

基于区域对比度增强的二值化算法

卢 迪 黄 鑫^{*} 柳长源 林 雪 张华玉 严 俊

(哈尔滨理工大学电气与电子工程学院 哈尔滨 150080)

摘 要: 降质文档图像二值化问题是图像处理领域的一个难点。该文通过分析图像不同区域灰度对比度的差异, 为降质文档图像提出了新的二值化算法。首先利用四叉树原理自适应划分区域, 再对不同灰度对比度区域采用不同对比度增强法以调整局部区域内的灰度对比度, 最后根据灰度值出现的频率确定局部阈值。该文测试了随机拍摄的降质图像及 DIBCO (Document Image Binarization COntest) 图像集中的 50 幅图像。与 4 种经典算法比较, 所提算法处理的降质图像具有最高 F-measure 值和峰值信噪比 (PSNR 值)。

关键词: 图像处理; 二值化; 区域对比度增强; 局部阈值; 四叉树法

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2017)01-0240-05

DOI: 10.11999/JEIT160197

Binarization Method Based on Local Contrast Enhancement

LU Di HUANG Xin LIU Changyuan LIN Xue ZHANG Huayu YAN Jun

(School of Electrical and Electronic Engineering, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: Binarization for degraded document images is a difficult point in image processing. This paper presents a new binarization method for the degraded document images by analyzing the differences of image grayscale contrast in different areas. Firstly, theory of quadtree is used to divide areas adaptively. Secondly, various contrast enhancements are selected to adjust local grayscale contrast for different contrast areas. Lastly, the frequency of gray value is utilized to calculate threshold. The proposed algorithm is tested on random shooting degraded images and datasets of Document Image Binarization COntest (DIBCO). Compared with other four classical algorithms, the binarized images using the proposed algorithm gain the highest F-measure and PSNR (Peak Signal-to-Noise Ratio).

Key words: Image processing; Binarization; Local contrast enhancement; Local threshold; Quadtree

1 引言

光学字符识别技术 (Optical Character Recognition, OCR) 是将纸质文字信息录入计算机的最高效方法, 但其主要不足表现为无法从降质文档图像中提取清晰字符, 即降质文档图像二值化过程存在问题。目前图像二值化方法主要为基于阈值的二值化方法, 又可分为全局阈值法和局部阈值法。全局阈值法 (如 Otsu 法^[1]) 对处理目标和背景灰度分布均匀且类偏差不大的图像有良好效果, 但对降质文档图像, 全局阈值法会出现大量的误分割^[2]。局部阈值法是根据局部像素点信息确定阈值, 更适用于二值化降质文档图像。1986 年, 文献[3]提出利用局

部均值和标准差计算阈值的自适应二值化算法, 该算法二值化的图像字符笔画较完整, 但图像背景会产生连续大块非目标黑色区域。2000 年, 文献[4]在文献[3]算法基础上进行改进, 引入标准方差动态范围参数 R 降低文献[3]算法中对修正参数 K 的依赖。该算法对大部分降质文档图像有较好的处理效果, 但对背景特别亮或特别暗的图像, 处理效果较差。2007 年, 文献[5]提出的改进 FCM 算法利用偏移场有效地降低了非均匀光照干扰, 但分割结果目标丢失严重。2010 年, 文献[6]提出基于图像区域的二值化算法, 根据预先设定的分块数将图像分块, 再采用 Otsu 法全局分割, 由于其需事先指定分块数, 所以自适应性受到了很大程度限制^[7]。2014 年, 文献[8]提出一种自适应二值化算法, 分 4 步对图像进行处理, 依次是: 局部对比分析、对比度扩展、阈值选取和噪声移除。文献[8]算法对文档填充充实的降质图像分割效果好, 但处理图像中只有少量字符的文

收稿日期: 2016-03-03; 改回日期: 2016-07-12; 网络出版: 2016-09-30

*通信作者: 黄鑫 scorpion_hx@163.com

基金项目: 哈尔滨市科技创新人才项目 (2014RFQXJ163)

Foundation Item: The Science and Technology Innovation Talents Project of Harbin (2014RFQXJ163)

档图像，会产生大量的噪点。

当前技术大多是针对某一特定降质原因的图像有较好的二值化效果^[9]，但降质原因多种多样，好的二值化算法应能处理多种原因形成的降质图像。降质文档图像的特点是部分区域目标和背景的灰度对比度明显，部分对比不明显或整体对比均不明显^[10]。基于此，本文设计了一种基于区域对比度增强的二值化算法，可解决受多类原因影响的降质文档图像的二值化问题。

2 本文工作

2.1 区域划分

区域划分的好坏直接影响着不同对比度增强法能否在相应区域发挥最大效用，灰度对比度可作为划分依据。设 $F(x, y)$ 是一幅 $M \times N$ 的灰度图像，定义像素点 (x, y) 处灰度对比度 $D(x, y)$ 为

$$D(x, y) = \max[D_h(x, y), D_v(x, y)] \quad (1)$$

$$D_h(x, y) = |f(x, y) - f(x-1, y)| \quad (2)$$

$$D_v(x, y) = |f(x, y) - f(x, y-1)| \quad (3)$$

其中， $D_h(x, y)$ 和 $D_v(x, y)$ 分别为像素点 (x, y) 水平方向和竖直方向的灰度对比度。对于降质文档图像，需要对图像进行划分来确定不同区域的灰度对比度分布情况。

2.1.1 区域粗划分 本文利用四叉树原理，根据灰度对比度对图像进行区域划分如图 1。第 1 次四分将待处理图像分为 A, B, C, D 4 个子区域，若子区域 C 内的最大灰度对比度小于 a 倍整体区域的最大灰度对比度(如式(4))，说明该区域灰度变化不显著，认为该子区域不含字符均为背景，直接输出该子区域。

$$D_{C\max}(x, y) \leq a \cdot D_{et\max}(x, y) \quad (4)$$

其中， $D_{C\max}(x, y)$ 为第 1 次四分后子区域 C 的最大灰度对比度， $D_{et\max}(x, y)$ 为整幅图像的最大灰度对比度。 a 为前、背景划分系数。经此步骤后可剔除大块背景区域，减少计算量。

2.1.2 区域细划分 粗划分后不满足式(4)的区域均被认为是含字符的目标区域，本文假设图 1 中 A, B,

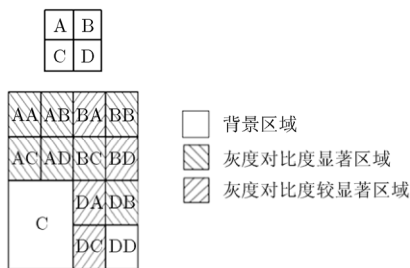


图 1 区域划分原理示意图

D 为目标区域。对降质文档图像，粗划分后的剩余区域间的灰度对比度也可能存在明显差异，因此需要对剩余区域再次细分。以区域 D 为例(A, B 区域的细分同 D 区域)，若第 2 次四分后子区域 DD 的最大灰度对比度小于 a 倍四分前区域的最大灰度对比度(如式(5))，直接输出该区域。

$$D_{DD\max}(x, y) \leq a \cdot D_{D\max}(x, y) \quad (5)$$

若子区域 DB 的最大灰度对比度大于 b 倍四分前区域的最大灰度对比度(如式(6))，说明该区域的灰度变化非常显著，则对该区域进行弱灰度对比度增强算法处理。

$$D_{DB\max}(x, y) \geq b \cdot D_{D\max}(x, y) \quad (6)$$

若子区域 DA 内的最大灰度对比度介于 a 倍和 b 倍的四分前区域的最大灰度对比度之间(如式(7))，说明该区域的灰度变化较显著，则对该区域进行强灰度对比度增强算法处理。

$$a \cdot D_{D\max}(x, y) \leq D_{DA\max}(x, y) \leq b \cdot D_{D\max}(x, y) \quad (7)$$

其中， $D_{DD\max}(x, y)$ ， $D_{DB\max}(x, y)$ ， $D_{DA\max}(x, y)$ 分别为第 2 次四分后子区域 DD, DB 及 DA 的最大灰度对比度， $D_{D\max}(x, y)$ 为第 1 次四分后子区域 D 的最大灰度对比度， b 为灰度变化显著与较显著区域的划分系数。

实验表明 $a \in [0, 0.4]$ ， $b \in [0.7, 1]$ ，且对图像进行两次四分来确定区域的灰度变化属性是最佳四分次数。过多的四分会导致计算量过大，容易将噪声点误分成目标，无法将噪点处理掉。过少的四分虽然会减小计算量，但同时将丢失过多细节，达不到理想的分割效果。

2.2 区域对比度增强

2.1 节将图像划分成灰度变化不显著区域、非常显著区域及较显著区域。文档图像通常为黑色目标及白色背景，因此对灰度变化不显著区域，将区域内各像素点灰度值调整为

$$f'(x, y) = 255 \quad (8)$$

对于灰度变化非常显著区域，采用弱灰度对比度增强方式^[11]，将区域内各像素点灰度值调整为

$$f'(x, y) = (n-1) \times \frac{f(x, y) - f_{\min}(x, y)}{f_{\max}(x, y) - f_{\min}(x, y)} \quad (9)$$

对于灰度变化较显著区域，本文基于弱灰度对比度增强方式提出了强灰度对比度增强方式，其目的是更大程度地拉大区域内各像素点间的灰度对比度。此时区域内各像素点灰度值调整为

$$f'(x, y) = (nn-1) \times \left(\frac{f(x, y) - f_{\min}(x, y)}{f_{\max}(x, y) - f_{\min}(x, y)} \right)^2 \quad (10)$$

式(9)，式(10)中， $f(x, y)$ 为原始区域图像像素点灰度

值, $f_{\max}(x, y)$ 和 $f_{\min}(x, y)$ 为原始区域图像像素点最大灰度值和最小灰度值。 n 和 m 为调整后区域图像的灰度级。

本文的灰度增强法可有效调整非均匀光照图像、透背图像及变化背景图像像素点的灰度值, 进而解决这 3 类图像的二值化问题。对由字迹透背或受非均匀光照引起的降质文档图像, 图像中字迹透背、光照太亮或太暗区域, 其前景与背景间对比度相差不大, 可归为对比度较显著区域, 采用式(10)强对比度增强法可分离出清晰字符。对文档降质不严重的区域, 其前景与背景间灰度对比度存在明显差异, 此时采用式(9)弱对比度增强法可减少噪声。对存在变化背景的降质文档图像, 其背景间的对比度变化远比不上前景与背景之间的对比度变化。因此, 若区域中不存在字符只存在对比度变化相对较大的背景时, 通过式(8)即可将它剔除; 若区域中既存在字符又存在变化背景, 则可以通过两类对比度增强法, 拉大像素点间的对比度, 进而从背景中分离出前景。图 2 为强、弱灰度对比度增强法的二值化效果对比图。通过图 2 可以看出, 对灰度变化程度不同的区域, 应用不同灰度对比度增强法可取得较好的分割效果。

2.3 区域阈值选取

经对比度增强法调整后的灰度图像直观上可明显地区分出目标和背景。分析文档图像特点可知, 通常情况下, 文档图像字符像素点数目小于背景像素点数目, 极少数可能存在大于的情况。因此可通过累加灰度直方图中灰度值所对应的像素点数目来确定代表前背景的灰度值。对于经灰度对比度增强法处理后的子区域图像(其大小为 $s \times t$), 寻找累加值最接近 $0.5(s \times t)$ 时所对应的灰度值 n_{hf} , 在 $0 \sim n_{\text{hf}}$ 中出现频率最高的灰度值即为前景点的灰度值 f'_{fg} , 在 $n_{\text{hf}} \sim n$ 中出现频率最高的灰度值即为背景点的灰度值 f'_{bg} 。若在 $0 \sim n_{\text{hf}}$ (或 $n_{\text{hf}} \sim n$) 中存在多个频率最高的灰度值, 则选取最小(或最大)的灰度值来代表前景(或背景灰度值)。子区域的阈值 T 为 f'_{fg} 和 f'_{bg} 的均值, 即: $T = 0.5(f'_{\text{fg}} + f'_{\text{bg}})$ 。最终, 可获得子区域的二值化图像 $g(x, y)$ 为

$$g(x, y) = \begin{cases} 1, & f'(x, y) > T \\ 0, & \text{其它} \end{cases} \quad (11)$$

3 实验结果分析

3.1 实验环境和测试图像库

本文采用如下仿真实验环境: Intel(R) Core(TM)i3-3240 CPU, 4G 内存, Matlab(R2011a) 编译环境。本文测试了随机拍摄的手写体图像和印刷体图像以及来自 DIBCO 图像集^[12]的 50 幅图像。

3.2 分割效果测试

分别应用 Otsu 算法, Niblack 算法, Sauvola 算法和 Singh 算法与本文算法进行对比实验。图 3~图 5 为各算法处理后的二值图像。除 Otsu 算法外均存在参数选取问题, 因此本文结合参考文献^[2,3,6], 对 Niblack 算法、Sauvola 算法选取窗口大小为 5×5 , 15×15 , 25×25 , 35×35 和 50×50 做 5 组实验, 对 Singh 算法选取子块大小为 32×32 , 64×64 , 128×128 , 256×256 和 512×512 做 5 组实验, 本文算法以 0.1 为步长, 随机在 $k_1 \in [0, 0.4]$, $k_2 \in [0.7, 1]$ 中选取 5 组 a , b 值测试, 以上实验均选取具有最佳 F - measure 值的二值图像作为最终实验结果。

图 3~图 5 分别为受光照不均匀、字迹透背及字迹模糊影响的降质文档图像。由图 3~图 5 可以看出, 无论是传统的 Otsu 算法、Niblack 算法或 Sauvola 算法还是先进的 Singh 算法, 经它们处理后降质文档图像的二值图像, 或存在笔画断裂、模糊不清的现象, 或存在大块黑色非目标区域的现象, 或存在笔画浓度不一致的现象。针对降质文档图像, 本文算法可智能地筛选目标区域及非目标背景区域, 避免噪声干扰。经本文算法处理后的二值图像, 保留了字符的完整性, 分离出清晰的字符, 大大地方便了后续计算机的识别。

3.3 评估指标

高质量的二值图像具有较高的 F - measure 值和 PSNR 值^[6], 计算公式如式(12), 式(13):

$$FM = \frac{2 \times RC \times PR}{RC + PR} \quad (12)$$

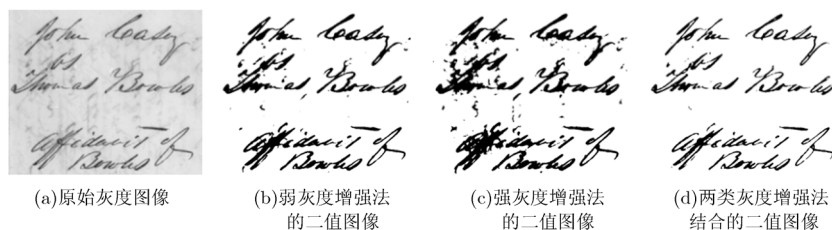


图 2 强、弱灰度对比度增强法的二值化效果对比图

JPEG 标准本身只是定义了一个规范的编码数据流,并没有规定图像数据文件的格式。Cable Microsystems 公司定义了一种 JPEG 文件交换格式(JFIF)。JFIF 图像是一种或者使用灰度表示,或者使用 Y、Cb、Cr 分量彩色表示的 JPEG 图像。它包含一个与 JPEG 兼容的文件头。一个 JFIF 文件通常包含单个图像,图像可以是灰度图像,其中的数据为单个分量;也可以是彩色图像,其中的数据包括 3 个分量。

JPEG 标准本身只是定义了一个规范的编码数据流,并没有规定图像数据文件的格式。Cable Microsystems 公司定义了一种 JPEG 文件交换格式(JFIF)。JFIF 图像是一种或者使用灰度表示,或者使用 Y、Cb、Cr 分量彩色表示的 JPEG 图像。它包含一个与 JPEG 兼容的文件头。一个 JFIF 文件通常包含单个图像,图像可以是灰度图像,其中的数据为单个分量;也可以是彩色图像,其中的数据包括 3 个分量。

JPEG 标准本身只是定义了一个规范的编码数据流,并没有规定图像数据文件的格式。Cable Microsystems 公司定义了一种 JPEG 文件交换格式(JFIF)。JFIF 图像是一种或者使用灰度表示,或者使用 Y、Cb、Cr 分量彩色表示的 JPEG 图像。它包含一个与 JPEG 兼容的文件头。一个 JFIF 文件通常包含单个图像,图像可以是灰度图像,其中的数据为单个分量;也可以是彩色图像,其中的数据包括 3 个分量。

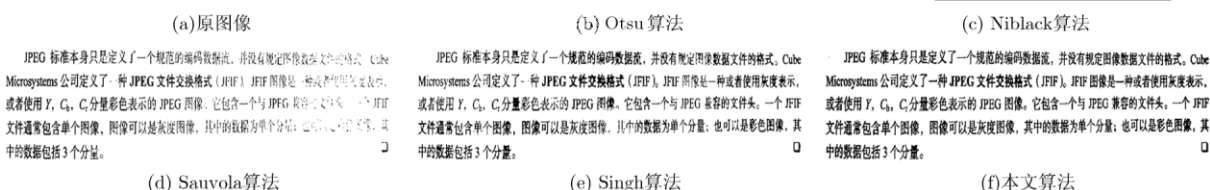


图 3 随机拍摄印刷体文字图像二值化结果

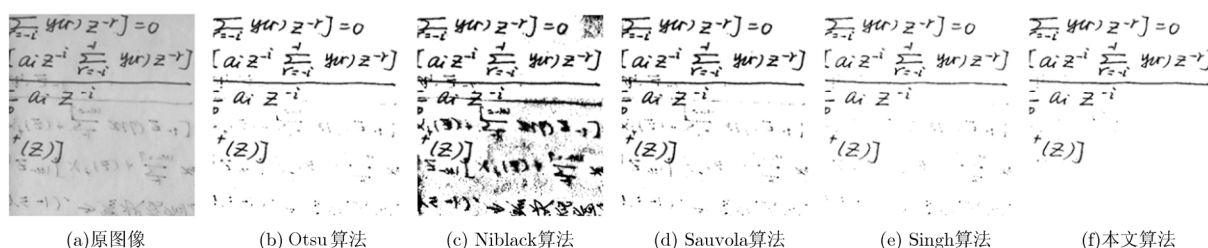


图 4 随机拍摄手写体文字图像二值化结果

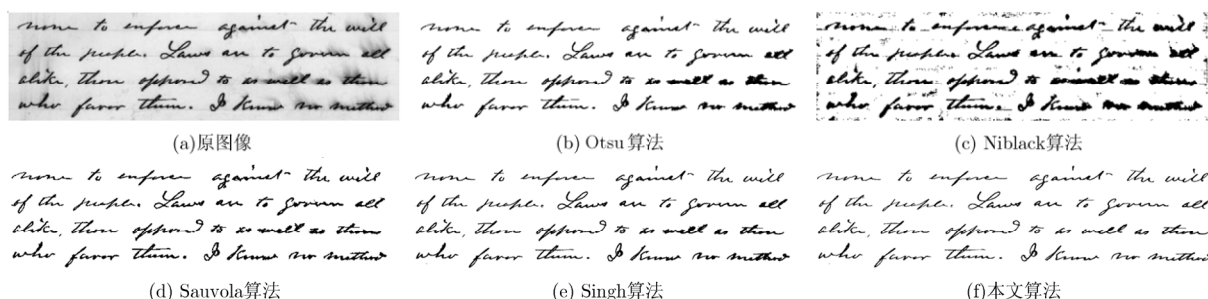


图 5 H-DIBCO2012 H08 二值化结果

$$\text{PSNR} = 10 \lg \left(\frac{C^2}{\text{MSE}} \right) \quad (13)$$

其中,PR 和 RC 分别代表准确率及召回率,MSE 为标准二值图像与输出二值图像间的均方误差。表 1 为 5 种算法处理 DIBCO 图像集 50 幅降质图像的平均 F - measure 值和 PSNR 值。

由表 1 可以看出,在 5 种算法处理降质文档图像的测试结果中,应用本文算法处理的二值图像,具有最高的 F - measure 值(高于次高 Otsu 法 4%)、最高的 PSNR 值(高于次高 Sauvola 法 5%),说明本文算法可最好地还原目标像素点,同时说明应用本文算法输出的二值图像最接近实际未降质文档的二值图像。

3.4 基于 OCR 评估

测试各算法二值图像在 OCR(识别正确率)上的识别效果可作为评价各算法二值效果好坏的重要依据。因此,本文利用图 3,测试各算法处理的二值化图像在 ABBYY Fine Reader 12 上的识别正确率(具体数据见表 2)。本文定义 OCR 识别正确率为

$$\text{识别正确率} = \frac{\text{正确识别字符数}}{\text{字符总数}} \times 100\% \quad (14)$$

结合图 3 及表 2 可知,虽然原灰度图像受非均匀光照影响,但在 ABBYY Fine Reader 上仍有 88.94% 的识别正确率。经 Otsu 算法及 Sauvola 算法

表 2 图 3 中各图像在 ABBYY FineReader 上的识别正确率

表 1 不同算法对 DIBCO 图像集的测试结果

方法	Otsu 算法	Niblack 算法	Sauvola 算法	Singh 算法	本文 算法
F-measure(%)	77.81	48.97	72.93	70.50	80.94
PSNR (dB)	15.07	8.09	15.14	13.74	15.90

	原始 灰度 图像	Otsu 算法	Niblack 算法	Sauvola 算法	Singh 算法	本文 算法
识别正 确率 (%)	88.94	77.89	96.98	57.29	91.46	98.49

处理后的二值图像,由于图像右下角区域字符笔画断裂,因此在 ABBYY Fine Reader 上识别效率较低。经 Niblack 算法处理后的二值图像,虽然图像中存在黑色噪声块,但由于带字符的前景区域字符笔画连续,因此识别率可达到 90%以上。经 Singh 算法处理后的二值图像,大部分字符保留较完整,因此识别率也可达到 90%以上。经本文算法处理后的二值图像,字符清晰、完整、易于识别,在 ABBYY Fine Reader 上可达到 98%以上的识别正确率,说明了本文二值化算法的有效性。

4 总结

本文根据区域间灰度对比度的差异,为降质文档图像提出新的二值化算法。该算法自适应地划分灰度对比度显著与较显著区域,对灰度对比度显著区域,本文采用弱对比度增强法调整前景与背景间的灰度差异,有效地抑制了噪声的干扰。对灰度对比度较显著区域,本文提出强对比度增强法来调整灰度差异,进而分离出清晰字符。因此不论任何区域(包括变化背景、非均匀光照、透背等区域),总有适合的方法去处理。且本文算法对由严重非均匀光照及透背引起的降质文档图像尤为有效。实验结果表明,在 5 种算法处理 DIBCO 图像集得到的结果中,本文算法得到的二值图像,字符笔画清晰、完整,背景极少出现噪声,且具有最高的 F-measure 值及 PSNR 值。将各种算法得到的二值图像应用到 ABBYY FineReader12,本文算法具有最高的识别正确率。

参 考 文 献

- [1] OTSU N. A threshold selection method from gray level histograms[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1979, 9(1): 62-66. doi: 10.1109/TSMC.1979.4310076.
- [2] 申铨京, 龙建武, 陈海鹏, 等. 三维直方图重建和降维的 Otsu 阈值分割算法[J]. *电子学报*, 2011, 39(5): 1108-1114.
SHEN Xuanjing, LONG Jianwu, CHEN Haipeng, *et al.* Otsu thresholding algorithm based on rebuilding and dimension reduction of the 3-dimensional histogram[J]. *Acta Electronica Sinica*, 2011, 39(5): 1108-1114.
- [3] NIBLACK W. An Introduction to Digital Image Processing [M]. Englewood Cliffs, NJ, US, Prentice-Hall, Inc., 1986: 115-116.
- [4] SAUVOLA J and PIETIKAINEN M. Adaptive document image binarization[J]. *Pattern Recognit*, 2000, 33(2): 225-236. doi: 10.1016/S0031-3203(99)00055-2.
- [5] MA L and STAUNTON R C. A modified fuzzy C-means image segmentation algorithm for use with uneven illumination patterns[J]. *Pattern Recognition*, 2007, 40(11): 3005-3011. doi: 10.1016/j.patcog.2007.02.005.
- [6] CHOU C H, and LIN W H, and CHANG F. A binarization method with learning-build rules for document images produced by cameras[J]. *Pattern Recognition*, 2010, 43(4): 1518-1530. doi: 10.1016/j.patcog.2009.10.016.
- [7] 龙建武, 申铨京, 臧慧, 等. 高斯尺度空间下估计背景的自适应阈值分割算法[J]. *自动化学报*, 2014, 40(8): 1773-1782. doi: 10.3724/SP.J.1004.2014.01773.
LONG Jianwu, SHEN Xuanjing, ZANG Hui, *et al.* An adaptive thresholding algorithm by background estimation in Gaussian scale space[J]. *Acta Automatica Sinica*, 2014, 40(8): 1773-1782. doi: 10.3724/SP.J.1004.2014.01773.
- [8] SINGH B M, SHARMA R, GHOSH D, *et al.* Adaptive binarization of severely degraded and non-uniformly illuminated documents[J]. *International Journal of Document Analysis and Recognition*, 2014, 17(4): 393-412. doi: 10.1007/s10032-014-0219-6.
- [9] MESQUITA R G, MELLO C A B, and ALMEIDA L H E V. A new thresholding algorithm for document images based on the perception of objects by distance[J]. *Integrated Computer-Aided Engineering*, 2014, 21(2): 133-146. doi: 10.3233/ICA-130453.
- [10] MILYAEV S, BARINOVA O, NOVIKOVA T, *et al.* Fast and accurate scene text understanding with image binarization and off-the-shelf OCR[J]. *International Journal of Document Analysis and Recognition*, 2015, 18(2): 169-182. doi: 10.1007/s10032-015-0240-4.
- [11] ROSENFELD A and KAK A C. Digital Picture Processing [M]. 2nd ed. New York, Morgan Kaufmann: Academic Press, 1982: 92-95.
- [12] PRATIKAKIS I, GATOS B, and NTIROGIANNIS K. DIBCO 2013-Document Image Binarization Contest [EB/OL]. <http://utopia.duth.gr/~ipratika/DIBCO2013/>, 2013.

卢迪: 女, 1971 年生, 教授, 研究方向为数据融合、故障检测、图像处理、智能交通系统。

黄鑫: 女, 1991 年生, 硕士生, 研究方向为图像处理。

柳长源: 男, 1971 年生, 副教授, 研究方向为模式识别与智能信息处理技术、机器学习与人工智能、数字图像处理。