National Journal of Parallel and Soft Computing, Volume 01, Issue 01, March-2019

A Robust Local Thresholding Method for Segmentation of Non-Uniform Coloured Documents

Thidar Aung, Thin Lai Lai Thein
University of Computer Studies, Yangon,Myanmar
moethidar@ucsy.edu.mm, tllthein@ucsy.edu.mm

Abstract

Text document segmentation is one of the essential steps text document recognition and extraction systems. The existing segmenting methods are not much reliable for text with colour gradient and texture background. Also, long processing time of existing methods is unfavorable.

Thus, aiming to have reliable segmentation and less processing time, a new local thresholding method is proposed in this study. The proposed method is based on pixel intensity magnification concept. Non-overlapping local window is used. The performance of proposed method is compared with that of Otsu's method and Niblack's method. The results show that the proposed method give more reliable segmentation compared to Otus' method, 1.2 times longer processing time. Compared to Niblack's method, the segmented results are comparable and the processing time is 5 times faster.

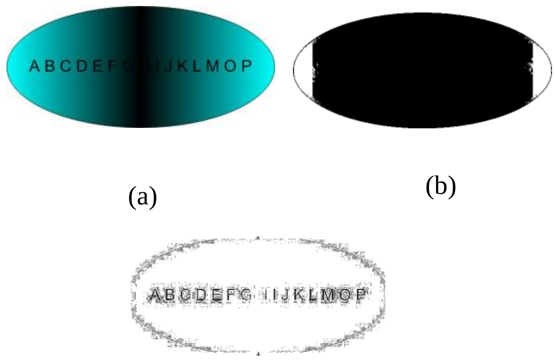
Keywords*: segmentation, local threshloding, pixel magnification, text document recognition, character recognition*

1. Introduction

Today, OCR (Optical Character Reorganization) becomes popular for automatic retrieval information from image version of text documents, vehicle license plates, business cards, invoices, ID cards, driver licenses, etc. Generally, there are three main steps to perform optical character recognition. These are pre-processing of image that contains targeted characters, segmentation of characters and character recognition. In the first step, some processes such as image quality enhancement, skewness correction are performed if required. In the second step, the characters are segmented from the background and further fragmented as individual character. Finally, in the last step, each character is recognized by matching with pre-known features of each character.

As stated above, the character segmentation involves segmenting all contained characters from the background and fragmentation all characters as individual one. There are many methods to segment the characters from the background. Most frequently used methods are global thresholding method and colour based character detection. However, the drawback of these methods is the failure of segmenting characters when there is colour gradient in text images. Moreover, these methods are also not applicable if there is non-uniform illumination in text images.

For this reason, in these days, many researchers started to propose local thresholding methods for segmenting characters from the background colour. The local thresholding perform better than global thresholding but it is even not perfect for some conditions.



(a) (b)

(c)

Figure 1. Text image segmentation (a) original image (b) result using global thresholding (c) result using local thresholding

Figure 1 shows the original image, and segmented result using global thresholding (Otsu's method) and local thresholding (Niblack's method). It can be seen that local thresholding shows the better performance than global thresholding. However the result has not reached to a satisfactory level. There are some noises around each character, which can interrupt in further processes of OCR. The other critical of local thresholding method is long processing time which is not acceptable in real-time applications.

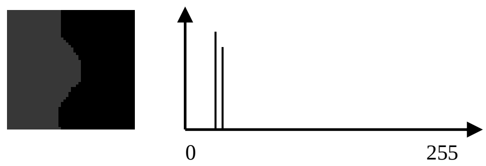
There have been many works in which character segmentation (binarization) was performed by using different thresholding methods. A very popular thresholding method was proposed by Otsu [1]. However, this method is global thresholding method and it is suitable only for document images without illumination effect. Then, Niblack [2] developed a local thresholding technique which is based on local mean and standard deviation. Its default local window size is 15×15. Sauvola and Pietikainen [3] also proposed local threholding method by improving Niblack's method. Thus, this improved method is also based on local mean and standard deviation. In the method proposed by Singh et al.,[4], only local mean is used without standard deviation. Then, LABB (Local Adaptive Automatic Binarizationan) method proposed by Singh [5] is also related to local mean. In the work of Bernsen [6], the local thresholding method is based on local contrast. If local contrast is greater than 15, the local threshold is half of sum of local maximum and local minimum. Otherwise, Otsu's global threshold is used as a local threshold. Imocha Singh [7] proposed a threshold method which is based on local mean, local maximum and local minimum.

According to this survey, most local thresholding methods in previous works are based on local mean and a bias constant (k). In most methods, the required bias constant is described with a range, which necessitates the condition-dependent adjustment of k's value. The selected value of bias constant (k) can affect the performance of the method. Thus, a simpler, more robust, more time-efficient local thresholding method is still in demand.

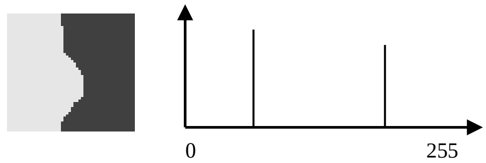
Thus, the objectives of this study are to propose a robust local thresholding method and to experimentally verify the performance of proposed method by comparing with the performance of existing methods, (Otsu's method, Niblack’s method). The rest of this paper is organized as follows. The system bock diagram of proposed method is explained in section 2. In section 3, the detail of proposed method is described. The tests conducted and the results are discussed in section 4 which is followed by conclusion.

2. Concept of Magnification

Since the concept of magnification is used in this proposed method, it is explained in advance. When a region has slightly different pixel intensity it is very tricky to segment using a certain threshold. As shown in Figure 2(a), the dominant intensity levels in the region are nearly equal and it is difficult to set a threshold for segmentation. Thus, it can easily fail in segmentation. When these dominant pixel intensities are magnified with a certain factor, the pixel intensities become higher. However, the higher pixel intensities will go higher while the lower intensities will just remain at slightly higher level compared to original level. As depicted in Figure 2(b), the two intensity levels become far different after magnification. It makes the segmentation easy. Here, the task is to find an appropriate magnification factor for the region according to pixel intensity contained. It is proposed in this work.



(a)



(b)

Figure 2. Concept of magnification (a) before magnification (b) after magnification

3. Proposed Method

3.1 Flow Chart

The overall flow chart of proposed method is shown in Figure 3. Firstly, image to be segmented is acquired. Then, the image is enhanced by using edge sharpening method. The image is chopped into non-overlap multiple windows. Afterwards, the number of dominant pixels, minimum pixel intensity, maximum pixel intensity and intensity range in each window are calculated. Based on these parameters, the magnification factor for each window is calculated. Then, each window is segmented by multiplying with corresponding magnification factor. Finally, dilation process is performed for the image if required. Each process in the flow chart is explained in details in the following section.

3.2 Acquiring Images

The offline image is read by program. The image is true colour image and image type is JPEG. Different image sizes are used. For every image has the resolution is 96 dpi.

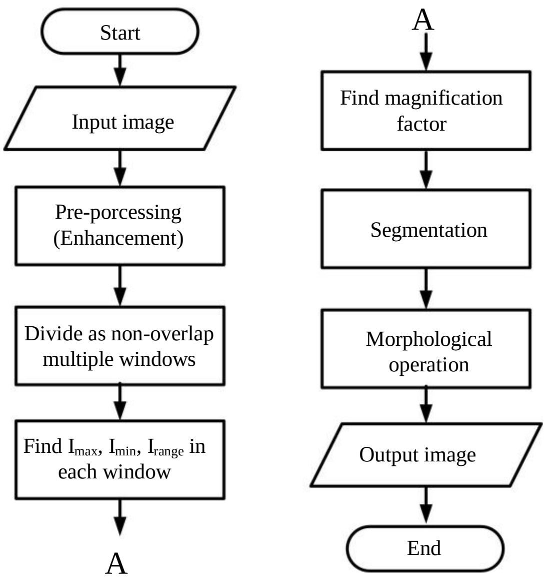


Figure 3. Flow chart of proposed method

3.3 Edge Sharpening

The image quality is enhanced by using edge sharpening method. Edge sharpening is the process for enhancing edges of the regions in the image. The mathematical description of edge sharpening can be written as below.

$$I_n = k\,(I_g - \alpha\, I_f) \qquad (1)$$

where, I_n is sharpened image, I_g is original grey image, I_f is smoothened image, k and α are scale factors. The value of k and α can be taken inversely. For example, when the value of α is 0.5, k should be 2 so that the original image is obtained back. The detailed discussion for the edge sharpening can be seen in [8]. For smoothing image, average (mean) filter with window size of 30×30 is used. The true colour image is converted to grey scale with the following relationship.

$$I_g = 0.3\,R + 0.6\,G + 0.1\,B \qquad (2)$$

where, R is red intensity, G is green intensity and B is blue intensity of the true colour image.

3.4 Dividing Image into Subwindows

The enhanced image is divided into multiple non-overlapped subwindows. Using non-overlapping subwinodws is one of the distinct features of proposed method. The objective of using non-overlapped windows is to reduce processing time. It significantly reduces the processing time compared to exiting methods.

For any size (m×n) of input image, the subwindow size is 30x30. Thus, the number of subwindows for the whole Figure is (m/30 × n/30). If the last windows have less than 30 rows or 30 columns are added to the second-last window. One example of this process is show in Figure.4. For the image shown in Figure.4, there are 80 subwindows. Here, the size of windows 1, 2, 3, 4 are 30×30, but the rest have more than 30×30 since they are combination of the second-last window and the last window.

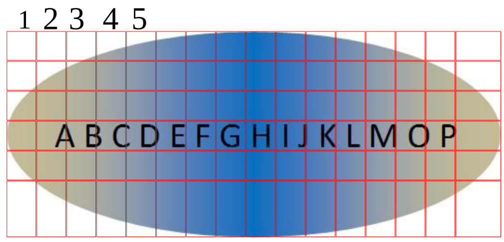


Figure 4: Dividing image into non-overlapping windows

3.5 Searching Dominant Intensity Levels

The number of dominant intensities in a subwindow is searched by creating pixel histogram for the window. The window size is 30×30 and there is a total of 900 pixel. From the intensity histogram of these 900 pixels, the dominant pixels are searched. In the histogram, if the frequency of a pixel value is more than 10, it is accounted as dominant pixel. As shown in Figure 5(a), in windows-1, 2, there is only one dominant intensity level, and in windows-3, 4, 5 there are two dominant intensity levels as shown in Figure 5(b). In a window, there can also be many dominant intensity levels.

From this dominant intensity levels, maximum intensity level (I_{max}), minimum intensity level (I_{min}), intensity range (I_{range}) are extracted.

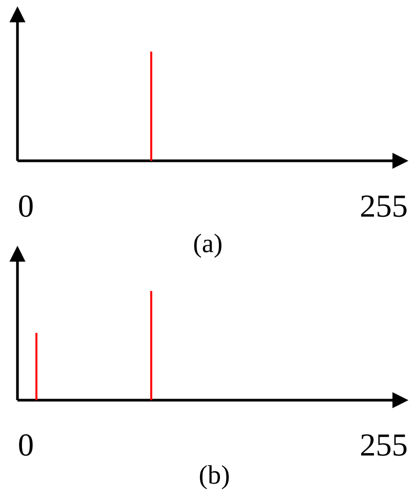


Figure.5 Dominant intensity level in windows (a) for window-1 (b) for window-7

3.6 Calculating Magnification Factor

In this proposed method, the value of magnification factor depends on the number of dominant pixel intensity, maximum intensity, minimum intensity and intensity range as follows.

For single dominant intensity level,

$$F = \begin{cases} 255 / I_{min} & \text{if } I_{min} \geq 1 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

For two dominant intensity levels,

$$F = \begin{cases} 0 & \text{if } I_{min} < 1 \ \& \ I_{range} < 20 \\ 255 / I_{min} & \text{if } I_{min} \geq 1 \ \& \ I_{range} < 20 \\ 255 / I_{max} & \text{if } I_{min} < 1 \ \& \ I_{range} \geq 20 \\ 255 / I_{max} & \text{if } 1 < I_{min} < 30 \ \& \ I_{range} \geq 20 \\ 255 / I_{min} & \text{if } I_{min} > 30 \ \& \ I_{range} \geq 20 \end{cases} \qquad (4)$$

For more than two dominant intensity levels,

$$F = \begin{cases} 0 & \text{if } I_{min} < 1 \ \& \ I_{range} < 20 \\ 255 / I_{min} & \text{if } I_{min} \geq 1 \ \& \ I_{range} < 20 \\ 255 / I_{k-n} & \text{if } I_{min} < 1 \ \& \ (I_{k-1} - I_{min}) < 10 \ \& \ I_{range} \geq 20 \\ 255 / I_{k-1} & \text{if } I_{min} < 1 \ \& \ (I_{k-1} - I_{min}) \geq 10 \ \& \ I_{range} \geq 20 \\ 255 / I_{k-n} & \text{if } 1 < I_{min} < 30 \ \& \ (I_{k-1} - I_{min}) < 10 \ \& \ I_{range} \geq 20 \\ 255 / I_{k-1} & \text{if } 1 < I_{min} < 30 \ \& \ (I_{k-1} - I_{min}) \geq 10 \ \& \ I_{range} \geq 20 \\ 255 / I_{min} & \text{if } I_{min} \geq 30 \ \& \ I_{range} \geq 20 \end{cases} \qquad (5)$$

As one example, the values of magnification factors (F) for windows 1 to 96 for the image shown in Figure 4 are shown in the following matrix.

1	1	1	2	2	2	2	3	3	3	2	2	2	2	1	1
1	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2
2	1	2	2	2	2	2	2	3	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2
1	2	2	2	2	2	3	3	3	3	2	2	2	2	2	2

3.7 Segmentation and Morphological Operation

After obtaining the value of F for the corresponding subwindow, all of the pixels in the subwindow are multiplied by F. The magnified window is segmented by setting a threshold of 150. After differentiation the intensities as minimum (0) and maximum (255), a half value (128) is the suitable and acceptable threshold value. Thus, a somewhat larger value 150 is used for segmentation.Afterwards, the binary image undergoes dilation process. The detailed dilation process can be seen in [8].

4. Results and Discussion

4.1 Text Images

A numbers of images are tested to verify the performance of proposed method. Images are created for different conditions such as text in uniform colour, text in multi colour, text in gradient background, text in national ID, text with watermark, highlight text, text in light illumination, text in passport, text in bank card. Some images are created and some are taken from image source.

4.2 Segmented Results

The original images and segmented image using proposed method is shown in Figure.6(a)-(f). For text in uniform colour and multi-colour background, shown in Figure.6(a), (b), (c), the segmented results are very clear and no text is lost. For the text in gradient colour background, shown in Figure.6(d), the segmented result is acceptable since all text are successfully segmented. However, the result is not 100% perfect since some noises are left at some locations in the result. It is due to colour gradient interface. It also depends on coluor and gradient direction. For the results depicted in Figure. 6(e), the segmented image is perfect. Also, it can bee seen that the proposed method give perfect result for multi-colour highlighted text in Figure. 6(f).

One can also notice that the watermark in Figure. 6(g) is perfectly removed. Then, the proposed method is tested for national ID which have pink texture background. The result proves that the proposed method is still perfect for removing texture background from national ID.

Due to limited space, the results of other test images are not shown here. The other text images are text in light illumination, text in passport, text in bank card. The proposed method still performs well for those types of images.

4.3 Performance Evaluation

The performance of proposed method is compared with that of Otsu's method and Niblack's method. The performance parameters are accuracy and processing time. The accuracy is calculated as equation (6).

$$\beta(\%) = \frac{\Psi - \Phi}{\Psi} \times 100 \qquad (6)$$

where, β is the accuracy, Ψ is the number of characters in original image, Φ is the number of characters in segmented image. The processing time is CPU time in second.

Table 1 shows the comparison of accuracy and processing time for eight different test images. Here, one can see that Otsu's method is not totally reliable for gradient images. For tested images, the lowest accuracy of proposed method is 84% for Figure 6(e) and and it can give mostly 100% accuracy for Figure 6(a), (b), (c), (f), (g), (h). The accuracy of proposed method is comparable that of Niblack's method.

Otsu's method the smallest processing time for every image since it is global threhsolding method. Compared to Otsu's method, the processing time of proposed method is 1.2 times longer. However, Compared to Niblack's method, the processing time is 5 times faster.

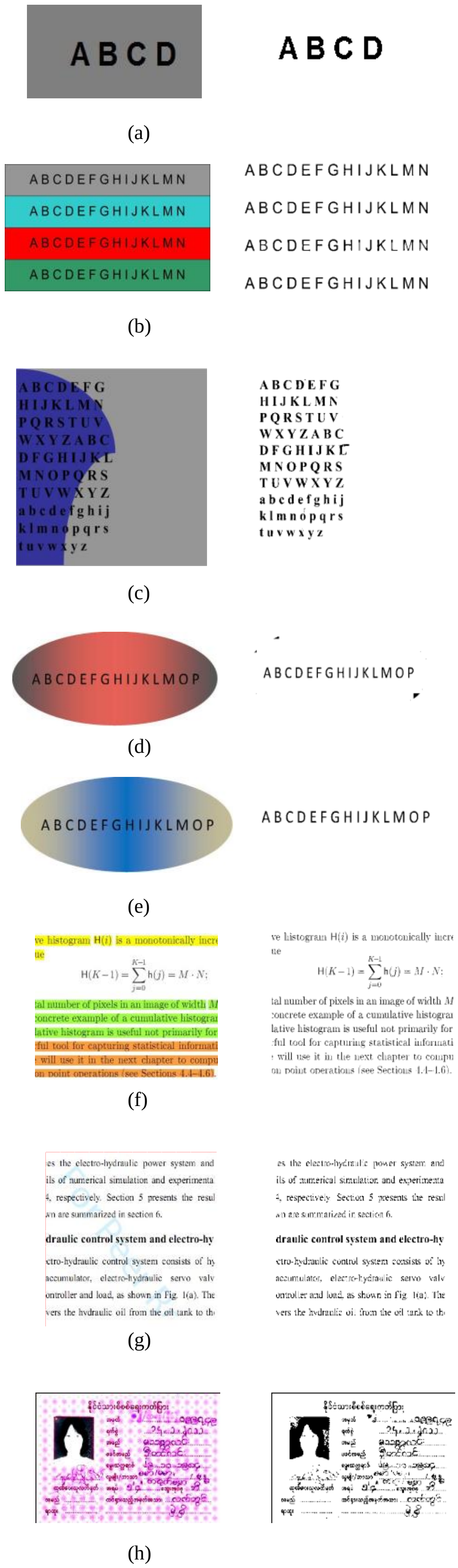


Figure 6: Test images and results (Original images are in left column and results are in right column)

Table 1. Performance comparison

Image No	Otsu's method		Niblack's method		Proposed method	
	β (%)	Time (s)	β (%)	Time (s)	β (%)	Time (s)
(a)	100	0.1094s	100	0.7969s	100	0.2969s
(b)	75	0.1406s	75	5.2656s	100	0.6250s
(c)	31	0.1406s	31	5.2500s	100	0.9373s
(d)	0	0.0938s	0	2.6406s	95	0.3906s
(e)	0	0.1094s	0	2.0301s	84	0.4063s
(f)	63	0.1875s	0	3.5000s	100	0.4375s
(g)	100	0.0781s	60	2.1719s	100	0.6719s
(h)	100	0.1250s	30	1.0781s	100	0.5781s

5. Conclusion

In this work, the proposed method is based on pixel intensity magnification concept. Non-overlapping local window is used.The performance of proposed method is tested with a numbers of images and compared with the performance of existing methods; Otsu's and Niblack's method. The proposed method is applicable for segmenting different kinds of text images such as text in uniform colour, text in multi colour, text in gradient background, text in national ID, text with waterrmark, highlight text, text in light illumination, text in passport, text in bank card.

Compared to Otsu's method, the proposed method is much more reliable for segmenting gradient colour images. The processing time is 2 time longer. The accuracy of proposed method is comparable to that of Niblack's method while it is giving 5 times faster processing time.

References

[1] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Systems, Man, and Cybernetics 9(1), pp.62-66 , 1979.

[2] Niblack, W.: 'An introduction to digital image processing' (Prentice- Hall, Englewood Cliffs, NJ, 1986), pp. 115–116

[3] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," Pattern Recognition 33(2), pp. 225–236, 2000.

[4] T. Romen Singh, Sudipta Roy, and Kh.Manglem Singh, "Local Adaptive Automatic Binarization (LAAB)", International Journal of Computer Applications 40(6), 2012

[5] J. Bernsen,v'Dynamic thresholding of gray-level images'. Proc. 8th Int. Conf. on Pattern Recognition, Paris, 1986, pp. 1251–1255

[6] F. Rukhsar Firdousi and P. Shaheen," Local Thresholding Techniques in Image Binarization", International Journal Of Engineering And Computer Science, 3(3), pp.4062-4065, 2014

[7] F. Rukhsar Firdousi and P. Shaheen," Local Thresholding Techniques in Image Binarization", International Journal Of Engineering And Computer Science, 3(3), pp.4062-4065, 2014

[8] F. Rukhsar Firdousi and P. Shaheen," Local Thresholding Techniques in Image Binarization", International Journal Of Engineering And Computer Science, 3(3), pp.4062-4065, 2014