1 次细化迭代后, 权值图的变化, 其他图类

似,不再阐述.最终经过4次细化迭代, map_4 展示了最终的权值图,统计该权重图可以获得平均笔画宽度 $l = 3$.

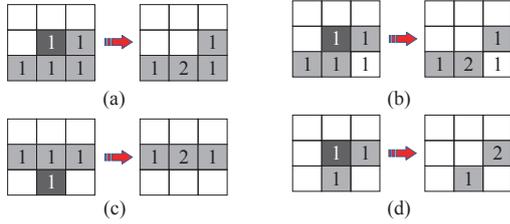


图2 权值合并模板

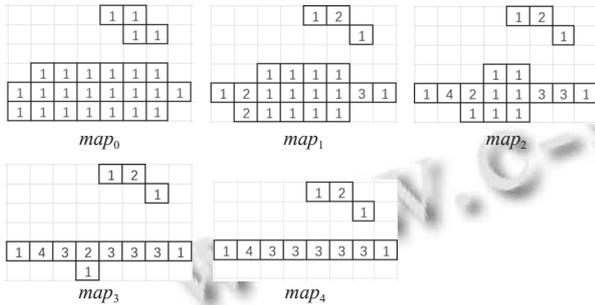


图3 笔画权重图变化过程

1.2 基于 SWF 的特征聚类

本节首先针对暗部低对比度区域进行分析,图4为图1所示图片中亮部高对比度局部区域的直方图和暗部低对比度局部区域的直方图,结果表明,暗部低对比度区域的前后景灰度值差异极小,为了有效地提取书法字笔画段的信息,本文提出的CS-SWF算法首先使用文献[25]所使用的SWF算子来获取特征信息,其核心思想是将当前像素位于局部窗口的边缘,其目的在于对同一类像素保持粘连性,对不同类像素呈现互斥性.这样就吧单像素二值化问题转化为聚类像素分类问题.

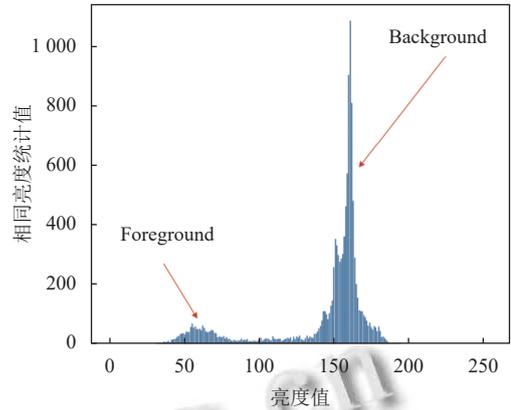
SWF算子模板如图5所示,设 $F1$ 代表 $\{U,R,D,L\}$ 4个模板内任意一个的像素点集合,则滤波器 U,R,D,L 的计算方式定义为:

$$v_{i,j} = \frac{\sum_{(i,j) \in F1} \|x(i,j) - x(r+1, r+1)\|}{(2r+1)(r+1)} \quad (5)$$

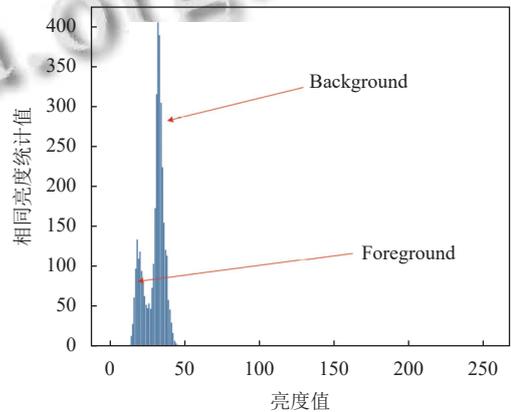
设 $F2$ 代表 $\{NW,NE,SE,SW\}$ 模板内任意一个像素点集合,则滤波器 NW,NE,SE,SW 的计算方式定义为:

$$v_{i,j} = \frac{\sum_{(i,j) \in F2} \|x(i,j) - x(r+1, r+1)\|}{(r+1)^2} \quad (6)$$

由此可得到当前像素点的8维SWF特征值 $V_{SWF} = \{U,R,D,L,NW,NE,SE,SW\}$.



(a) 高亮度对比度区域直方图



(b) 低亮度对比度区域直方图

图4 亮度高对比度区域的局部直方图和暗部低对比度局部区域的直方图

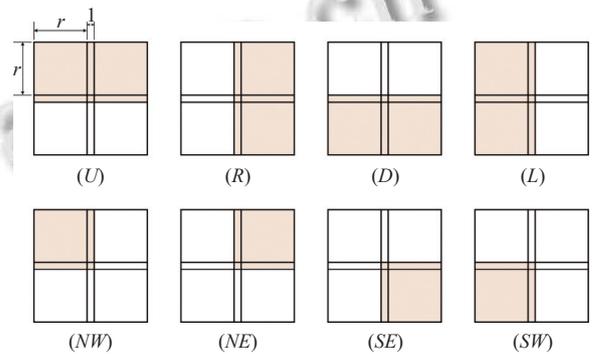


图5 SWF模板图

接着,根据当前像素的SWF特征值,进行局部的形态学聚类.设 i,j 表示当前像素的坐标位置, $C_{dir}(i,j)$ 表示在 i,j 位置的像素点特征值 V_{SWF} 最大的模板序号, $C_{sim}(i,j)$ 表示在 i,j 位置的像素点特征值 V_{SWF} 最小的模板序号.遍历每一个像素点的 V_{SWF} ,计算差异模板序号 $C_{dif}(i,j)$ 与 $C_{sim}(i,j)$,计算公式如下所示:

$$C_{dif}(i,j) = \operatorname{argmax}(V_{SWF}(i,j)) \quad (7)$$

$$C_{sim}(i, j) = \operatorname{argmin}(V_{SWF}(i, j)) \quad (8)$$

像素点的特征可以表示为 $F_{i,j} = \{C_{dif}, C_{sim}\}$, 共有 8×8 种组合. 将 $F_{i,j}$ 相同的像素点进行四邻域聚类, 完成相似像素块的划分, 相似块特征图 F 如图 6(a) 所示, 可以观察到汉字字体周围展现出连续、规则的聚类结果, 图 6 中同一类相似块共享同一种颜色, 相似块的轮廓可以突出字体形状的形态学信息.

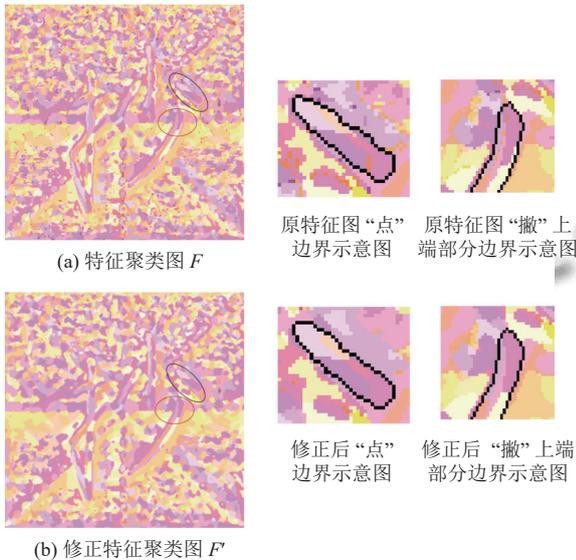


图 6 相似像素块特征图

1.3 相似块特征图修正

通过观察第 1.2 节生成的相似块特征图, 在一些光线较暗区域的笔画边缘会出现一定的像素抖动, 导致笔画边界不够清晰, 如图 6 绿色框位置所示. 为了修复汉字笔画边缘不清晰的情况, CS-SWF 算法对相似块特征图进行修正, 具体步骤如下.

(1) 对比特征图计算. 很容易假设图像相邻像素之间往往具有高度相关及相似的灰度值, 而噪声像素是非结构化且较为独立的^[28], 在此前提下, 引入 Zero-shot Noise2Noise 模块来获得下采样后的图片 q_1 和 q_2 得到对比特征图. 下采样的卷积核如图 7 所示.

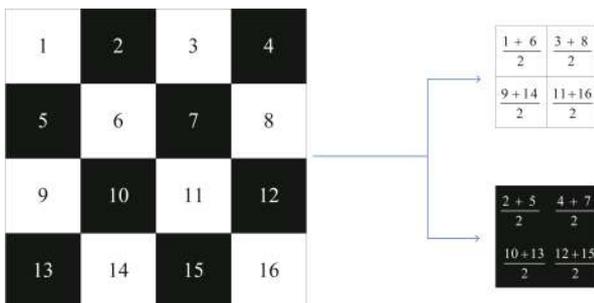


图 7 Zero-shot Noise2Noise 卷积核变化过程

接着分别对图片 q_1 和 q_2 进行 SWF 的形态聚类, 得到 2 幅对比特征图 F_1 和 F_2 .

(2) 特征图修正规则. 根据对比特征图 F_1 和 F_2 对原特征图 F 的缺陷进行修正, 同时, 定义原图的相似块特征图中每个相似块中像素个数小于 2 的相似块为空白块 **blank**, 需要重新分类, 遵循 CS-SWF 算法提出的规则.

在描述规则之前, 先定义如下符号.

设原特征图 F 中某 2×2 卷积块为:

$$T = \begin{bmatrix} t_1 & t_2 \\ t_4 & t_3 \end{bmatrix} \quad (9)$$

则特征图 F_1 和 F_2 会有唯一一个像素点 $u(u_1, u_2)$ 、 $v(v_1, v_2)$ 与该区域对应.

T_u, T_d 分别表示矩阵 T 的左上三角矩阵和右下三角矩阵, 定义如下:

$$T_u = \begin{bmatrix} t_1 & t_2 \\ t_4 & 0 \end{bmatrix} \quad (10)$$

$$T_d = \begin{bmatrix} 0 & t_2 \\ t_4 & t_3 \end{bmatrix} \quad (11)$$

$\operatorname{sum}(T == u)$ 表示矩阵 T 中元素与 u 相等的个数, 定义如下:

$$\operatorname{sum}(T == u) = \sum_{i=1}^4 (t_i == u) \quad (12)$$

T_{blank} 表示矩阵 T 中标记为空白块的元素.

CS-SWF 算法提出 3 条规则修正特征图, 规则如修正规则 1 所述.

修正规则 1. 特征修复规则

1) 若卷积块 T 中包含空白块 **blank**, 则更新空白块的值, 更新规则如下:

$$\begin{cases} \text{if } \operatorname{sum}(T == u) > \operatorname{sum}(T == v): \\ \quad \text{then } T_{\text{blank}} = u \\ \text{else if } \operatorname{sum}(T == u) \leq \operatorname{sum}(T == v): \\ \quad \text{then } T_{\text{blank}} = v \end{cases} \quad (13)$$

这一规则可以修正随机噪点特征;

2) 下采样特征图中对应像素特征值相同, 表明一定程度上, 原图对应区域具有同一特征, 此时修正原特征图中对应区域的异常噪声, 具体更新规则如下:

$$\begin{cases} \text{if } u == v: \\ \quad \text{then } T = u = v \end{cases} \quad (14)$$

3) 下采样特征图对应像素特征值不相同, 表明在某个区域内, 特征值出现波动需要修正. 具体更新规则如下:

$$\begin{cases} \text{if } u \neq v \&\& \operatorname{sum}(T_u == u) > \operatorname{sum}(T_d == v): \\ \quad \text{then } T_u = u \\ \text{else if } u \neq v \\ \quad \&\& \operatorname{sum}(T_u == u) \leq \operatorname{sum}(T_d == v): \\ \quad \text{then } T_d = v \end{cases} \quad (15)$$

经过修正后,对比于原特征图,部分笔画边缘更加清晰平滑,异常点减少,如图6(b)特征图 F' 所示

1.4 区域分类二值化

CS-SWF 算法同时使用一种相似块前后景分离方法对前文中聚类形成的相似块进行分类.修正后的特征图 F' 中的相似块被称为一个平滑区域,依据局部的原图灰度值信息对平滑区域进行分类,确定前景还是背景,以此来得到一个边缘信息连续清晰的二值化图,算法如算法2.

算法2. 区域分类二值化算法

- 1) 设图像中的平滑区域为 $CS=\{cs_1, cs_2, cs_3, \dots, cs_n\}$, n 为相似块个数.
- 2) 依次遍历 CS 中的区域,以其平滑区域的质心为中心点,建立窗口大小为 $2r+1$ 的滑动窗口,计算窗口内连通域的均值 Mrc ,以及窗口均值 m ,标准差 v .设 $len(cs_i)$ 为平滑区域 cs_i 的长度, $wid(cs_i)$ 平滑区域 cs_i 的宽度,窗口半径 r 取值如下:

$$r = \begin{cases} l, & \max(len(cs_i), wid(cs_i)) \leq l \\ \max(len(cs_i), wid(cs_i)), & \max(len(cs_i), wid(cs_i)) > l \end{cases} \quad (16)$$

- 3) 设 $T(i, j)$ 为分类阈值,利用式(17)计算当前平滑区域的阈值.

$$T(i, j) = m - k \times m \times \frac{v}{128} \quad (17)$$

- 4) 若阈值小于连通域的均值 Mrc ,该平滑区域作为背景,反之,该平滑区域作为前景, Bin 分类公式如下:

$$Bin = \begin{cases} 1, & Mrc \leq T \\ 0, & Mrc > T \end{cases} \quad (18)$$

经过本节的分类算法,可以粗略提取出汉字“抄”的前景,如图8(a)所示,可以看到字体边缘较为圆滑,结构完整.

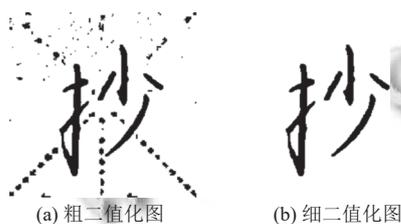


图8 基于SWF特征的二值化图

1.5 二值化图像去噪

针对第1.4节中得到的粗二值化图像,当窗口内不包含前景时,窗口内部会不可避免地引入一些不规则的噪声,为了移除噪点.本文利用在第1.1节中求出的笔画宽度 l ,定义 $Area$ 为噪点面积大小,将面积小于 $Area$ 的连通域作为噪点去除, $Area$ 的面积定义为:

$$Area = l^2 \quad (19)$$

$$Bin = \begin{cases} a \leq Area, & \text{remove noise} \\ a > Area, & \text{reserve} \end{cases} \quad (20)$$

通过上述方法去除噪点后,可以看到汉字周围的噪声都被移除,显示出字体的前景,最终的二值化结果图如图8(b)所示.

2 实验及结果分析

本节进行了大量的实验来评估所提出的方法性能.首先介绍了用于测试的自建数据集,用于智能书法评价系统的不良光照样例图片,接着将本文的二值化方法与其他经典算法进行了定量比较.本节的所有工作均在PC(2.30 GHz, 16 GB RAM的i7-12700H)上完成,算法实现环境为Python 3.6, OpenCV 3.3.0.11

2.1 数据集

在课题的前期准备工作中,为提供智能书法评价辅助系统应对不良光照条件下对书法字文档的处理条件,本文联合了书法教学公司建立了自己的数据集,均写在田字格字帖上,通过扫描、拍摄等方式存储到电子设备,最后利用分割方法^[29,30]切分出单字图片.为了模拟用户在不同光照环境下的拍摄条件,随机抽取出20张单字图片,通过固定摄像位置并对同一个书法字施加不同的光源拍摄多组照片作为测试集,由于拍摄的照片仅仅是光源位置发生改变,因此认为同一个书法字的不同光源位置的照片共享同一个GT图,而标准的GT图可由均匀光照下的该书法字图片经过二值化算法并人工微调得到.最终创建出包含100张图片的测试集,其中共含有20个均匀光照的不同书法字图像以及80张随机光源的书法字图像.部分数据集图片如图9所示.

2.2 定量评价

选取自制测试集进行测试,并与Otsu、Niblack、Sauvola、Bernsen及文献[25]中的方法进行比较.其中Otsu是全局二值化方法,剩余的3种则是局部二值化方法.Niblack、Sauvola算法的参数的取值依照文献中的建议, R 取值一般为128,超参数 k 的取值,经过大量实验,取 $k=0.5$ 和 $k=0.2$ 效果作为对比图较好.Bernsen方法中超参数 c 按照文献中的建议,取 c 为15.文献[25]同样采用了SWF算子,与本文方法相似.

图10展现了部分测试图片在这4种算法以及本文提出的算法二值化结果对比图.对上述的测试结果

进行比较,可以明显地发现,基于全局的二值化方法在处理光线分布不均匀的图片时会完全丧失暗部图像的信息,Niblack 在窗口内部没有前景像素时表现出大量噪声, Sauvola 算法受到参数以及局部对比度的影响,图片之间不同光照条件的差异会带来不同的效果,但是可以从汉字某些笔画的边缘看出,在一些暗部低对比度区域,汉字边缘界限模糊,像素边缘毛刺严重,笔

画也出现不同程度的断裂,而 Bernsen 算法受到局部窗口的最大最小值影响,在窗口无前景区域以及低对比度区域表现出大量的噪声.文献[25]由于自身算法中选取了阈值,导致 SWF 算子在低对比度区域表现不佳,笔画出现空洞.本文提出的方法对不同光照情况,尤其是低对比度区域表现出较好的适应性,在均匀光照条件也能完成前景分割任务.

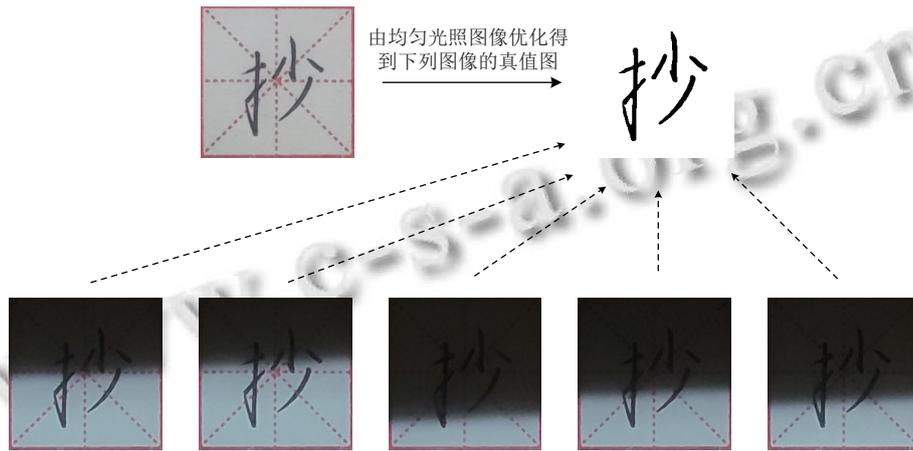


图9 部分数据集示例图

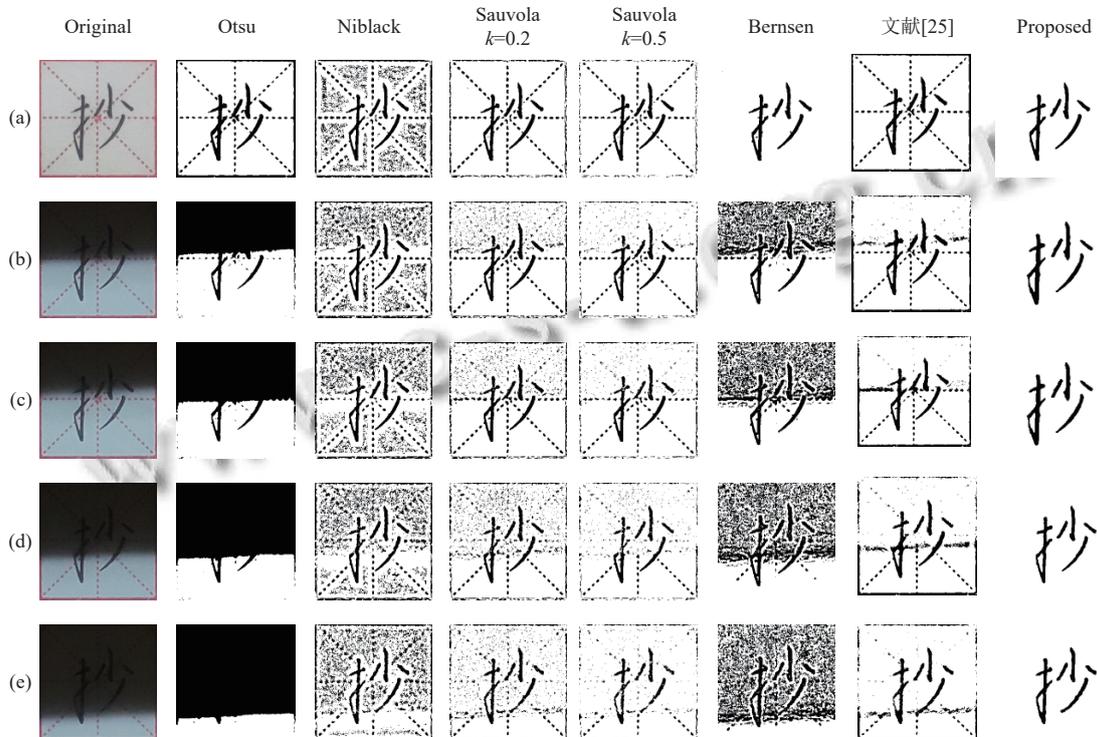


图10 算法效果对比图

为了客观评价算法的优劣,采用一系列在国际二值化竞赛中使用的评估指标,包括 FM (F -measure)、

$PSNR$ (峰值信噪比) 和 DRD (距离倒数失真)^[31].

FM 是准确率 P 、召回率 R 的综合评价指标,这

3个指标的值越高,表明二值化图与真值图GT之间越相近,当 $FM = 1$ 时,二值化图与真值完全相同,效果最好。

$$FM = \frac{2 \times R \times P}{R + P} \quad (21)$$

PSNR是测量二值化图像与GT图像的接近程度,指标的值越高,表明图像质量越好。

$$PSNR = 10 \times \log \left(\frac{255^2}{MSE} \right) \quad (22)$$

$$MSE = \frac{\sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (f(i, j) - f^0(i, j))^2}{N_i N_j} \quad (23)$$

其中, N_i 、 N_j 分别表示图像的长度和宽度, $f(i, j)$ 表示GT图,表示处理后的二值化图在 (i, j) 中的值。

DRD用于测量二值图像的视觉失真,这种方法侧重于人类感知图像的表现,当二值化效果较好时,DRD的值较低。

$$DRD = \frac{\sum_k DRD_k}{NUBN} \quad (24)$$

其中, DRD_k 是第 k 个翻转像素的失真,NUBN是GT图像非均匀(不是全部黑色或白色像素) 8×8 窗口的数量。

与其他5种二值化算法相比,表1统计了样例图9中第2幅测试图片在不同二值化算法情况下的各项数据,表2统计了包含100张图片的自制数据集的各项平均指标,本文是算法在各个评价指标中表现较好,证明了本文算法对退化文档图像,尤其是暗部低对比度区域的二值化效果更好、适应性强,二值化分割精度高。

表1 图10(b)的定量评价指数对照的定量评价指数对照

Algorithm	FM (%)	PSNR (dB)	DRD ($\times 10^{-2}$)
Otsu	0.332	52.102	97.081
Niblack	0.687	56.341	59.323
Sauvola, $k=0.5$	0.899	59.237	21.215
Sauvola, $k=0.2$	0.847	58.204	25.190
Bernsen	0.782	56.433	8.790
文献[25]	0.910	54.289	31.813
Proposed	0.932	64.783	9.310

3 结语

针对不良光照的退化文档图片,尤其是光线昏暗低对比度区域的汉字,提出了基于多维侧窗聚类分块

(CS-SWF)的二值化方法。该方法提升了传统方法在低对比度区域对信息的提取效果,克服了传统方法缺少汉字形态学的缺点。通过利用8维SWF卷积核描述图片像素的形态学特征进行相似块聚类,并提出修正规则,利用低维特征图的特点来修正相似块特征图,使得其符合汉字的形态学特征,最后对相似块进行分类,得到最终的二值化图。通过模拟用户的拍照环境,拍摄出不同光照条件的退化文档图片,构建出适用于退化文档二值化图像的测试数据集,通过定量评价来验证算法的有效性,实验结果表明,本文算法对光线昏暗低对比度区域的笔画细节有更好的提取效果,同时对于光线分布不均的退化文档二值化具有更好的精度和鲁棒性。

表2 数据集剩余图片的平均定量评价指数

Algorithm	FM (%)	PSNR (dB)	DRD ($\times 10^{-2}$)
Otsu	0.401	51.781	89.403
Niblack	0.718	57.313	60.020
Sauvola, $k=0.5$	0.913	60.829	26.243
Sauvola, $k=0.2$	0.861	57.901	21.270
Bernsen	0.802	54.212	13.377
文献[25]	0.882	56.381	27.829
Proposed	0.928	62.087	11.271

参考文献

- 雷实. 书法教育主要是写好汉字. 人民教育, 2012(10): 35-37.
- 冀晓萍. 中办国办印发《关于实施中华优秀传统文化传承发展工程的意见》要求 将中华优秀传统文化传承贯穿国民教育始终. 人民教育, 2017(3-4): 27.
- Otsu N. A threshold selection method from gray-level histogram. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66. [doi: 10.1109/TSMC.1979.4310076]
- 魏兴凯, 蒋峥, 傅呈勋, 等. 基于光照影响因子的动态Niblack算法研究及应用. 计算机工程与设计, 2022, 43(4): 1066-1073.
- Kaur A, Rani U, Josan GS. Modified Sauvola binarization for degraded document images. Engineering Applications of Artificial Intelligence, 2020, 92: 103672. [doi: 10.1016/j.engappai.2020.103672]
- 李艺杰, 邹坤霖, 孙炜, 等. 基于Sauvola算法和神经网络的图像自适应二值化方法. 测控技术, 2020, 39(08): 62-69, 75.
- 乔志凯, 陈世才, 蒙子昕, 等. 复杂光照下的船舶水尺图像二值化新方法. 中国航海, 2021, 44(4): 80-85, 93. [doi: 10.3969/j.issn.1000-4653.2021.04.013]
- Pai YT, Chang YF, Ruan SJ. Adaptive thresholding

- algorithm: Efficient computation technique based on intelligent block detection for degraded document images. *Pattern Recognition*, 2010, 43(9): 3177–3187. [doi: [10.1016/j.patcog.2010.03.014](https://doi.org/10.1016/j.patcog.2010.03.014)]
- 9 Singh BM, Sharma R, Ghosh D, *et al.* Adaptive binarization of severely degraded and non-uniformly illuminated documents. *International Journal on Document Analysis and Recognition (IJ DAR)*, 2014, 17(4): 393–412. [doi: [10.1007/s10032-014-0219-6](https://doi.org/10.1007/s10032-014-0219-6)]
- 10 Bataineh B, Abdullah SNHS, Omar K. Adaptive binarization method for degraded document images based on surface contrast variation. *Pattern Analysis and Applications*, 2017, 20(3): 639–652. [doi: [10.1007/s10044-015-0520-0](https://doi.org/10.1007/s10044-015-0520-0)]
- 11 Calderon F, Garnica-Carrillo A, Reyes-Zuñiga C. Binarization of images with variable lighting using adaptive windows. *Signal, Image and Video Processing*, 2022, 16(7): 1905–1912. [doi: [10.1007/s11760-022-02150-1](https://doi.org/10.1007/s11760-022-02150-1)]
- 12 刘朋远, 田瑞, 周媛奉, 等. 基于自适应混合阈值的智能电表图像二值化. *计算机应用与软件*, 2023, 40(1): 210–215. [doi: [10.3969/j.issn.1000-386x.2023.01.033](https://doi.org/10.3969/j.issn.1000-386x.2023.01.033)]
- 13 Abdullah SNHS, Ismail SM, Hasan MK, *et al.* Novel adaptive binarization method for degraded document images. *Computers, Materials & Continua*, 2021, 67(3): 3815–3832.
- 14 Lin SCF, Wong CY, Rahman MA, *et al.* Image enhancement using the averaging histogram equalization (AVHEQ) approach for contrast improvement and brightness preservation. *Computers & Electrical Engineering*, 2015, 46: 356–370.
- 15 何皇兴, 陈爱国, 王蛟龙. 背景估计和局部自适应集成的手写图像二值化. *计算机科学*, 2022, 49(11): 163–169. [doi: [10.11896/jsjcx.210900225](https://doi.org/10.11896/jsjcx.210900225)]
- 16 冯炎. 基于背景估计和对比度补偿的退化古籍图像二值化算法. *科学技术与工程*, 2015, 15(34): 105–109.
- 17 李志杰, 王力, 张习恒. 改进樽海鞘群优化 K-means 算法的图像分割. *包装工程*, 2022, 43(9): 207–216.
- 18 Jacobs BA, Celik T. Unsupervised document image binarization using a system of nonlinear partial differential equations. *Applied Mathematics and Computation*, 2022, 418: 126806. [doi: [10.1016/j.amc.2021.126806](https://doi.org/10.1016/j.amc.2021.126806)]
- 19 Du ZJ, He CJ. Nonlinear diffusion system for simultaneous restoration and binarization of degraded document images. *Computers & Mathematics with Applications*, 2024, 153: 237–248.
- 20 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015. 3431–3440.
- 21 Zhang ZA, Wu CD, Coleman S, *et al.* DENSE-INception U-net for medical image segmentation. *Computer Methods and Programs in Biomedicine*, 2020, 192: 105395. [doi: [10.1016/j.cmpb.2020.105395](https://doi.org/10.1016/j.cmpb.2020.105395)]
- 22 王红霞, 武甲礼, 陈德山. 分离复杂背景下的文档图像二值化方法. *中国图象图形学报*, 2023, 28(7): 2011–2025.
- 23 Ramya BN, Singh S. A novel approach to implement binarized neural network to enhance accuracy using machine learning techniques. *SN Computer Science*, 2023, 4(2): 174. [doi: [10.1007/s42979-022-01640-w](https://doi.org/10.1007/s42979-022-01640-w)]
- 24 Yin H, Gong YH, Qiu GP. Side window filtering. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 8750–8758.
- 25 Ren H, Wang YJ, Dong X. Binarization algorithm based on side window multidimensional convolution classification. *Sensors*, 2022, 22(15): 5640. [doi: [10.3390/s22155640](https://doi.org/10.3390/s22155640)]
- 26 Lu KL, Liu EH, Zhao RJ, *et al.* Star sensor denoising algorithm based on edge protection. *Sensors*, 2021, 21(16): 5255. [doi: [10.3390/s21165255](https://doi.org/10.3390/s21165255)]
- 27 Castleman KR. *Digital Image Processing*. Englewood: Prentice Hall, 2011. 74–120.
- 28 Mansour Y, Heckel R. Zero-shot Noise2Noise: Efficient image denoising without any data. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 14018–14027.
- 29 陈艳红, 崔晓光, 张吉祥. 硬笔书法田字格切片图像提取方法、系统、设备: 中国, 113901936A. 2022-01-07.
- 30 徐占洋, 张家瑞, 秦飞扬, 等. 一种汉字字体二值化切片图像的提取方法及系统: 中国, 115909369A. 2023-04-04.
- 31 Pratikakis I, Gatos B, Ntirogiannis K. ICDAR 2013 document image binarization contest (DIBCO 2013). *Proceedings of the 12th International Conference on Document Analysis and Recognition*. Washington: IEEE, 2013. 1471–1476.

(校对责编: 孙君艳)