2009 10th International Conference on Document Analysis and Recognition

A Modified Adaptive Logical Level Binarization Technique for Historical Document Images

K. Ntirogiannis, B. Gatos and I. Pratikakis
Computational Intelligence Laboratory,
Institute of Informatics and Telecommunications,
National Center for Scientific Research "Demokritos",
GR-153 10 Agia Paraskevi, Athens, Greece
{kntir,bgat,ipratika}@iit.demokritos.gr

Abstract

In this paper, a new document image binarization technique is presented, as an improved version of the state-of-the-art adaptive logical level technique (ALLT). The original ALLT depends on fixed windows to extract essential features such as the character stroke width. Since characters with several different stroke widths may exist within a region, this can lead to erroneous results. In our approach, we use local adaptive binarization as a guide to our adaptive stroke width detection. The skeleton and the contour points of the binarization output are combined to identify locally the stroke width. Additionally, we introduce an adaptive local parameter "β" that enhances the characters and improves the overall performance. In this way, we achieve more accurate binarization results in both handwritten and printed documents with a particular focus on degraded historical documents. Experimental results prove the effectiveness of the proposed technique compared to other state-of-the-art methodologies.

1. Introduction

Document image binarization is of great significance for the Optical Character Recognition (OCR). It is an OCR preprocessing stage hence its outcome directly affects further stages of the OCR analysis pipeline. Binarization methodologies can be classified into global and local adaptive. In the case of local adaptive, the adaptation is based on parameters which are usually related to document characteristics. Representative examples of those methodologies are the logical level techniques (LLT) which take into account the character stroke width (SW).

Previously, other authors have addressed LLTs for binarization. Kamel and Zhao [1] presented a LLT for document image binarization with the main idea of character SW preservation. Yang and Yan [2] developed an adaptive version of the former named as ALLT (Adaptive Logical Level Technique), in which improvements comprise the automatic SW detection and the adaptive local threshold.

The binarization techniques based on the SW are robust against large stains, noise and borders. Problems arise in documents with different font sizes where the use of a single SW value is not effective. In particular, characters with original SW larger than the estimated tend to result in characters with different topology, while characters with original SW smaller than the estimated result in broken characters. In [3] an improvement of the ALLT is presented in which different SW is assigned in different blocks. Although this technique, in several cases overcomes the problem of font size variability, it is likely to introduce the blocking effect due to the image block division. Moreover, in the case of historical documents which contain characters with holes or faint characters along with noise, the gray-level and run-length histogram of the ALLTs may lead to erroneous SW detection.

In this paper, we further improve the original ALLT by using a state-of-the-art adaptive binarization technique [4] as a guide to our adaptive SW detection. The skeleton and the contour points of the binarization output are combined to identify locally the SW at pixel and connected component (CC) level for the handwritten and printed historical documents respectively. Additionally, we introduce an adaptive local parameter "*β*" that leads to character enhancement and improves the overall performance. Furthermore, for a fast implementation of our technique, we adopted an approach similar to [5].

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.225

The remainder of the paper is organized as follows. In Section 2, the description of current logical level binarization techniques is given. In Section 3, our modified logical level binarization technique is detailed and in Section 4 the experimental results are demonstrated. The conclusion is drawn in Section 5.

2. Current Logical Level Techniques

Originally, Kamel and Zhao [1] used the character SW in the document image binarization. Computations are performed according to SW and a global threshold, which are both predetermined by the user. In particular, the gray level or the smoothed gray level of each processing point P is compared with four local averages located in (2SW+1)x(2SW+1) windows centered at two pairs of diametric opposites.

In mathematical terms, the logical level principle is denoted as in the following.

$$b(x,y)=\begin{cases}1, \text{ if } \bigcup_{\mathrm{i}=0}^{3}[L(P_i)\cap L(P_i^{'})\cap L(P_{i+1})\cap L(P_{i+1}^{'})] \\ \quad \text{is true} \\ 0, \text{ otherwise}\end{cases} \tag{1}$$

$$L(P)=\begin{cases}true, \text{ if } \mathrm{ave}(P)-g(x,y)>T \\ false, \text{ otherwise}\end{cases} \tag{2}$$

$$\mathrm{ave}(P)=\frac{\sum\limits_{-sw\le i\le sw}\ \sum\limits_{-sw\le j\le sw} g(P_x-i,P_y-j)}{(2\mathrm{SW}+1)^2} \tag{3}$$

where $\mathrm{P}_{\mathrm{i}}^{'}=\mathrm{P}_{(\mathrm{i}+4)\mathrm{mod}8}$, P_{x} and P_{y} are the coordinates of the point P and g(x,y) is the grey level of the original image or its smooth counterpart.

The LLT [1] is user-dependent since both the SW and the global threshold are assigned by the user. Moreover, a global threshold is difficult or even impossible to be set for historical document images. In this respect, Yang and Yan [2] proposed the Adaptive LLT (ALLT) in which the SW is automatically detected and the threshold is locally adapted.

In ALLT [2], the input image is divided into NxN (N=4,…8) regions with the aim of finding some local areas with quasi-bimodal histogram (bi-level sub-images). Specifically, the histogram analysis is performed within regions of the two diagonal directions if N is even, while if N is odd, the analysis is additionally performed in the vertical and horizontal directions. The quasi-bimodal regions are used for the run-length histogram analysis and the SW is defined as the run-length with the highest frequency.

The other improvement proposed in ALLT concerns the local adaptive threshold. For each processing point P, the min, max and average (ave) grey value are calculated within a (2SW+1)x(2SW+1) window centered at P. It is worth mentioning that if |max-ave| = |min-ave| then the window expands to (2SW+3)x(2SW+3) and calculations are performed once more. Overall, the adaptive threshold T is produced as follows:

$$T=\begin{cases}a\cdot(\frac{2}{3}\min+\frac{1}{3}\mathrm{ave}), \text{ if } |\max-\mathrm{ave}|>|\min-\mathrm{ave}| \\ a\cdot(\frac{1}{3}\min+\frac{2}{3}\mathrm{ave}), \text{ if } |\max-\mathrm{ave}|<|\min-\mathrm{ave}| \\ a\cdot\mathrm{ave}, \text{ if } |\max-\mathrm{ave}|=|\min-\mathrm{ave}|\end{cases} \tag{4}$$

where *α* is a global predetermined parameter between 0.3 and 0.8, while 2/3 is recommended for most cases.

Problems arise in documents with variable font sizes where the single SW value of the ALLT is not effective. In those cases, characters with SW larger than the one detected tend to result in characters with different topology while characters with original SW smaller than the estimated result in broken characters.

In [3], Badekas and Papamarkos proposed a new ALLT method to preserve the characters in documents with font size variability. Specifically, the image is divided in NxM sub-images and the global binarization technique of Otsu [6] is performed at the sub-images which have almost bi-level histograms. The proper SW deriving from the corresponding local run-length histogram is given to each bi-level sub-image. The average of the SW values from neighboring regions is assigned to regions without a SW value.

The purpose of the previously mentioned technique [3] was to correctly assign SWs at the sub-images. However, it is not guaranteed that the initial bi-level sub-images contain characters of one SW, neither the neighboring sub-images include characters of the similar size to the bi-level ones. Additionally, in noisy sub-images with bi-level histograms the run-length procedure will favor the noisy parts and destroy the characters.

3. The Proposed Modified Adaptive Logical Level Technique

Our proposed modified ALLT does not take into account any sub-mages which may mislead the SW detection and create binarization outputs with blocking effects. On the contrary, the proposed adaptive SW detection is performed using the skeleton and the contour points of a guide local adaptive binarization output, avoiding the image block division and run-length histograms. Moreover, a modified local adaptive threshold offers robustness and enhances the performance. A flowchart of the proposed technique is presented in Fig. 1.

Since our technique is based on local averaging, we incorporate the acceleration similar to [5] for fast implementation. According to [5], for any window size the local averaging requires only four additions operations instead of summation over all window pixels.

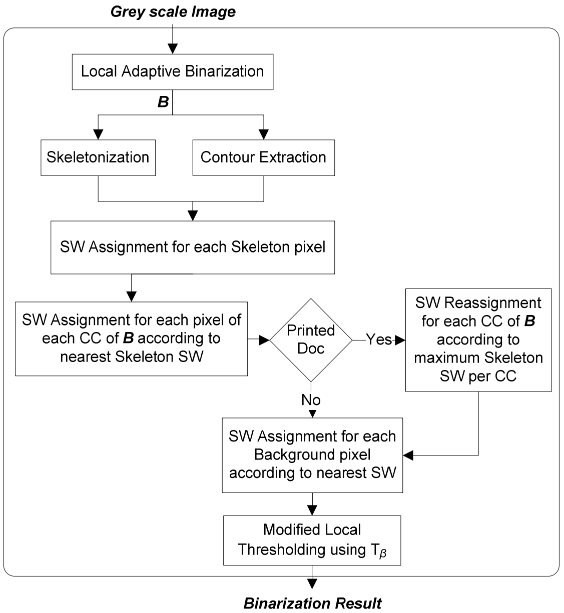


Figure 1. Flowchart of the proposed modified adaptive logical level binarization technique.

3.1. Adaptive Stroke Width Detection

Instead of using a global threshold or searching for bi-level regions we used an adaptive local thresholding technique for the SW detection. Taking into account [7], which is a recent evaluation study of several binarization techniques, we selected the GPP algorithm [4] which performs best and has the lowest false alarms rate.

Once the binarization using [4] is performed, we extract the contour points C(x,y) and the skeleton S using the skeletonization method of [8]. For each skeleton point, we calculate the smallest distance R from the contour points C(x,y) and the corresponding SW is equal to 2R+1. The remainder points of each CC inherit the SW value of nearest skeleton point. Finally, the background points inherit the nearest foreground SW found. Representative SW maps are shown in Fig. 2(e) and 3(c).

Our extensive study on degraded and historical document images has shown that the most difficult cases concern printed characters with holes and faint handwritten ones. In those cases, the current logical level techniques that use local run-length histograms for the SW detection are highly prone to errors.

In the case of printed characters with holes, the adaptive SW detection will not effectively assign the proper SW at each point, since the skeleton will surround the holes and be closer to the contour points (Fig. 2(d)). Hence, the adaptive stroke width detection is performed at CC level. Specifically, right after the foreground SW detection, each CC is assigned with the maximum of its SW values (Fig. 2(f)). Figure 2 demonstrates the effectiveness of the SW detection at CC level for the degraded printed document images. The final image has less and smaller holes and less noise than the initial binary image.

As far as the faint handwritten characters are concerned, the initial adaptive pixel-level SW detection deals effectively with them, since the broken part of the binary image will inherit the proper nearest SW (Fig. 3(c)).

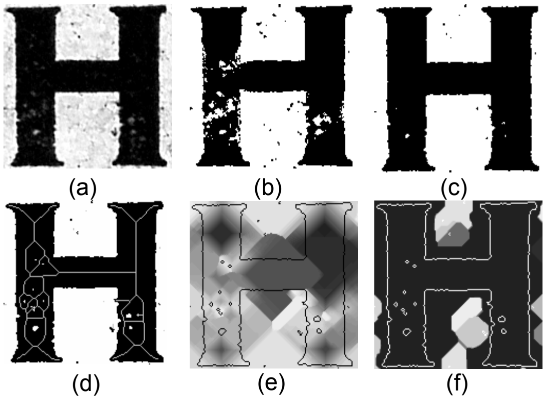


Figure 2. (a) The original image, (b)-(c) the binarization output using the SW image at pixel and connected component level respectively, (d) the binary image along with the skeleton, (e)-(f) the SW image at pixel and connected component level respectively along with the contour points.

3.2. Local Adaptive Threshold Modification

The global parameter $\boldsymbol{\alpha}$ of the adaptive local threshold (Eq. 4) cannot be effectively tuned for both precision and recall. Therefore, we introduce an adaptive local parameter $\boldsymbol{\beta}$ that decreases the threshold T according to the contrast level. In this way, we achieve higher recall rates at the same precision enhancing the overall performance (Fig. 3(e), (f)).

Eq. 4 that denotes the local adaptive threshold T is modified to T_β as follows:

$$T_{\beta} = \begin{cases} \beta \cdot T, & \text{if } |\max - \text{ave}| \neq |\min - \text{ave}| \\ T, & \text{if } |\max - \text{ave}| = |\min - \text{ave}| \end{cases} \quad (5)$$

where β is denoted as:

$$\beta = \begin{cases} (\frac{\text{ave}}{\max})^2, & \text{if } |\max - \text{ave}| < |\min - \text{ave}| \\ (\frac{\min}{\text{ave}})^2, & \text{if } |\max - \text{ave}| > |\min - \text{ave}| \end{cases} \quad (6)$$

The interpretation of parameter β is related to the text enhancement against the background noise. In homogenous areas (background, stains) the min, ave and max values are similar and β is almost equal to one, while in text areas, they are different and β decreases as the max or min deviate from the average.

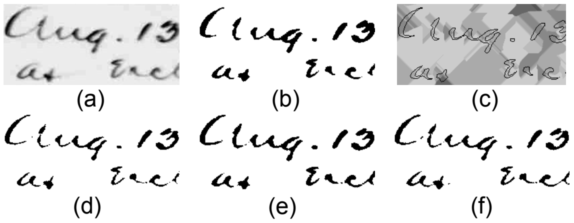

(a) (b) (c)
(d) (e) (f)

Figure 3. (a) The original image with faint characters, (b) the binary image, (c) the pixel-level SW image along with the contour points, (d) ALLT binarization using adaptive SW only, (e) ALLT binarization using the proposed modified adaptive local threshold only, (f) ALLT binarization using both the adaptive SW and the proposed modified adaptive local threshold.

4. Experimental Results

All experiments were conducted using 40 historical printed and handwritten images from both the Library of Congress [9] and European Libraries [10]. Most of the images contain severe degradations, such as shadows, non-uniform illumination, stains, bleeding and faint characters, low quality and other artifacts. All experimental results were addressed in the evaluation framework proposed in [7].

We studied the performance of the original ALLT compared to the proposed modified ALLT. In Fig. 4 and 5 the F-Measure is plotted as a function of the global parameter "*a*" concerning the handwritten and the printed document images, respectively. The use of our adaptive SW improves the performance but remains close to the original ALLT performance. The incorporation of the adaptive local parameter "β" boosts the performance and offers robustness. The combination of the aforementioned improvements offers even higher performance. Our experiments have shown that the best performance is achieved for values equal to 0.2 and 0.4 for the global parameter "*a*" concerning the historical handwritten and printed images respectively.

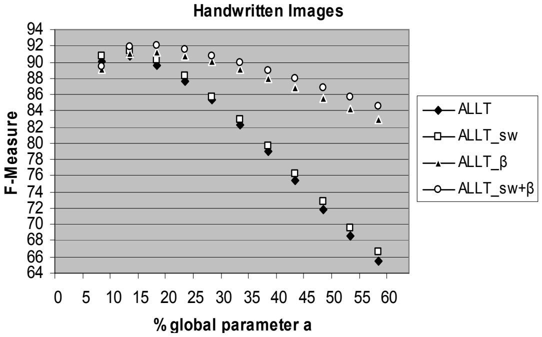


Figure 4. The performance of the ALLT along with the proposed improvements concerning the handwritten images.

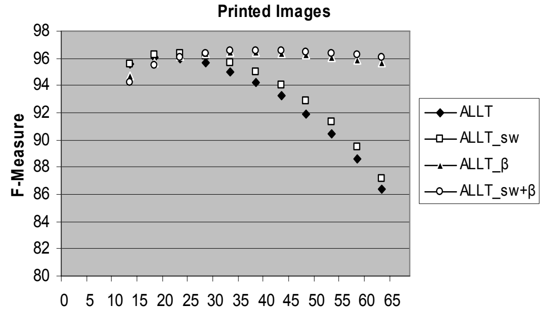
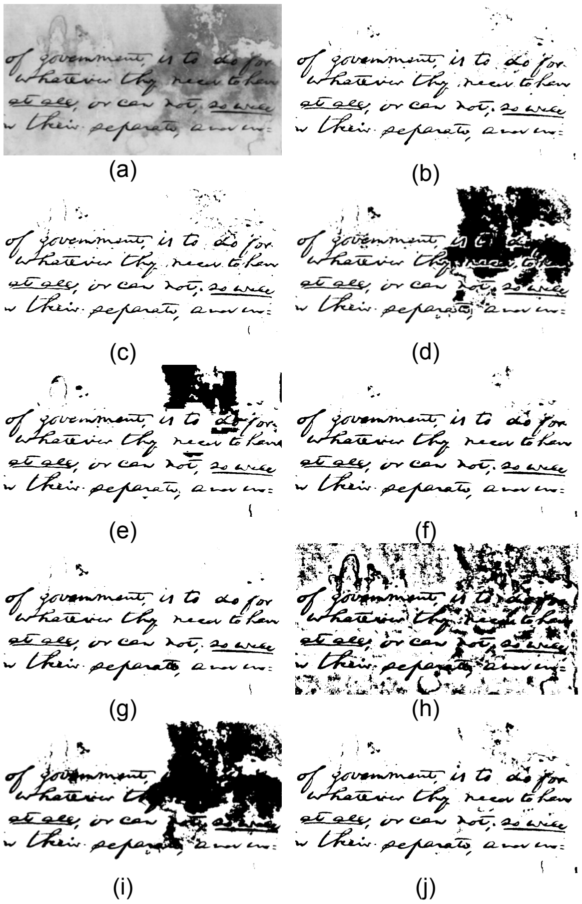


Figure 5. The performance of the ALLT along with the proposed improvements concerning the printed images.

Several state-of-the-art binarization methods are compared with our proposed Modified ALLT (MA):
1) ALLT (AL) [2], 2) Bernsen (BR) [11], 3) FineReader (FR) [12], 4) Gatos et. al. (GP) [4], 5) Kim et. al (KM) [13], 6) Niblack (NB) [14], 7) Otsu (OT) [6], 8) Sauvola et. al. (SV) [15].
In Fig. 6 representative results of all the binarization techniques are demonstrated.

For the evaluation, we used approximately the recommended by the authors values when predetermined parameters were required. Specifically, for the AL, we used the results where the precision rate is similar with our best to clarify the improvement. In those values, the AL had the best performance. Table 1 summarizes the evaluation results.

Furthermore, in Table 1 we present the evaluation results of MA for three different values of ***a***. In handwritten images with faint characters and degradations the recall rate is important. In Table 1 MA0 has high recall rate while others (BR, KM, SV) achieve this rate with lower precision rates. On the other hand, most of the printed document images do not include faint characters, thus the recall rate is very high. In Table 1 MA1 and MA2 achieve the highest precision rate and the highest performance.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 6. (a) Original image, (b) MA binarization, (c) AL binarization, (d) BR binarization, (e) FR binarization, (f) GP binarization, (g) KM binarization, (h) NB binarization, (i) OT binarization and (j) SV binarization.

Table 1. The evaluation results using the F-Measure (FM), Recall (Rec) and Precision (Prec).

	Handwritten				Printed		
	FM	Rec	Prec		FM	Rec	Prec
MA0	91.93	92.10	91.76		96.54	97.64	95.47
MA1	92.02	90.32	93.77		96.59	97.24	95.94
MA2	91.48	88.55	94.61		96.55	96.79	96.31
AL	90.71	88.16	93.41		96.05	96.27	95.84
BR	89.52	92.08	87.09		95.79	97.93	93.74
FR	88.71	93.21	84.62		94.87	98.34	91.63
GP	91.49	91.21	91.78		96.10	98.86	93.49
KM	91.51	92.25	90.78		95.50	96.83	94.22
NB	42.97	99.23	27.42		74.63	99.87	59.57
OT	88.09	93.82	83.02		94.39	98.92	90.27
SV	90.79	92.10	89.51		95.14	99.08	91.50

5. Conclusion

In this paper a new improved adaptive logical level technique is presented that detects effectively the text information and improves the performance. The improvements comprise the adaptive SW detection at pixel and CC level and the new efficient adaptive local threshold. The experimental results along with the performance study of our technique prove its effectiveness in the document image binarization.

Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n° 215064 (project IMPACT).

References

[1] M. Kamel and A. Zhao, "Extraction of Binary Character/Graphics Images from Grayscale Document Images", *CVGIP: Comput. Vision Graph. Imag. Process.*, vol. 55, no. 3, pp. 203-217, May 1993.

[2] Y. Yang and H. Yan, "An Adaptive Logical Method for Binarization of Degraded Document Images", *Pattern Recognition*, vol. 33, pp. 787-807, 2000.

[3] E. Badekas and N. Papamarkos, "A System for Document Binarization", *Proc. of the 3rd Inter. Symposium on Image and Signal Processing and Anal.*, 2003, pp. 909-914.

[4] B. Gatos, I. Pratikakis and S. J. Perantonis, "Adaptive Degraded Document Image Binarization", *Pattern Recognition*, vol. 39, no. 3, pp. 317-327, Mar. 2006.

[5] F. Shafait, D. Keysers and T.M. Breuel, "Efficient Implementation of Local Adaptive Techniques Using Integral Images", *SPIE Proc., Document Recognition and Retrieval XV*, Jan. 2008, vol. 6815, pp. 681510-6.

[6] N. Otsu, "A Thresholding Selection Method from Gray-level Histogram", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, pp. 62-66, Mar. 1979.

[7] K. Ntirogiannis, B. Gatos and I. Pratikakis, "An Objective Evaluation Methodology for Document Image Binarization Techniques", *Proc. of the 8th IAPR Inter. Workshop on Document Anal. Sys.,* Sept. 2008, pp. 217-224.

[8] H.J. Lee and B. Chen, "Recognition of Handwritten Chinese Characters via Short Line Segments", *Pattern Recognition*, vol. 25, no. 5, pp. 543-552, May 1992.

[9] Library of Congress (http://memory.loc.gov/ammem)

[10] Impact Project (http://www.impact-project.eu)

[11] J. Bernsen, "Dynamic Thresholding of Grey-level Images", *Proc. of the 8th Inter. Conf. Pattern Recognition*, Paris, France, 1986, pp. 1251-1255.

[12] ABBYY (www.finereader.com)

[13] I.K. Kim, D.W. Jung and R.H. Park, "Document Image Binarization Based on Topographic Analysis Using a Water Flow Model", *Pattern Recognition*, vol. 35, pp. 265–277, 2002.

[14] W. Niblack, An Introduction to Digital Image Processing, Englewood Cliffs, NJ PrenticeHall, 1986, p. 115.

[15] J. Sauvola and M. Pietikainen, "Adaptive Document Image Binarization", *Pattern Recognition*, vol. 33, no. 2, pp. 225-236, 2000.