

背景估计和局部自适应集成的手写图像二值化

何皇兴, 陈爱国, 王蛟龙

引用本文

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](#)[J]. 计算机科学, 2022, 49(11): 163-169.

HE Huang-xing, CHEN Ai-guo, WANG Jiao-long. [Handwritten Image Binarization Based on Background Estimation and Local Adaptive Integration](#)[J]. Computer Science, 2022, 49(11): 163-169.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[高分辨率斜视聚束 SAR 回波仿真加速算法研究](#)

Study on Acceleration Algorithm for Raw Data Simulation of High Resolution Squint Spotlight SAR

计算机科学, 2022, 49(8): 178-183. <https://doi.org/10.11896/jsjcx.210600066>

[指静脉识别技术研究综述](#)

Survey on Finger Vein Recognition Research

计算机科学, 2022, 49(6A): 1-11. <https://doi.org/10.11896/jsjcx.210400056>

[视频流的实时景深延拓算法](#)

Real-time Extend Depth of Field Algorithm for Video Processing

计算机科学, 2022, 49(6A): 314-318. <https://doi.org/10.11896/jsjcx.201100187>

[基于多重差异特征网络的街景变化检测](#)

Street Scene Change Detection Based on Multiple Difference Features Network

计算机科学, 2021, 48(2): 142-147. <https://doi.org/10.11896/jsjcx.200500158>

[图像去雨算法在云物联网应用中的研究综述](#)

Survey of Intelligent Rain Removal Algorithms for Cloud-IoT Systems

计算机科学, 2021, 48(12): 231-242. <https://doi.org/10.11896/jsjcx.201000055>

背景估计和局部自适应集成的手写图像二值化

何皇兴¹ 陈爱国¹ 王蛟龙²

1 江南大学人工智能与计算机学院 江苏 无锡 214122

2 江南大学物联网工程学院 江苏 无锡 214122

(hxx97@qq.com)

摘要 手写文档图像中存在光照不均、笔墨浸染、纸张退化、阴影等复杂情况,针对文档图像在复杂背景下二值化后 OCR 效果不理想的问题,提出了一种对改进的背景估计和局部自适应集成的二值化方法。首先利用局部自适应方法得到具有高召回率的二值化图像,然后对背景估计的方法进行改进得到具有高精确率的二值化图像,最后基于连通域的方法将两种类型的图像集成得到结果。使用 4 种评价指标在 DIBCO2013 和 DIBCO2016 手写数据集上进行了对比实验,结果表明该方法整体性能优于 Otsu, Wolf, Niblack, Sauvola, Singh 和 Howe 等经典算法。

关键词 图像处理; 手写文档图像; 二值化; 背景估计; 二值化集成

中图法分类号 TP391

Handwritten Image Binarization Based on Background Estimation and Local Adaptive Integration

HE Huang-xing¹, CHEN Ai-guo¹ and WANG Jiao-long²

1 School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, Jiangsu 214122, China

2 School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract In handwritten document images, there are often uneven lighting, ink stains, paper degradation, shadows and other complex conditions. In order to solve the problem that OCR effect of document image is not ideal after binarization in complex background, a binarization method of improved background estimation and local adaptive integration is proposed. In this method, the local adaptive binarization method is first used to obtain the binarization image with high recall rate, then the improved background estimation method is used to obtain the binarization image with high accuracy rate, and finally the two types of binarization images are integrated based on the connected domain method to obtain the final binarization image. Experimental results on DIBCO2013 and DIBCO2016 handwritten data sets show that the proposed method has better overall performance than Otsu, Wolf, Niblack, Sauvola, Singh and Howe.

Keywords Image processing, Handwritten document image, Binarization, Background estimate, Binary integration

1 引言

文档图像的处理技术被广泛应用于图像处理、模式识别等领域,如文档图像的数字化、快递单识别、车牌识别等。文档图像的处理一般包括图像采集、二值化、光学字符识别(Optical Character Recognition, OCR)这 3 个步骤。其中,图像的二值化是将采集到的原始图像转换为二进制的图像。图像二值化的结果将直接影响图像 OCR 识别的性能。

文档图像包括印刷文档图像和手写文档图像。印刷文档图像的笔画、格式较为统一,而手写文档图像字体、笔画都不固定,因此手写图像较印刷图像往往更难处理。手写

文档图像还存在更多的复杂情况,比如光照不均匀、图像低质量、对比度低、污点、重影等非文本噪声^[1]。因此,针对手写文档图像的二值化依然是一个具有挑战性的研究课题。

针对图像的二值化算法,国内外学者已经进行了大量的研究。目前对二值化算法的分类大致如下:

(1)基于阈值的二值化算法,即为图像的每个像素点设置一个阈值以区分前景和背景,如基于全局阈值的 Otsu 算法^[2]。Otsu 算法通过找到一个阈值使类内方差最小,类间方差最大,其算法具有速度快的优点,但对光照不均匀、背景复杂的图像二值化效果不理想。基于局部阈值的算法能较好地处理光照不均匀的问题,如 Niblack^[3]用每个像素点邻域的

到稿日期:2021-09-27 返修日期:2022-03-15

基金项目:国家自然科学基金(61702225)

This work was supported by the National Natural Science Foundation of China(61702225).

通信作者:陈爱国(agchen@jiangnan.edu.cn)

平均灰度值和标准差确定局部的阈值,其算法能将对比度较低的前景分离出来,提高二值化图像的召回率,但会因此产生较多噪声。Sauvola 等^[4]在 Niblack 的基础上采用了不同的阈值选取方法,在相同标准差的情况下,对于灰度值较高的背景点,修改阈值使其更小,能过滤更多噪声,但对低对比度的图像效果欠佳^[5]。文献^[6]提出了一种基于积分图的自适应多阈值二值化方法,该方法根据均值松弛因子判断光照情况,并修改阈值以适应不同光照情况。

(2)基于数学理论的二值化方法,如基于条件随机场的 Howe^[7]算法,其使用拉普拉斯算子来估计前景和背景的局部似然性,并找到目标函数的最小能量,该算法在大多数情况下有比较稳定的效果。基于博弈理论的 Gib^[8]算法利用博弈论的思想,从预处理图像中提取像素特征后使用 K -means 进行聚类。文献^[9]介绍了 3 种新的基于广义模糊积分图像的自适应二值化模型,并通过改进积分图来优化算法,该方法适用于模糊积分。文献^[10]通过估计检测高对比度像素的平均值和标准偏差来估计局部阈值,但无法处理图像对比度较低的文档图像。

(3)基于像素分类的二值化算法,即通过机器学习的方法对图像中的每个像素点进行分类,以确定该点是前景还是背景。文献^[11]针对仪表图像提出了一种基于卷积神经网络(Convolutional Neural Network,CNN)的二值化方法,适用于光照不同的复杂环境。Westphal 等^[12]针对历史文档图像提出了基于循环神经网络(Recurrent Neural Network,RNN)的二值化方法。基于机器学习的方法需要大量已标记的文档图像来训练,文献^[13]提出了一种集成多个预训练的 U-Net 模块来解决训练图像较少的问题。文献^[14]提出了一种结合神经网络和域适应(Domain Adaptation,DA)的方法来无监督地进行文档图像的二值化。虽然基于机器学习的二值化方法与传统方法相比准确率更高,但是其训练速度与计算速度都较慢,并且需要大量的训练数据,仅适用于部分场景^[15]。

上述二值化方法虽然在某些场景下具有较好的效果,但都有一定的局限性。对于存在噪声和对比度低^[16]的图像,一般都较难处理。为了能够更好地解决复杂背景下的二值化问题,本文提出了一种二值化集成的方法,通过集成具有高召回率和高精确度的二值化图像,来优化最终的二值化结果。实验结果表明,本文方法在背景复杂的手写文档图像数据集下效果较稳定。

2 二值化集成方法

2.1 基本框架

如图 1 所示,本文提出的二值化方法包含以下 4 个步骤:

(1)预处理。首先对原始图像进行灰度化,再对灰度图像进行一次粗二值化,利用此二值化图像计算笔画宽度。

(2)局部自适应二值化。该算法的参数依据笔画宽度自适应设置,将参数设置得比正常二值化稍大,可得到具有高召回率的二值化图像。

(3)改进的背景估计。采用背景估计^[17]的方法并进行

适当的改进可以得到具有高精确率的二值化图像。

(4)二值化的集成。将高召回率二值化图像与高精度二值化图像进行基于连通域的二值化集成,得到最终的结果。

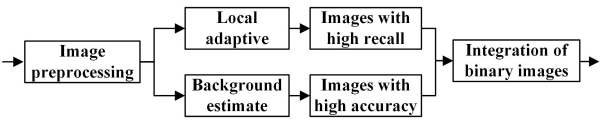


图 1 算法处理流程
Fig. 1 Processing flow of algorithm

2.2 预处理

(1)灰度化

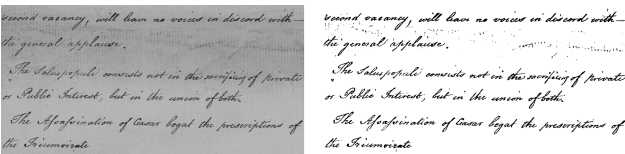
尽管有些学者已经研究了将彩色图像转为灰度图像对二值化性能的影响,但他们提出的方法需要对特定的数据集进行训练,耗时也更长^[18]。本文也尝试使用主成分分析(Principal Component Analysis,PCA)的方法生成灰度图像^[19],经实验,其对二值化图像质量并没有太大提升。综合考虑后,我们使用经典的光学公式^[20]来进行灰度化,其公式如下:

$$Y=0.299\times R+0.587\times G+0.114\times B$$
 (1)

其中, Y 表示该点灰度图像最终的亮度值, R,G,B 分别表示彩色图像对应的红、绿、蓝 3 通道的数值, Y,R,G,B 的范围为 0255。

(2)粗二值化

粗二值化是为了计算笔画宽度,对二值化的效果要求不高,因此本文选用了计算速度较快的 Otsu 算法,最终得到的二值化图像前景值为 1,背景值为 0,如图 2 所示。



(a) Grayscale image (b) Otsu binarization

图 2 粗二值化

Fig. 2 Rough binarization

(3)计算笔画宽度

本文在粗二值化后的图像上计算文本的笔画宽度。如果对于不同的文本图像使用固定的参数,得到的结果将不够理想,因此,依据笔画宽度来自适应设置参数,能够适应更多的图像,鲁棒性更好。确定文本笔画宽度的方法有许多,本文不需要特别精确的笔画宽度,因此采用了较为简单的横向扫描的方式来计算。本文的笔画宽度计算方法流程如下:

1)横向扫描。对二值化图像进行从左到右的横向扫描,统计图像中值连续为 1 的数量,作为笔画宽度并将其记录下来。从上到下对每一行都做同样的扫描并记录下所有笔画宽度的值。

2)估计笔画宽度。通过第 1)步统计能够得到每一个笔画宽度的对应数量,数量最多的即视为当前图像的笔画宽度。如图 3 所示,右边的数字表示每行的笔画宽度,对整个图片统计完成后,笔画宽度为 3 的共有 2 个,笔画宽度为 2 的共有 4 个,因此该示例图像的估计笔画宽度为 2。

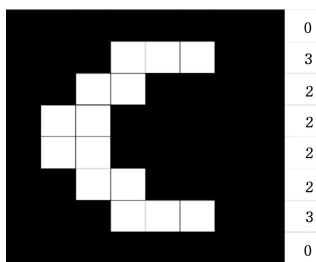


图3 笔画宽度计算示例

Fig. 3 Example of stroke width calculation

2.3 局部自适应二值化

这一步需要得到召回率较高的二值化图像,本文使用的是局部自适应均值二值化,该方法与其他具有高召回率的方法相比,运算速度快且引入的噪声更少,能够降低后期处理难度,也能提升最终二值化结果的准确率。

局部自适应二值化需要设置两个参数,一个是局部窗口的大小 bs ,根据局部自适应方法的特性和实验验证,窗口设置过小,会导致召回率较低,但窗口设置过大,提高召回率的同时也会导致精度下降。经实验,将窗口设置为笔画宽度的25倍较为合适,本文实验时将 bs 设置为笔画宽度的2倍;另一个是常数 C ,表示局部平均值减去的常数,使用均值通常不能去除噪声,设置常数 C 可以调整阈值以覆盖噪声范围,经实验,针对文档图像, C 设置为810较为合适,本文实验时设置为8。局部自适应均值二值化方法如下:

(1) 计算阈值

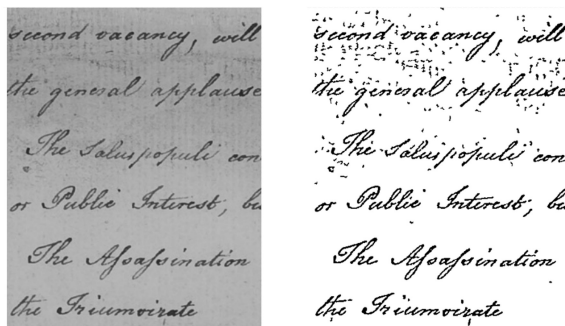
对图像中的一个点 (x, y) 的阈值计算如下:以点 (x, y) 为中心,计算其周围 $bs \times bs$ 大小的区域均值,将均值减去常数 C 得到点 (x, y) 的阈值 $T(x, y)$ 。

(2) 二值化

根据当前点 (x, y) 的像素值与第(1)步计算的阈值 $T(x, y)$ 进行比较得到二值化的结果,其公式如下:

$$dst(x, y) = \begin{cases} 255, & \text{if } src(x, y) > T(x, y) \\ 0, & \text{other} \end{cases} \quad (2)$$

其中, dst 表示最终得到的二值化图像, $src(x, y)$ 表示灰度图像上位于点 (x, y) 的亮度值, T 为第(1)步计算得到的阈值。最终得到的二值化图像如图4所示,可以看出,图像具有较高的召回率,其噪声也相对较多。



(a) Gray scale image

(b) High recall image

图4 高召回率二值化图像示例

Fig. 4 Example of binarized image with high recall

2.4 改进的背景估计

这一步需要得到精度较高的二值化图像,在保证高精度

的前提下,尽可能地降低噪声。因为基于背景估计的方法能够去除大量的非文本噪声,所以本文采用了基于背景估计的二值化来获取高精度的二值化图像。

本文使用的基于背景估计的二值化方法如下:首先使用局部自适应均值得到一张粗略的二值化图像;然后在其基础上进行背景估计,将原始灰度图减去估计的背景图,得到前景图像;最后对前景图像进行二值化即可得到具有较高精度的二值化图像。其具体处理流程如下:

(1) 局部自适应二值化与后处理

自适应均值二值化的具体实现方法在本文的2.3节有介绍,这里使用的方法相同。相较于得到高召回率图像,此处还加了一个后处理步骤:删除连通域面积较小的前景。图像中文本部分的面积远大于笔画宽度的方块,比笔画宽度方块更小的连通域通常是噪声,因此本文实验将删除最小面积参数设置为笔画宽度的平方。后处理可以在高召回率图像的基础上,去除部分较小的噪声,使下一步的背景估计更加准确。这一步最终得到的二值化图像为黑底白字,即前景像素值为1,背景像素值为0。

(2) 估计背景图像

估计图像的背景需要用到原始灰度图像和第(1)步得到的二值化图像,通过二值化图像的值来估计图像的背景。在点 (x, y) 处,如果该点的二值化图像值为0,表示该点为背景,估计的背景值为原始灰度图像的值。如果该点的二值化图像值为1,则表示当前点为前景,估计的背景为灰度图中该点周围邻域所有背景的中值。

估计背景图像的公式如下:

$$bg(x, y) = \begin{cases} src(x, y), & b(x, y) = 0 \\ median(t_bg), & b(x, y) = 1 \end{cases} \quad (3)$$

其中, bg 表示最终得到的背景估计图像, src 表示原始灰度图像, b 表示第(1)步得到的二值化图像, $median()$ 表示对矩阵中的非0值取中值的函数。 t_bg 表示一个矩阵,为灰度图与背景图矩阵乘积后的一块区域,该区域为点 (x, y) 周围固定大小的一个邻域,计算该矩阵的公式如下:

$$t_bg = (src * b)[x-d:x+d, y-d:y+d] \quad (4)$$

其中, $src * b$ 表示灰度图与背景图相乘后的图像, $[]$ 符号表示对矩阵切片,即提取矩阵的一部分数值,其中 d 表示窗口大小,本文参考了文献[17]中的参数取值,实验中的 d 取值为笔画宽度的5倍。最终得到的背景图效果如图5(b)所示。

(3) 估计前景图像

通过第(2)步可以得到估计的背景图,然后将原始灰度图减去背景图即可得到粗略估计的前景图像,但是此时的粗前景图一般不够准确,如果估计的背景和真实的背景是相同的,那么相减后得到的前景图像该点的像素值为0。但往往估计的背景值与真实的背景值在数值上有一定差异,如一个背景点的真实像素值为200,估计的背景像素值为198,两者相减后不是0。因此,为了得到更精确的前景图,本文还设置了一个阈值对估计的前景图像进行过滤,以便正确分离出背景和前景。

为了能够自适应不同的文本图像,本文通过原始灰度

图像的平均灰度值来设置阈值,即将小于平均灰度值的像素点视为背景,对大于平均灰度值的像素点不做处理。计算图像的平均灰度值的公式如下:

$$avg_gray = \frac{\sum_{ix=0}^{height} \sum_{iy=0}^{width} src(ix, iy)}{height \times width}$$

(5)

其中, avg_gray 表示计算得到的平均灰度值, src 表示原始灰度图像, $width$ 和 $height$ 分别表示原始图像的宽和高。最终得到的前景图效果如图 5(c) 所示。

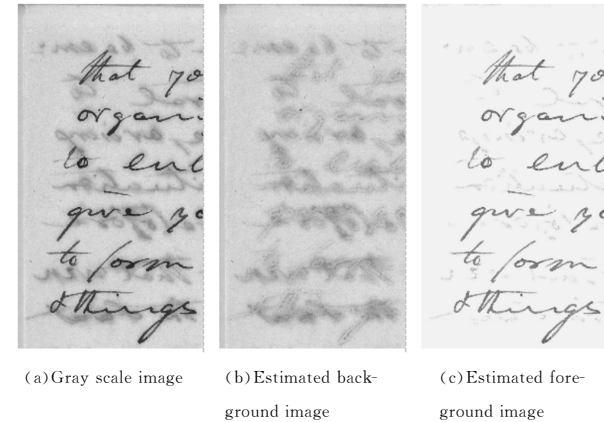


图 5 估计背景和前景图像

Fig. 5 Estimate background and foreground images

(4) 二值化

通过第(3)步可以得到去除了背景的前景图像,对前景图像进行二值化,即可得到一张精度较高的二值化图像。前景图像中的噪声已经得到有效抑制,因此在这一步的二值化方法选用了计算速度较快、二值化效果也相对稳定的 Otsu 方法。

2.5 二值化的集成

具有高召回率的二值化图像,其引入的噪声一般较多;具有高精度的二值化图像,其召回率一般较低。为了弥补两种类型的二值化图像的缺点,本文提出了一种基于连通域的二值化集成方法,该方法集成的思想是在高召回率图像的基础上去除噪声,通过与高精度图像的对比,保留高召回率图像中的前景文本信息,去除噪声的同时,也保证了最终结果的高召回率。

基于连通域的二值化集成方法如下:

(1) 与运算

将高精度二值化图像与高召回二值化图像做一次与运算,与运算后的图像中的前景是高召回率图像和高精度图像共同存在的前景,因此,与运算后得到的二值化图像噪声更少,精度也更高。

(2) 二值化集成

在第(1)步二值化图像的基础上,检测出它所有的连通域,然后遍历每一个连通区域块,将当前区域块与高召回率图像位于同一位置的区域块进行与运算,如果与运算的结果不为空,则从高召回率图像中取出包含该区域块的连通域。遍历完成后,即可得到最终的二值化图像。

二值化集成的结果如图 6(c) 所示。

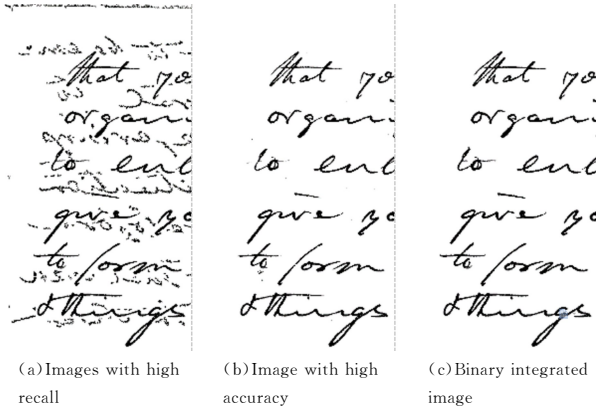


图 6 二值化集成结果图

Fig. 6 Binary integration results

3 实验结果

3.1 实验环境

本文实验环境如表 1 所列。

表 1 实验环境

Table 1 Experimental environment

Environment	Version
OS	Windows 10
CPU	Intel(R) Core(TM) i5-10210U CPU @1.60 GHz 2.11 GHz
RAM/GB	24
ROM/GB	512
Compile Soft-ware	Pycharm Community Edition 2020.2.3
OpenCV	3.3.0.10
Python	3.5.2

3.2 数据集与评价方法

为了验证所提方法的效果,本文在 DIBCO2013^[21] 和 DIBCO2016^[22] 手写文档图像数据集上进行了相关实验。

DIBCO2013 的数据集共有 16 张文本图像,其中 8 张为手写文档图像,8 张为印刷文档图像,本文实验采用的是其中的 8 张手写文档图像,这些文档图像都有严重的退化现象以及很多的噪声,如:文本笔迹细弱、背景颜色复杂、背景有文本重影、纸张颜色退化严重、光照不均匀、部分阴影遮挡、纸张褶皱等。

DIBCO2016 共有 10 张手写文档图像,同样存在光照不均匀、重影、污点、对比度低、纸张退化等问题。这些文本图像包含了现实复杂环境下可能遇到的各种情况。部分复杂背景的图像如图 7 所示。



图 7 复杂情况下的部分手写图像

Fig. 7 Partial handwritten images in complex situations

为了客观评价不同算法的优劣,本文采用的 DIBCO 推荐的指标作为评价指标,包括 F 值(F-Measure, FM),伪 F 值(Pseudo-FMeasure, Fps),峰值信噪比(Peak Signal to Noise Ratio, PSNR),距离倒数失真(Distance Reciprocal Distortion, DRD)^[21]。在文献[21]中可以找到这些评价指标的定义和计算公式。

依据定义,较大的 FM 值和 Fps 值表示更好的二值化图像质量。 $PSNR$ 值越大,则图像失真越小,图像质量越高。 DRD 的值越小,目标图像的视觉失真就越小。

3.3 实验结果与分析

为了测试所提算法的性能,选取了一些有代表性的算法与本文算法进行对比,包括:基于全局阈值的 Otsu 方法、基于局部阈值的 Wolf^[23]方法、经典的 Niblack 方法和 Sauvola 方法、Singh^[24]方法以及在 DIBCO 中取得较好成绩的 Howe 方法。

其中基于局部阈值的方法还涉及参数的设置,本实验中的参数均按照原论文中的建议参数进行设置,具体参数设置如表 2 所列。

表 2 算法参数设置

Table 2 Algorithm parameter settings

Algorithm	Parameter
Wolf	$K=0.5$
Niblack	$K=-0.2$
Sauvola	$K=0.35$
Singh	$K=0.85$

在 DIBCO2013 手写数据集上的实验结果如表 3 所列,在 DIBCO2016 数据集上的实验结果如表 4 所列,表中的数值均为对数据集上的所有图像进行实验后计算得到的平均值。结果表明,本文方法的各个评价指标均优于其他算法,结合评价指标的定义,说明本文方法得到的图像与标准 GT 图像具有较高的相似度且图像质量较好。

表 3 DIBCO2013 手写数据集上的实验结果

Table 3 Experimental results on DIBCO2013 handwritten data

Algorithm	FM	Fps	PSNR	DRD
Otsu	78.191	81.805	18.524	0.2360
Wolf	76.076	81.068	18.646	0.0976
Niblack	22.401	22.683	5.072	0.1601
Sauvola	73.437	80.033	18.109	0.2085
Singh	71.550	67.493	13.659	0.2004
Howe	84.413	85.595	20.016	0.0951
Our	85.236	89.271	20.333	0.0627

表 4 DIBCO2016 手写数据集上的实验结果

Table 4 Experimental results on DIBCO2016 handwritten data

Algorithm	FM	Fps	PSNR	DRD
Otsu	86.586	89.916	17.185	0.3240
Wolf	80.202	88.421	16.888	0.2038
Niblack	34.047	34.750	5.767	0.2792
Sauvola	73.418	81.951	16.377	0.2482
Singh	75.654	76.522	13.790	0.3620
Howe	86.264	90.221	17.245	0.1493
Our	88.280	92.326	17.877	0.0963

考虑到本文方法的预处理流程中使用了 Otsu 进行二值化,为排除这部分算法对结果的影响,本文还使用了 Wolf,

Sauvola, Singh 等二值化算法进行预处理,并在 DIBCO2013 数据集上进行实验,使用不同二值化方法进行预处理的实验结果如表 5 所列。表中的第一行表示 4 种算法得到结果的平均值,其他行数据表示使用该算法进行预处理得到的结果与平均值的差值。由表 5 可以看出, F 值结果在 0.752% 以内浮动, Fps 的浮动范围最大是 1.217%, $PSNR$ 最大浮动值为 0.408, 距离倒数失真结果在 0.026 以内浮动。各项评价指标的结果整体较为稳定,虽然不同的预处理方法对结果有一定影响,但整体的影响较小。

表 5 不同预处理方法在 DIBCO2013 数据集上的实验结果

Table 5 Experimental results of different preprocessing methods on

DIBCO2013 dataset

Algorithm	FM	Fps	PSNR	DRD
Average	84.539	88.575	20.082	0.0367
Otsu	+0.697	+0.696	+0.251	+0.0260
Wolf	+0.043	+0.245	+0.088	-0.0158
Sauvola	+0.013	+0.269	+0.070	-0.0157
Singh	-0.752	-1.217	-0.408	+0.0056

为了更直观地对比不同二值化算法的优劣,我们选取了 DIBCO2013 中的一张手写图像来进行对比分析。不同算法的二值化结果如图 8 所示,图 8(a)是原始图像,其特点是背景含有笔迹浸染的噪声;图 8(b)为 GT 图像,即原始图像对应的标准二值化图像。

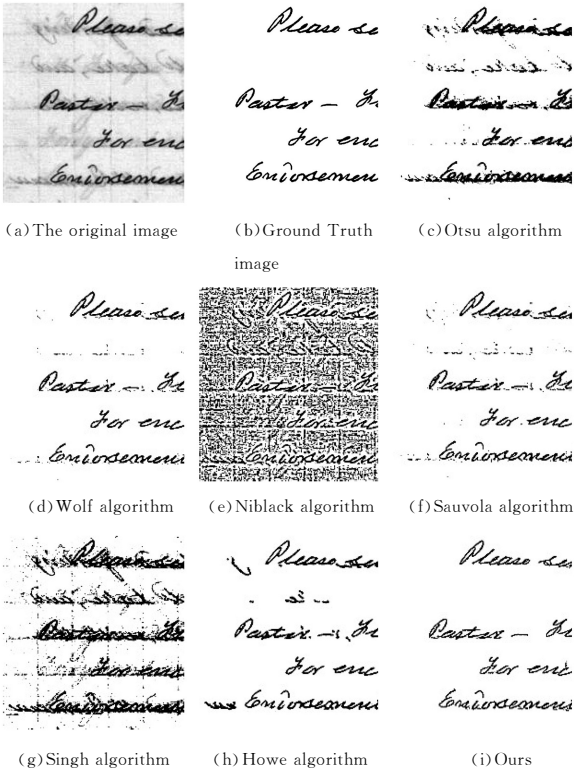


图 8 不同算法的二值化结果图

Fig. 8 Binarization results of different algorithms

图 8(c)是使用 Otsu 算法得到的二值化结果,该方法处理复杂背景图像会产生很多噪声;图 8(d)为 Wolf 算法处理后的结果图,其能正确分辨大部分文本信息,但这种对比度较低的图像中依然会产生少量的噪声;图 8(e)为用 Niblack 算法处理后的结果图,该方法能够很好地保留文本的细节,但

对复杂背景的图像极为敏感,很容易产生椒盐噪声;图 8(f)是 Sauvola 算法的结果图,其作为 Niblack 的改进方法,能够在 Niblack 的基础上正确分辨背景和前景,去除大量非文本噪声,但是该方法仍然会产生一些非文本噪声;从图 8(g)中可以看出,Singh 方法处理复杂背景和低对比度图像后二值化图像质量较差,该方法容易使背景和前景粘连在一起,丢失了笔画细节,也会产生大量的噪声;图 8(h)是使用 Howe 算法得到的结果图,其结果与 Wolf 和 Sauvola 方法的结果相差不大,与它们相比,Howe 能够过滤掉背景中细小的噪声,但也会保留一些较大的背景噪声。

图 8(i)为本文方法得到的结果图,可以看出,本文方法能够很好地处理复杂背景,在保留文本的同时去除掉大量非文本噪声,精度较高。

从图 9 可以看出,本文提出的二值化方法可以处理背景复杂的手写文档图像。它可以消除背景中的许多非文本噪声,而仅将手写笔画保留在前景中。

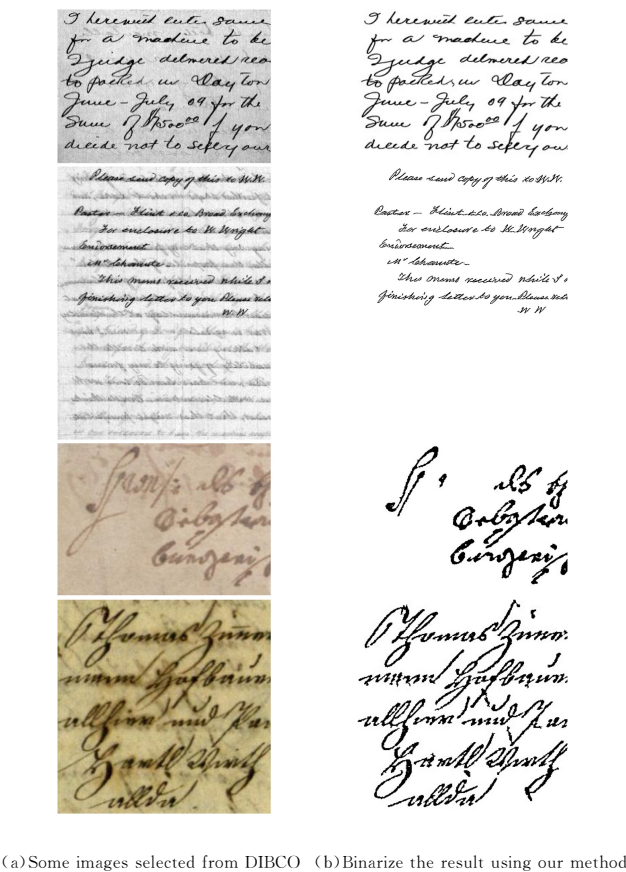


图 9 复杂背景下的二值化

Fig. 9 Binarization in a complex context

表 6 列出了使用不同二值化方法在 DIBCO2013 数据集上进行实验时统计得到的总时间。Otsu 是基于全局阈值的算法,仅计算一次阈值便可以得到结果,因此效率更高。而 Wolf,Niblack,Sauvola 是基于局部阈值的算法,需要不断滑动窗口进行多次阈值计算,耗时更多。Singh 利用局部对比度和均值来计算阈值,由于多次计算对比度非常耗时,因此该算法耗时也较长。Howe 需要多次自动调整参数,用时也较长。本文方法在以下 4 个流程需要进行二值化:预处理、

高召回率图像、背景估计、集成。除此之外,计算笔画宽度和基于连通域的集成都需要遍历图像,这两个步骤也将耗费大量时间。因此,本文方法相比传统二值化方法更加耗时。

表 6 DIBCO2013 手写数据集上的时间对比

Table 6 Time comparison on DIBCO2013 handwritten dataset (单位:s)

Algorithm	Time
Otsu	1.122
Wolf	6.231
Niblack	5.865
Sauvola	5.968
Singh	28.659
Howe	48.063
Our	58.614

本文算法有效的原因在于以下几点:1)结合了高召回率图像和高精度图像的优点,使得集成后的图像同时具有高召回率和高精确率,因此整体图像效果较好;2)使用基于背景估计的方法计算高精度图像,背景估计能够去除大量的非文本噪声,本文也基于背景估计进行了改进,能够自适应得到前景图像,使得最终得到的高精度图像具有较高的召回率;3)根据笔画宽度的不同自适应设置对应的参数,能够适应更多类型的图像,鲁棒性更好。

结束语 本文提出了一种基于改进的背景估计和局部自适应集成的二值化方法,其适用于光照不均、有污点、有重影、纸张退化等背景较为复杂的手写文档图像。当然,该方法也有一定的局限性,如需要进行多次二值化,因此比其他只需一次二值化即可得到结果的方法耗时更长。本文是基于连通域对二值化进行集成,如果高召回率图像中笔画与噪声连接在一起,也会影响最终的二值化结果。因此后续的工作将研究如何更有效率地去除噪声和保留图像中的文本信息。

参 考 文 献

[1] SULAIMAN A,OMAR K,NASRUDIN M F. Degraded historical document binarization: A review on issues, challenges, techniques, and future directions[J]. Journal of Imaging, 2019, 5(4): 48.

[2] OTSU N. A Threshold Selection Method from Gray-Level Histograms[J]. IEEE Transactions on Systems Man & Cybernetics, 1979, 9(1): 62-66.

[3] NIBLACK W. An Introduction to Digital Image Processing [M]. Englewood Cliffs, NJ: Prentice-Hall, Inc, 1986: 115-116.

[4] SAUVOLA J, PIETIKÄINEN M. Adaptive document image binarization[J]. Pattern Recognition, 2000, 33(2): 225-236.

[5] HADJADJ Z, MEZIANE A, CHERFA Y, et al. ISauvola: Improved Sauvola's algorithm for document image binarization [C]// International Conference on Image Analysis and Recognition. Póvoa de Varzim, Portugal: Springer, 2016: 737-745.

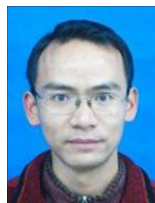
[6] ZHU D L, YANG D G, HU R, et al. Adaptive Multi-level Threshold Binaryzation Method for Optical Character Recognition in Mobile Environment [J]. Computer Science, 2019, 46(8): 315-320.

[7] HOWE N R. Document binarization with automatic parameter tuning[J]. International Journal on Document Analysis and Reco-

- gnition, 2013, 16(3):247-258.
- [8] BHOWMIK S, SARKAR R, DAS B, et al. GiB: a Game theory In-spired Binarization technique for degraded document images [J]. IEEE Transactions on Image Processing, 2018, 28(3):1443-1455.
- [9] BARDOZZO F, DE LA OSA B, HORANSKÁ L', et al. Sugeno integral generalization applied to improve adaptive image binarization [J]. Information Fusion, 2021, 68:37-45.
- [10] RANI U, KAUR A, JOSAN G. A New Contrast Based Degraded Document Image Binarization [J]. Cognitive Computing in Human Cognition, 2020, 17(1):83-90.
- [11] SUN G D, XU Y, XU L, et al. Binarization method of instrument image with uneven illumination and ghosting [J]. Journal of Applied Optics, 2020, 41(1):74-78.
- [12] WESTPHAL F, LAVESSON N, GRAHN H. Document image binarization using recurrent neural networks [C] // 2018 13th IAPR International Workshop on Document Analysis Systems (DAS). Vienna, Austria: IEEE, 2018:263-268.
- [13] KANG S, IWANA B K, UCHIDA S. Complex image processing with less data—Document image binarization by integrating multiple pre-trained U-Net modules [J]. Pattern Recognition, 2021, 109:107577.
- [14] CASTELLANOS F J, GALLEGO A J, CALVO-ZARAGOZA J. Unsupervised neural domain adaptation for document image binarization [J]. Pattern Recognition, 2021, 119(2):108099.
- [15] TENSMEYER C, MARTINEZ T. Historical document image binarization: A review [J]. SN Computer Science, 2020, 1(3):1-26.
- [16] MUSTAFA W A, KADER M M M A. Binarization of document images: A comprehensive review [J]. Journal of Physics: Conference Series, 2018, 1019(1):012023.
- [17] XU H Y, MA L L, WU J, et al. Document image binarization based on background estimation and edge detection [J]. Computer Applications and Software, 2014, 31(8):196-200.
- [18] HEDJAM R, NAFCHI H Z, KALACSKA M, et al. Influence of color-to-gray conversion on the performance of document image binarization: Toward a novel optimization problem [J]. IEEE Transactions on Image Processing, 2015, 24(11):3637-3651.
- [19] WOLD S, ESBENSEN K, GELADI P. Principal component analysis [J]. Chemometrics and Intelligent Laboratory Systems, 1987, 2(1/2/3):37-52.
- [20] ITU-R R, BT. Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios [S]. Geneva, Switzerland: ITU, 2011.
- [21] PRATIKAKIS I, GATOS B, NTIROGIANNIS K. ICDAR 2013 Document Image Binarization Contest (DIBCO 2013) [C] // 2013 12th International Conference on Document Analysis and Recognition. Washington, DC, USA: IEEE, 2013:1471-1476.
- [22] PRATIKAKIS I, ZAGORIS K, BARLAS G, et al. ICFHR2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016) [C] // 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR). Shenzhen, China: IEEE, 2016:619-623.
- [23] WOLF C, JOLION J M. Extraction and Recognition of Artificial Text in Multimedia Documents [J]. Formal Pattern Analysis & Applications, 2004, 6(4):309-326.
- [24] SINGH O I, JAMES O, SINAM T, et al. Local Contrast and Mean based Thresholding Technique in Image Binarization [J]. International Journal of Computer Applications, 2013, 51(6):4-10.



HE Huang-xing, born in 1997, master. His main research interests include image processing and machine learning.



CHEN Ai-guo, born in 1975, associate professor. His main research interests include machine learning, artificial intelligence and pattern recognition.

(责任编辑:何杨)