Asian Journal of Information Technology 15 (16): 2927-2936, 2016
ISSN: 1682-3915

A Novel Simple Thresholding for Uneven Illuminated Document Images Captured via Handheld Devices

[1]Musab Kasim Alqudah, [1]Mohammad Nasrudin, [2]Bilal Bataineh and [1]Khairuddin Omar
[1]Centre for Artificial Intelligence Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
[2]Department of Computer Science, Umm Al-Qura University, Mecca, Saudi Arabia

Abstract: The prevalence of handheld devices such as mobile phones for image capturing now a days is uncontested which is mainly contributed by their high image quality, low cost and portability. Hence, it is only natural that users would prefer the convenience of documents image capturing to photocopying or scanning. Images taken from these devices however, usually succumb to uneven illuminations of which photocopied or scanned documents are spared. This study expounds a simple thresholding technique to eliminate unwanted background of document images captured by handheld devices. The technique administers a local thresholding which is a simple computation of the mean and standard deviation of sliding windows procured from the whole image. We appraise the efficacy of method visually and analytically. In analytical experiments we use 5 uneven illuminated document images captured from a 5MP mobile phone camera and 13 images of DIBCO. Results of several performance measurements from the technique is compared with those generated from several state-of-the-art thresholding methods. The results attest to the efficacy of the Otsu technique in eliminating background, albeit in the absence of uneven-illumination and a large gap between image pixels contrast. In addition, the techniques propounded by Sauvola, Nick and Bataineh perform well with uneven illumination cases, but fall short in low contrast and when text pixels values are too close to the foreground or background pixels values. Unlike the proposed technique, which perform well in both circumstances. This substantiates the claim that performance of the proposed method is superior to that of the other techniques in overcoming uneven illumination especially on shadow cases and close value proximity of the text with that of foreground or background as well with low contrast.

Key words: Binarization, hand-held camera, local approach, thresholding, uneven illumination

INTRODUCTION

Image processing and analysis has become one of interesting research areas, for it accords a computer with the capability to interpret numerous varieties of images. Processing, analysing and acquiring the information from images are steps of documents analysis methods to interpret the textual and shapes of images (Marinai, 2008). The typical fashion of procuring images from documents is by scanning them via any commodity scanners. However, the emerging popular trend of document image acquisition as of late is via the handheld cameras (Hannuksela *et al*., 2007; Zhang *et al*., 2013). The quality of images captured from these handheld cameras however, varies as they are highly provisional on the capturing circumstances.

Camera is the basic feature of modern mobile phones. It comes in great handy for capturing and converting traditional documents into digital format images, which renders it the most convenient tool for said purpose. Nonetheless, these images are captured by a variety of image capturing software set to different resolutions and light conditions. In automatic document analysis case, these variations may negatively affect the quality of the resulting image in a way that the shadow or bright light present contributes uneven illumination (Bukhari *et al*., 2009; Hannuksela *et al*., 2007). One way to remedy the aforementioned deformation is by means of binarization or also known as thresholding.

Thresholding is geared towards capturing the black and white color from an image in accordance to specific equations. It is regarded as a mandatory step in the pre-processing stage of document analysis process wherein several other techniques to extract information and eliminate noises namely, filtering, clearing and image preparation, are also incorporated. This method attempts

Corresponding Author: Musab Kasim Alqudah, Centre for Artificial Intelligence Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

to recognize two layers: the foreground layer which contains the text black pixels and background layer which contains the white or not-black pixels of document image. Therefore, in order to distinguish those two layers, all thresholding algorithms are divided into two approaches: local and global (AI and SA, 2015; Bataineh *et al*., 2011; Gatos *et al*., 2006; Slimane *et al*., 2013).

The global thresholding algorithm computes the whole image as one window or at once. It is speedy and yields very decent result in cases where documents are cleanly captured. However, this approach is not without challenges that is for the case of document images containing degraded, shadow, shadow, noise, complex structure, scanning errors and bad quality illumination levels. Conversely, the local approach partitions the image into small windows enabling the thresholding algorithm to be computed in each individual window. It excels in special cases where the global approach fails, albeit requires high computations and parameter tunings. Recently, however, many studies are exploring the utility of simple or compound thresholding methods to rectify the issues in the pre-processing step (AI and SA, 2015; Bataineh *et al*., 2011; Gatos *et al*., 2006; Kefali *et al*., 2010; Zhang *et al*., 2013).

The simple thresholding isolates useful pixels into black and background while noise pixels to white through only one equation. It is fast even when incorporated in local approach whilst easy in the tuning of the equation to suit various thresholding cases. In most normal cases, it can render excellent performance in separating text from document images, albeit not without some issues (Alkhatatneh *et al*., 2014; Armanfard *et al*., 2009; Kefali *et al*., 2010; Ntogas and Veintzas, 2008; Stathis *et al*., 2008). The compound thresholding, on the contrary, is a thresholding algorithm comprising of several sub-steps in which some simple thresholding might be included. The number of steps is very minimal and might be set as pre or post processing (Chou *et al*., 2010; Zhang and Yang, 2010). Unfortunately, recently those two types of thresholding fail to make the grade in cases of illumination, shadow, low contrast, low intensity variations or uneven brightness. This results in undesired outputs for some cases, especially in documents images where the gap between intensity or contrast variations is high.

The crux of this study is to propose a novel simple thresholding algorithm that specifically addresses the issues hitherto unresolved by previous thresholding methods to cope with hand-held camera captured images. The algorithm will be appraised visually and analytically through a series of experiments which is based on self-collected captured images and benchmark images from the Document Image Binarization Contest (DIBCO).

Motivation: Now a days, transferring and receiving captures images from hand-held cameras is effortless (Doermann *et al*., 2003) inasmuch as the digital cameras are portable, handy, easy to use, compact, offer high speed connection and come with several image editing software to modify, store and share pictures directly (Hannuksela *et al*., 2007; Kasar and Ramakrishnan, 2007). Henceforth, hand-held cameras gain momentous traction as apt devices for acquiring images of documents within constrained and uncontrolled environments. This is appropriately corresponded with the upsurge of researches pertaining to document image capture by hand-held and portable cameras along with documents images analysis.

The use of hand-held cameras is indeed at an advantage as it surpasses traditional scanning in terms of time saving, high performance, portability and affordability in spite of the fact that they are small, light, easily connected and integrated with various networks. In context of documents image capturing, they are of vital convenience when the size of documents, newspapers or books renders the scanning process too cumbersome or even when it is outright un-scannable such as text typed on machines, vehicles, note-board, buildings or other objects around. This certainly attests to the increased utility and flexibility of hand-held cameras relative to traditional scanners (Kim *et al*., 2014; Zhang *et al*., 2013).

Even so, capturing document images through hand-held cameras is still confronted with severe encumbrances. The deformation in the resulting image can manifest in the form of perspective distortion, skew, blur, variance of illumination and low resolution as well as complex layout such as shadow and unwanted combination "text and objects” and background. Despite that it is often disregarded by users, the deformation should be reduced or, better yet, totally removed such that the readability of the texts or objects in the document images is not compromised. In light of that, the rectifying approach expounded by this study involves segmenting document objects from its background using binarization or thresholding in conforming to the necessity for a robust, simple and fast thresholding method.

State-of-the art: According to Kasar and Ramakrishnan (2007), there were a number of software for recognizing images in existence. However, they fell short in some instances where the texts are not in focus during the acquisition process, a phenomenon stemmed from the apathy of the bulk of lay users towards intricacies of photography. Consequently, the images produced may be marred by uneven illumination, non-uniform lighting, shadow, blur, low resolution, noise and perspective distortion. Therefore, an edge-based connected component approach was adopted to determine the threshold for each component individually on each channel of the color image. This, nonetheless resulted in poor performance of the method given that the texturing of background causes the edge components detection to be error-prone as it is subject to edges from the background as well. Opportunely, there are myriads of preceding studies which had delved into various thresholding techniques such as global and local approach with simple thresholding.

Pratikakis *et al.* (2011) posited that the inconsistent background intensity contributed profusely to bad non-uniform illumination or uneven lighting. It was purported that the use of single thresholding in pre-processing of document image is impractical and thusly causes said problems. Instead, an initial adaptive segmentation surface estimation and normalization were incorporated to boost the pre-processing task on the document images. However, this method was rendered useless by the copious amount of large images with varying intensity, contras and binarization in our dataset. The process may prematurely terminate at the segmentation step without yielding a satisfactory result.

Global thresholding method presented in Otsu was among those regarded the most important in extracting information from images. Otsu method was based on clustering of distributed pixels to extract the text from document images. This results reflected the excellent performance of the method in several binarization cases in addition to its speediness. Still, contrast, intensity, extraneous light and shadow or any uneven illumination states are among the factors, which Otsu method is heavily affected by, that would compromise the quality of the resulting image. All in all, the Otsu method was not the best way to extract text from document images captured from handheld camera. The Otsu equation is defined as follows:

$$\begin{gathered}\sigma^2_{Within}(T) = N_B(T)\sigma^2_B(T) + N_o(T)\sigma^2_o(T)\\ \sigma^2_{Between}(T) = N_B(T)\,N_o(T)\\ \left[\mu_B(T) - \mu_o(T)\right]^2\end{gathered} \quad (1)$$

$$\mu_o(T+1) = \frac{\mu_o(T)\,n_o(T) - n_T T}{n_o(T+1)} \quad (2)$$

Where:

$$N_B(T) = \sum_{t=0}^{T-1} P(i), N_o(T) = \sum_{i=T}^{n-1} P(i) \quad (3)$$

$\sigma^2_B(T)$ = The variance of the pixels in background (below threshold)
$\sigma^2_o(T)$ = The variance of the pixels in foreground (above threshold)

The mediocre result rendered by global thresholding methods is due to their treatment of document image as a whole despite that it is captured under poor lighting conditions such as uneven illumination, shine and shadows, or has accumulated and degraded text. Local and hybrid methods, alternatively, attempt to reduce the number of shades or other special cases in the document image, by means of windows or region techniques (Makridis and Papamarkos, 2010; Papamarkos, 2003). In one way or another, the main goal of these is to decrease the number of uneven illumination values in document image cases into only two values (black and white).

A local thresholding method, based on mean, standard deviation and fixed variable numbers was proposed by Sauvola and Pietikainen (2000). Although clear images are obtained through said method, the edge of characters was severely compromised as it deleted pixels in window edge that resembled the color of background images. Moreover, all the text having shine light in the document images are deleted too. Wolf and Jolion (2003), on the other hand, improved the thresholding by juxtaposing local contrast of local windows with the global contrast of whole image. In certain cases, it managed to produce the best result in text retrieval. Even so, issues such as the presence of blank area in document image, extraneous noises generation and tendency of text removal within area with additional light in document image carried significant weight on the performance of the method. The following thresholding describes the Sauvola and Wolf equations:

$$T_{sauvola} = m\left(1 - k\left(1 - \frac{S}{R}\right)\right) \quad (4)$$

$$T_{sauvola} = m\left(1 - k\left(1 - \frac{\sqrt{\sum_{t=1}^{n}(P_i - m)^2}}{R}\right)\right) \quad (5)$$

Where:
m = Local mean
s = Standard deviation

K = Fixed to 0.5 by the researchers
R = Fixed to 128 by the researchers
N = The number of pixels for slide windows
pi = Current pixel in slide windows

$$T_{wolf} = m\left(1 - k \times m + k \times M + k \times \frac{S}{R}(m - M)\right) \tag{6}$$

Where:
m = Window mean
s = Window standard deviation
K = Fixed to 0.5 by researchers
M = Minimum gray value of the image
R = Maximum gray-value standard deviation over all windows

Bataineh *et al.* (2011) propounded novel techniques based on local binarization which was inspired by Nick and Niblack equations (Niblack, 1986). The Niblack later is enhanced by Nick equation through the exploitation of the standard slide window. Nick method was specially designed for binarization of antiquated and mediocre quality printed documents. Local thresholding is a prerequisite in settling the degraded issues which in turn, exhibited good results in documents with small variations of contrast or low intensity to derive foregrounds from their backgrounds.

Consequently, substantial noises emerged in shadow or dark area of documents. Bataineh, instead exploits the areas of adaptive dynamic slide windows as method of enhancement and the result outperformed that of previous techniques in most cases. Still, it was influenced by uneven light within available gaps between contrast values. This results in noises in background area nearby text values and thereby creating black area in the edge between dark and light areas. The following thresholding describes the Nick and Bataineh equations:

$$T_{Niblack} = m + k \times S \tag{7}$$

$$T_{Niblack} = m + k\sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{n}(p_i - m)^2} \tag{8}$$

Where:
m = Local mean
s = Standard deviation
K = Fixed to -0.2 by the researchers
N = The number of pixels for slide windows
pi = Current pixel in slide windows

$$T_{Nick} = m + k \times \frac{\sqrt{\left(\sum\nolimits_{i=1}^{n} p_i - (m)^2\right)}}{R} \tag{9}$$

Where:
m = Global mean
K = Fixed to -0.1 and -0.2 is contingent on application defined by researchers
B = First thresholding rely on standard deviation
NP = The number of pixels for slide windows
pi = Current pixel in slide windows

$$T_{Bataineh} = m - \frac{m^2 * S}{(M + S) * (S + s)} \tag{10}$$

Where:
M = Global mean
m = Window mean
s = Standard deviation of slide window
S = Standard deviation for flexible slide “adaptive window”

Sauvola, Nick and Niblack were considered as good “simple” thresholding techniques where else, Wolf and Bataineh were considered as compound thresholding techniques of local binarization to extract text from images. Those methods had undergone myriads of laborious tasks which unequivocally failed in cases of uneven illumination, degraded, shadows, low contrast, smears and heavy noise. Some of the cases are shown in Fig. 1.

In line with the above-mentioned difficulties, researchers relied on advanced techniques that employed local thresholds of which the approximation was contingent on local intensity, contrast, brightness and spatial characteristic (Bataineh *et al*., 2011; Gatos *et al*., 2006; Gatos *et al*., 2009; Niblack, 1986; Sauvola and Pietikainen, 2000; Wolf and Jolion, 2003). Although, local techniques are almost tolerant to shadow, illumination changes and variance lighting, they were highly are critical and sensitive to degraded or non-uneven illuminations. Indeed, there were a wide-ranging different methods that incorporated simple and compound thresholding including local or global techniques but failed in obtaining good results for all kind of documents problems which is evidenced by the quality of the resulting image specifically when there are gaps between large amounts of values for intensity or contrast variations in documents images.

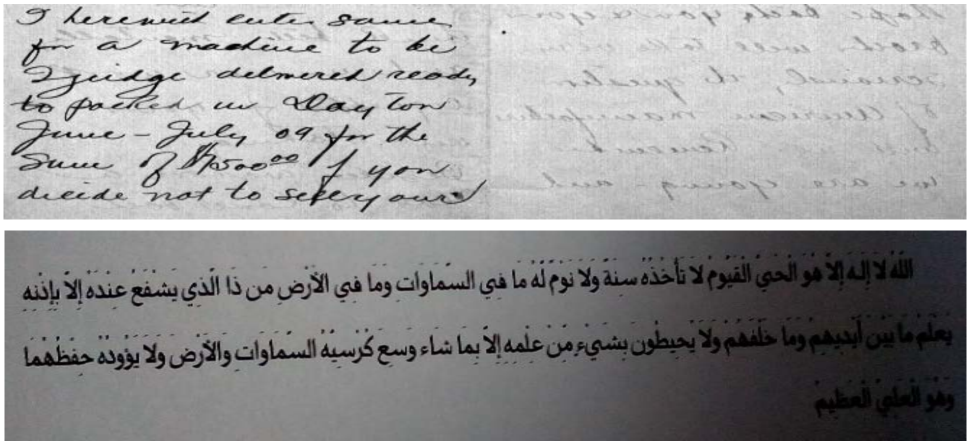

Fig. 1: Example of documents problem cases taken from DIBCO and our image dataset

MATERIALS AND METHODS

The materials of this study mainly use benchmark images from the Document Image Binarization Contest "DIBCO" benchmark dataset that is published from 2009 until 2013 (Gatos *et al*., 2009; Pratikakis *et al*., 2010, , 2011, 2012, 2013). Also, we captured another 5 images of Arabic documents with diacritics using a 5MP camera and added those additional images in this experiment. The images with their ground truths are available in our research center website at http://www.ftsm.ukm.my/cait/index.php/download/category/8-dataset.

The proposed method bypasses the use of global binarization because of its proven ineffectiveness for document images consisting of uneven illuminations or degraded problems. In its place, we use a simple binarization method, a fast and easy method to processing the images, which performs satisfactorily in many document image cases (Bataineh *et al*., 2011; Gatos *et al*., 2006; Sauvola and Pietikainen, 2000; Wolf and Jolion, 2003). Therefore, in this study, we propose the use of simple method based on local method to estimate a new path that can successfully address contemporary issues associated with capturing document images via hand-held cameras.

The proposed thresholding is inspired by the work of Niblack (1986) Wolf and Jolion (2003), Sauvola (2000), Khurshid *et al.* (2009), Niblack (1986), Sauvola and Pietikainen (2000) and Wolf and Jolion (2003) but without the use of fixed value of K (external factor). We also employ a SQRT function in the thresholding equation which is inspired from Otsu. With that, it implies that our method takes into account values of document image pixels only which are the mean and standard deviation of slide windows from the whole image. The method, known as Mosab thresholding, is defined as follows:

$$m = \sum_{i=0}^{n} \frac{x_i}{n} \tag{11}$$

$$S = \sqrt{\frac{\sum_{i=0}^{n} (x_i - m)^2}{n}} \tag{12}$$

$$T_{Msoab} = m \left(1 - \sqrt{\frac{m^2}{S^3}}\right) \tag{13}$$

Where:

m = Mean of local slide window
s = Standard deviation of slide window

The next pseudo code represents the steps of the proposed thresholding. We commence with the step to change the colored image to gray scale. Then, we segment the whole image into 40×40 slide windows. After that, for every slide window, we calculate the standard deviation and mean and threshold the slide window pixels into white or black based on Eq. 11. Finally, we store the new pixels that are obtained from the proposed thresholding into the image result. The pseudo code for our proposed method is given as follows:

Algorithm 1: Mosab's thresholding algorithm for segmenting a colored image into a binary image

```
Function Mosab thresholding (Color-Image)
Input  : A colored image
Output: A binary image
Initialize Gray-Image
Initialize n _ 40×40
Initialize Slide-Windows [n], sdv, mean, t
```

```
Gray-Image = Function RGB to Gray (Color-Image)
Slide-Windows _ Segment (Gray-Image)
for i _ to n do
sdv _ Standard Deviation (slide-Windows [i])
mean _ Average (Slide-Windows[i])
t _ m* (1 – Sqrt (mean^2 / (sdv ^3)))
Binary-Image _ Threshold (t, Slide-Windows, i)
End
```

RESULTS

Experiment: In this study, in order to evaluate the proposed thresholding method, we conduct two experiments to encompass visual and analytical aspects respectively.

Visual experiment: In visual experiment, four images are procured from a hand-held device camera. The thresholded images will be evaluated based on quality seen subjectively by our naked eyes. While capturing the images, we introduced several uneven illumination challenges such as shadow, additional light and reflection of glossy and transparent studys which are performed in both day and night condition. Subsequently another two images from DIBCO are added on which we label them by alphabets a to h. For the thresholding methods, we assign them from a to h, indicating original image, Otsu, Niblack, Sauvola, Wolf, Nick, Bataineh and Mosab respectively.

In this experiment, we use the 40×40 slide window size for all methods except the Bataineh and Otsu. In the Bataineh method, the slide window size is automatically determined by the algorithm, where else, the Otsu is based on global approach.

Analytical experiment: The analytical experiment, on the other hand, utilizes the images predominantly from DIBCO. The datas alllows for the use of ground truth information of the images for evaluation measurement and statistical measurement. We select 5 handwritten and 8 printed document images from DIBCO encompassing different challenges of binarization cases. Also, we captured another 5 images of Arabic documents with diacritics using a 5MP camera and added those additional images in this experiment. The images with their ground truths are available in the research center website at http://www.ftsm.ukm.my/cait/index.php/download/category/8-dataset.

The evaluation measurement in this experiment is subject to analytical equations mentioned in Gatos *et al*. (2009). They are prevalently used for the comparison between different binarization algorithms. It thereby follows that we used the F-measure, F-error and Peak Signal-to-Noise Ratio (PSNR) for the evaluation of the proposed thresholding (Makridis and Papamarkos, 2010; Pratikakis *et al*., 2010, 2011, 2012, 2013).

The following equations describe the F-Measure, where TP, FP and FN represent the true-positive (total number of matched foreground pixels), false-positive (total number of misclassified foreground pixels in binarization result as compared to ground-truth) and false-negative (total number of misclassified background pixels in binarization result as compared to ground-truth) values, respectively.

$$\text{F}-\text{Measure} = \frac{2\times\text{recall}\times\text{Precision}}{\text{recall}+\text{Precision}} \tag{14}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{15}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{16}$$

The quality of original image from noises is appraised through PSNR wherein the error percentage of result images to ground truth images is measured. Higher value of PSNR implies higher similarity between the two images, which is considered as a difference between foreground and background equals to MAX values of best ground truth image. The MAX is the maximum possible pixel value of the image. The following equations describe the PSNR equations:

$$\text{PSNR} = 10\log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right) \tag{17}$$

$$\text{MSE} = \frac{\sum_{x=1}^{M}\sum_{Y=1}^{N}\left(1(x,y)-1(x,y)\right)^2}{\sum_{x=1}^{M}\sum_{Y=1}^{N}1(x,y)} \tag{18}$$

Where:

MSE = Mean Squared Error
I = The ground truth Image
I = Image result acquired from the binarization method

Additionally, F-Error is used to find the error percentage of noises discerned in document images results, which is contingent on false positive and false negative of image binarization result to ground truth results. It is defined as the following equation:

$$\text{F}-\text{Error} = \frac{\text{FP}+\text{FN}}{\sum_{x=1}^{M}\sum_{Y=1}^{N}1(x,y)} \tag{19}$$

Table 1: The F-measure, Error, PSNR and F-Error results of experiment binarization methods for DIBCO Benchmark

Variables	F-Measure	Error	PSNR	F-Error
Bataineh	0.637	0.074	12.838	1.181
Mosab	0.858	0.015	18.275	0.253
Niblack	0.287	0.280	05.534	4.751
Nick	0.742	0.041	15.007	0.650
Otsu	0.206	0.432	03.736	7.329
Sauvola	0.844	0.018	17.692	0.299
Wolf	0.717	0.036	14.457	0.619

Table 2: The F-measure, Error, PSNR and F-Error results of experiment binarization methods for Arabic Images

Variables	F-Measure	Error	PSNR	F-Error
Bataineh	0.848	0.023	16.995	0.364
Mosab	0.859	0.023	17.803	0.257
Niblack	0.330	0.283	05.559	5.245
Nick	0.792	0.033	15.457	0.504
Otsu	0.850	0.027	17.587	0.379
Sauvola	0.859	0.022	17.673	0.304
Wolf	0.721	0.043	13.853	0.761

Table 1 presents the detailed results based on evaluation measurements, F-measure, PSNR, F-Error, Error, Recall and Precision. The proposed method as well as Sauvola's attain the best results in extracting text from images based on F-measure. This implies that our proposed method manages to gain the highest truth position pixels in image results as well as the most excellent execution of noise cleaning from images based on F-Error.

According to Table 2, our method produces the best result in all measurements. For example, our method acquires nearly perfect accomplishment, which is 94.1% in precision, 17.803 in PSNR and 0.257 in F-Error. The second best method is the Sauvola in most of the measurements. On the other hand, the worst results come from the Otsu (global approach) and Niblack (local approach).

DISCUSSION

This study comprises of two parts, namely visual and analytical experiments for comparing the results of conducted methods on document image.

Visual experiment: The visual experimental as depicted in Fig. 2, demonstrates how the proposed thresholding outperforms the other techniques. The quality of thresholded images from the proposed method is outstanding on uneven illumination as with the other binarization cases. Nonetheless, it is also discovered that the previous thresholding fall short in cases of low quality, low contrast and uneven illuminations such as shadow and extraneous light on transparent and glossy papers.

Indeed, the global approach is not affected by uneven illumination but it converts a darker area into black color in its entirety. Conversely, the previous local methods performs satisfactorily in extracting text from images, putting the research of enhancing and developing new thresholding based on local method, in the right direction with regards to. However, this is not the case when it comes to images that are rife with noises (Fig. 3-5) and extraneous light (Fig. 2 and 3). Still, in case of noise and dirt areas removal, the performance of our proposed method significantly surpasses that of the others. The next section probes into capability of the proposed method in statistically coping with the uneven illumination document image.

Analytical experiment: The analytical experiment results also attest to superior performance of the proposed method relative to the other methods as depicted in Fig. 3 and 4. The average of F-Measure for DIBCO images is 85.9% for Mosab, 85% for Otsu, 33% for Niblack, 85.9% for Sauvola, 72% for Wolf, 79.1% for Nick and 84.8% for Bataineh. In addition, the average of F-Measure for Arabic images is 85.8% for Mosab, 20.6% for Otsu, 28.7% for Niblack, 84.4% for Sauvola, 71.7% for Wolf, 74.2% for Nick and 63.7% for Bataineh.

The PSNR result illustrates the measure by which the proposed method achieved the best result relative to the previous methods. The results of DIBCO PSNR are sorted in ascension according to performance of the applied techniques, that is as follow: Mosab, Sauvola, Otsu, Bataineh, Nick, Wolf and Niblack 17.8, 17.7, 17.6, 17.0, 15.5, 13.9 and 5.6%. Whereas, the results of Arabic PSNR are in the following order: Mosab, Sauvola, Nick, Wolf, Bataineh, Niblack, Otsu, 18.3, 17.7, 15, 14.5, 12.8, 5.5 and 3.7%. The Otsu technique is greatly at disadvantage in captured images included illumination cases.

Generally, the global approach represented by Otsu's thresholding, succumbs seriously to uneven illumination. Niblack method as a local approach, is able to clear and extract the text, yet creates black pixel noises as its background. Even though Sauvola method manages to generate the second best results, extraneous light causes it to eliminate part of text from bright area. Notheless, noise generation is more minimal than other previous methods. Wolf evaded extraneous light challenges by using global standard deviation in thresholding equation. However, many problems in large images and large of noises amount are yet unresolved. Nick and Bataineh techniques which are derived from Niblack, accomplish textual extraction in low quality images, still noises in dark and low contrast part of image still appear. The visual and analytical experiments prove that our proposed simple thresholding is able to deal with the uneven illumination cases as well as other binarization challenge cases.

Fig. 2: Binarization results of uneven illumination document images capturing from hand phone and from DIBCO benchmark dataset

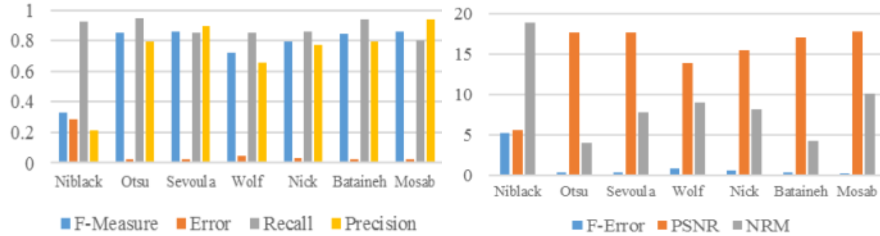

Fig. 3: The F-measure, Error, Recall, Precision, F-Error, PSNR and NRM of experiment binaraization methods for Thirteen DIBCO Benchmark

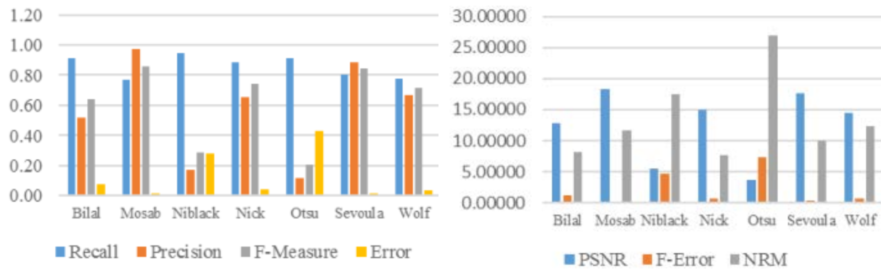

Fig. 4: The Recall, Precision, F-measure, Error, PSNR, F-Error and NRM of experiment binaraization methods for five Arabic Images

CONCLUSION

In this study, a novel simple binarization technique has been presented based on local binarization for captured document images via handheld camera. The proposed method segments the whole image into individual slide windows. Each window uses mean and standard deviation of image pixels without depending on fixed number K utilized by the previous methods. It had accomplished good results in all measurements for DIBCO and self-collected images with uneven illuminations and various binarization challenge cases. It succeeds in removing noises and uneven illumination, which is unsupportable by the other techniques. Since, the method is a simple thresholding, it is easy to be implemented in pre-processing step.

ACKNOWLEDGEMENT

The researchers would like to thank the International University of Malaysia (UKM) and Faculty of Information Science and Technology (FTSM). This work was supported in part by research grant number is FRGS/1/2014/ICT07/UKM/01/1.

REFERENCES

AI, A.K. and P. SA, 2015. Compound binarization for degraded document images. ARPN. J. Eng. Appl. Sci., 10: 594-599.

Armanfard, N., M. Valizadeh, M. Komeili and E. Kabir, 2009. Document image binarization by using texture-edge descriptor. Proceedings of the 14th CSI International Conference on Computer CSICC, October 20-21, 2009, IEEE, New York, USA., ISBN:978-1-4244-4261-4, pp: 134-139.

Bataineh, B., S.N.H.S. Abdullah and K. Omar, 2011. An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. Pattern Recognit. Lett., 32: 1805-1813.

Bukhari, S.S., F. Shafait and T.M. Breuel, 2009. Foreground-background regions guided binarization of camera-captured document images. Proceedings of the International Workshop on Camera Based Document Analysis and Recognition, July 25, 2009, Technical University of Kaiserslautern, Kaiserslautern, Germany, pp: 18-25.

Chou, C.H., W.H. Lin and F. Chang, 2010. A binarization method with learning-built rules for document images produced by cameras. Pattern Recognit., 43: 1518-1530.

Doermann, D., J. Liang and H. Li, 2003. Progress in camera-based document image analysis. Proceedings of the Seventh International Conference on Document Analysis and Recognition, August 6-6, 2003, IEEE, New York, USA., ISBN:0-7695-1960-1, pp: 606-616.

Gatos, B., I. Pratikakis and S.J. Perantonis, 2006. Adaptive degraded document image binarization. Pattern Recognition, 39: 317-327.

Gatos, B., K. Ntirogiannis and I. Pratikakis, 2009. ICDAR 2009 Document Image Binarization Contest (DIBCO 2009). Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR), July 26-29, 2009, IEEE, Barcelona, Spain, pp: 1375-1382.

Hannuksela, J., P. Sangi, J. Heikkila, X. Liu and D. Doermann, 2007. Document image mosaicing with mobile phones. Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP), September 10-14, 2007, IEEE, New York, USA., ISBN:978-0-7695-2877-9, pp: 575-582.

Kasar, T. and A.G. Ramakrishnan, 2007. Specialized text binarization technique for camera based document images. WISP., 2007: 28-31.

Kefali, A., T. Sari and M. Sellami, 2010. Evaluation of several binarization techniques for old Arabic documents images. Proceedings of the First International Symposium on Modeling and Implementing Complex Systems (MISC), May 30-31, 2010, Badji Mokhtar University, Annaba, Algeria, pp: 88-99.

Kim, W.T., H.S. Jeong, G.H. Lee and T.H. Kim, 2014. A high-speed and low-complexity lens distortion correction processor for wide-angle cameras. Proceedings of the 2014 19th Conference on Asia and South Pacific Design Automation (ASP-DAC), January 20-23, 2014, IEEE, New York, USA., ISBN:978-1-4799-2816-3, pp: 39-40.

Makridis, M. and N. Papamarkos, 2010. An adaptive layer-based local binarization technique for degraded documents. Int. J. Pattern Recognit. Artif. Intel., 24: 245-279.

Marinai, S., 2008. Introduction to Document Analysis and Recognition. In: Machine Learning in Document Analysis and Recognition, Simone, M. and H. Fujisawa (Eds.). Springer, Berlin, Gemany, ISBN:978-3-540-76279-9, pp: 1.

Niblack, W., 1986. An Introduction to Image Processing. Prentice-Hall, Englewood Cliffs, New Jersey.

Ntogas, N. and D. Veintzas, 2008. A binarisation algorithm for historical manuscripts. Proceedings of the 12th WSEAS international conference on Communications. July 23-25, 2008, WSEAS, Heraklion, Greece, ISBN:978-960-6766-84-8, pp: 41-51.

Papamarkos, N., 2003. Document gray-scale reduction using a neuro-fuzzy technique. Int. J. Pattern Recognit. Artif. Intell., 17: 505-527.

Pratikakis, I., B. Gatos and K. Ntirogiannis, 2010. H-DIBCO 2010-handwritten document image binarization competition. Proceedings of the 2010 International Conference on Frontiers in Handwriting Recognition (ICFHR), November 16-18, 2010, IEEE, New York, USA., ISBN: 978-1-4244-8353-2, pp: 727-732.

Pratikakis, I., B. Gatos and K. Ntirogiannis, 2011. DIBCO 2011: Document image binarisation contest. Int. J. Doc. Anal. Recognit., 14: 35-44.

Pratikakis, I., B. Gatos and K. Ntirogiannis, 2012. ICFHR 2012 competition on handwritten document image binarization (H-DIBCO 2012). Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition (ICFHR), September 18-20, 2010, IEEE, New York, USA., ISBN: 978-1-4673-2262-1, pp: 817-822.

Pratikakis, I., B. Gatos and K. Ntirogiannis, 2013. ICDAR 2013 document image binarization contest (DIBCO 2013). Proceedings of the 2013 12th International Conference on Document Analysis and Recognition, August 25-28, 2013, IEEE, New York, USA., ISBN: 978-0-7695-4999-6, pp: 1471-1476.

Sauvola, J. and M. Pietikainen, 2000. Adaptive document image binarization. Pattern Recognit., 33: 225-236.

Slimane, F., S. Kanoun, J. Hennebert, A.M. Alimi and R. Ingold, 2013. A study on font-family and font-size recognition applied to Arabic word images at ultra-low resolution. Pattern Recognit. Lett., 34: 209-218.

Stathis, P., E. Kavallieratou and N. Papamarkos, 2008. An evaluation technique for binarization algorithms. J. Universal Comput. Sci., 14: 3011-3030.

Wolf, C. and J.M. Jolion, 2004. Extraction and recognition of artificial text in multimedia documents. Formal Pattern Anal. Appl., 6: 309-326.

Zhang, C.Y. and J.Y. Yang, 2010. Binarization of document images with complex background. Proceedings of the 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), September 23-25, 2010, IEEE, New York, USA., ISBN: 978-1-4244-3708-5, pp: 1-4.

Zhang, J., T. Zheng, X. Fu, W. Wang and N. Huo, 2013. Image thresholding method based on uneven lighting. Proceedings of the 32nd Conference on Chinese Control Conference (CCC), July 26-28, 2013, IEEE, Xi'an, China, ISBN: 978-9-8815-6383-5, pp: 4549-4552.