Computer Engineering and Applications Vol. 2, No. 1, March 2013

An Analysis of Adaptive Approach for Document Binarization

Reza Firsandaya Malik, Saparudin, Intan Septyliana
Faculty of Computer Science, University of Sriwijaya
Email: reza.firsandaya@gmail.com

ABSTRAKSI

Binerisasi merupakan tahap awal pada analisis citra dokumen untuk membedakan area teks dari area latar belakang. Pemilihan teknik binerisasi sangat penting untuk memperoleh semua informasi teks yang dibutuhkan terutama pada citra dokumen yang mengalami degradasi. Makalah ini membahas mengenai pendekatan adaptif untuk binerisasi dokumen terdegradasi dengan menggunakan metode Gatos. Metode Gatos melakukan pra-pemrosesan, estimasi foreground dengan menggunakan metode Sauvola, background, upsampling, thresholding akhir dan tahap pasca-pemrosesan. Pada makalah ini, metode Sauvola merupakan thresholding akhir dari citra hasil filter Wiener dan citra asal, dan menghitung nilai F-Measure citra biner hasil keduanya. Dengan menggunakan nilai konstanta yang optimal pada nilai k, jendela lokal n, K_{sw} dan K_{sw1}, metode Gatos menghasilkan citra biner yang lebih baik daripada citra biner hasil metode Sauvola berdasarkan nilai F-Measure. Metode Sauvola menghasilkan rata-rata nilai F= 84,62%, metode Sauvola dengan filter Wiener menghasilkan rata-rata nilai F=99,06% dan metode Gatos menghasilkan rata-rata nilai F= 99,43%.

Kata kunci : Citra Dokumen Terdegradasi, Pendekatan Adaptif untuk Binerisasi, metode Gatos, dan metode Sauvola

ABSTRACT

Binarization is an initial step in document image analysis for differentiate text area from background. Determination of binarization technique is really important to retrieve all the text information especially from degraded document image. This paper explains about adaptive binarization using Gatos's method. Gatos's method is doing preprocessing, foreground estimation using Sauvola's method, background estimation, upsampling, final thresholding and postprocessing. In this paper, Sauvola's method is final thresholding from Wiener filter image result and source image, and count F-Measure from both of these binary image results. By using optimum constant value on k value, n local window, K_{sw} and $K_{sw1,}$ Gatos's method can produced binary image better than Sauvola's method based on F-Measure value. Sauvola's method produces average value F=84,62%, Sauvola's method with Wiener filter produces average value F=99.06% and Gatos's method produces average value F=99,43%.

Keywords*:* Degraded Document Image, Adaptive Approcah for Binarization, Gatos's Method and Sauvola's Method

1. Introduction

Binarization is done globally and locally [1]. As global binarization, single threshold value was counted for classify image pixel into object class or background. Meanwhile, as local, single threshold value retrieved from image local area information. Many available degraded document image binarization, but only few could improve the quality of degraded document image.

Method used by Otsu [2] is using global thresholding technique which is retrieved the threshold value based on histogram. Otsu's method is one of the best global thresholding and work well in clear scanned image, but produces unsatisfy output for poor quality image with low contrast and non-uniform illumination. Sauvola and Pietikainen [3] is doing classification to background, image and text from document image's local content.

The method which is used by Sauvola and Pietikainen solves the noise problem in background, but the character in the binarized image becomes thinner. Gatos, Pratikakis and Perantonis [4] use few steps in adaptive binarization, such as image preprocessing procedure using Wiener filter, foreground region estimation with Sauvola method, background surface estimation by doing interpolation to background intensity, thresholding by combine background surface estimated with upsampled image, then image postprocessing step for improves the quality of text region. Because according to He research, that *n* and *k* value has an effect on foreground estimation, so This research is doing observation on Gatos and Sauvola method by observes the parameter local window (*n)* in preprocessing step, *k* in foreground estimation, also parameter local window (*n)* in preprocessing step, k_{sw} dan k_{sw1} in postprocessing step.

2. Research Method

2.1 Image Sampel

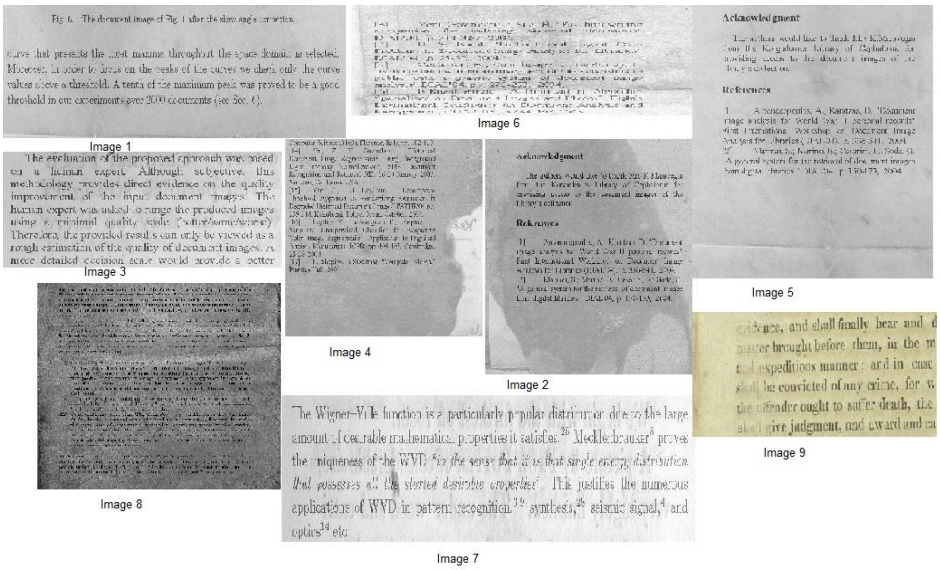


Figure 1. Image Sampel

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Figure 1 shows data that are used in this research. The data contain non-uniform illumination background, low contrast, smear, shadows, stain and other noises.

2.2 Preprocessing

Preprocessing step for poor quality and degraded image is needed for eliminate the noise, background texture smoothing, and contrast enhancement between background and text area. The using of Wiener filter has been proved efficient in image preprocessing. Adaptive filter is applied for refines the degraded image from random noise, so the noise in the background decreased. The grayscale image $I_s(x,y)$ is transformed into image result $I\ (x,y)$ with equation 1:

$$I(x,y) = \mu + (\sigma^2 - v^2)(I_s(x,y) - \mu)/\sigma^2 \qquad (1)$$

μ was local mean, σ^2 was local variance and v^2 was average from all the estimated local variance in neighborhood 3x3 around each pixel.

2.3 Foreground Estimation

Threshold value is estimated in the dynamics of the standard deviation (R) and constant (k). m coefficient decreases the threshold value in background area so it can remove the noise in image. Equation 2 is a equation used by Sauvola and Pietikakinen for counts threshold value for foreground estimation. In this research, we do evaluation in k from 0.04 to 0.09 in local window n 3, 5 and 7.

$$\mathrm{T(x,y)=m(x,y).}\left[1+\mathrm{k.}\left(\frac{\mathrm{s(x,y)}}{\mathrm{R}}-1\right)\right] \qquad (2)$$

2.3 Background Estimation

Background intensity is estimated by using the neighborhood pixels and produces the background surface $B(x,y)$ [5]. Interpolation is used for determine intensity pixel based on gray value from neighbor pixel which is corresponded with the original pixel [6]. The equation III is used for background estimation. Window size 6x6 is used in background surface estimation [7].

$$\mathrm{B(x,y)}\begin{cases} \mathrm{I(x,y)} & \text{if } \mathrm{S(x,y)=0} \\ \dfrac{\sum_{ix=x-dx}^{x+dx}\sum_{i_y=y-dy}^{y+dy}\left(\mathrm{I(ix,iy)}\left(1-\mathrm{S(ix,iy)}\right)\right)}{\sum_{ix=x-dx}^{x+dx}\sum_{iy=y-dy}^{y+dy}\left(1-\mathrm{S(ix,iy)}\right)} & \text{if } \mathrm{S(x,y)=1} \end{cases} \qquad (3)$$

2.4 Upsampling

Bicubic interpolation gives the good result because more pixels are used for consideration, so it produces more accurate effect to upsampled image [8]. The upsampled image with bicubic interpolation is more smooth and sharp [9]. It estimates image pixel intensity from 16 nearest pixels in 4x4 neighborhood on original image. Final binarized image Intensity $T(x',y')$ after uses upsampled image $I_U(x',y')$ are retrieved from equation 4:

$$T\ (x',y') = \begin{cases} 1 \text{ jika } B(x,y)-I_U(x',y')>d(B(x,y)) \\ 0 \qquad\qquad\qquad \text{for other pixel} \end{cases} \tag{4}$$

T(*x',y'*) is binarized image with M = 2 from the original image, $x = (int)x'/M$, $y = (int)y'/M$, $a = (x'/M)-x$, $b = (y'/M)-y$ dan $I_u(x',y')$ are estimated with equation V:

$$I_u(x',y') = -b(1-b)^2\ F(x',y-1) + (1-2b^2+b^3)F(x',y) + b(1+b-b^2)F(x',y+1) - b^2(1-b) \\ F(x',y+2)+b(1+b-b^2)F(x',y+1) - b^2(1-b)F(x',y+2) \tag{5}$$

F() is estimated with equation VI:

$$F(x',m) = -a(1-a)^2I(x-1,m) + (1-2a^2+a^3)I(x,m) \\ + a(1+a-a^2)I(x+1,m) - a^2(1-a)I(x+2,m) \tag{6}$$

2.5 Final Thresholding

Local thresholding uses variance of threshold value and is estimated based on local content from image. Local thresholding has better performance to noise compared to global threhsolding. For finding the threshold value (*d*), we use equation 9. The average distance between foreground and background (δ) and the average of background value (*b*) are needed and estimated using equation 7, 8:

$$\delta = \frac{\sum_x \sum_y (B(x,y)-I(x,y)}{\sum_x \sum_y S(x,y)} \tag{7}$$

$$b = \frac{\sum_x \sum_y (B(x,y)-I(x,y)}{\sum_x \sum_y (1-S(x,y))} \tag{8}$$

$$d\big(B(x,y)\big)=q\delta\left(\frac{(1-P_2)}{1+\exp(\frac{-4B(x,y)}{b(1-P_1)}+\frac{2(1+P_1)}{(1-P_1)})}+P_2\right) \tag{9}$$

B(x,y) is estimated background, *S(x,y)* is estimated foreground and *I(x,y)* is original image which has been done preprocessing use Wiener filter. *q* is weight parameter 0.6 and P_1 also P_2 is determined 0.5 dan 0.8. Final binarized image intensity retrieved with equation 10 :

$$T\ (x,y) = \begin{cases} 1 \text{ jika } B(x,y) - I(x,y) > d(B(x,y)) \\ 0 \qquad\qquad\qquad \text{for other pixels} \end{cases} \tag{10}$$

2.6 Postprocessing

Image postprocessing is done in the last step for produces good binarized image by eliminates noise, improves the quality of text region and preserves the connectivity of characters. Shrink, Swell and Extension of Swell is used in postprocessing.

2.6.1 Shrink Filter

Shrink filter is used for remove the noise in background. Each foreground pixel is scanned. If the number of background pixels in window *n x n* which

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

has foreground pixel as the central pixel (P_{sh}) more than K_{sh} then that central pixel changed into background pixel. Local window used in this research was 5 or 7 and $K_{sh} = 0.9n^2$

2.6.2 Swell Filter

Swell filter is used for fills the gaps or hole in character from Shrink image result. If the number of foreground pixel in window $n \times n$ which has background pixel (x,y) as the central pixel (P_{sw}), and x_a, y_a are average coordinate value of all foreground pixel coordinate in window $n \times n$. If $P_{sw} > K_{sw}$ and $|x-x_a| < dx$ and $|y-y_a| < dy$, the central pixel transformed into foreground pixel. K_{sw} used in this research are 0.8 to 1 and $dx = dy = 0.25n$ also local window (n) 5 and 7.

2.6.3 Extension Swell Filter

The extension swell filter is used to improve the quality of character. Every background pixel is scanned. If the number of foreground pixel in window $n \times n$, which has background pixel as the central pixel (P_{sw1}) and the value is more than K_{sw1} then that central pixel transformed into foreground pixel. K_{sw1} used in this research are 0.8 to 1 and local window (n) 5 and 7.

2.7 Evaluation Measure

F-Measure is used for measure performance of final binarized image [10], it is using equation 11 where TP (True Positive) is total number of matched foreground pixels, FP (False Positive) is total number of misclassified foreground pixels in binarization result as compared to ground-truth and FN (False Negative) is total number of misclassified background pixels in binarization result as compared to ground-truth.

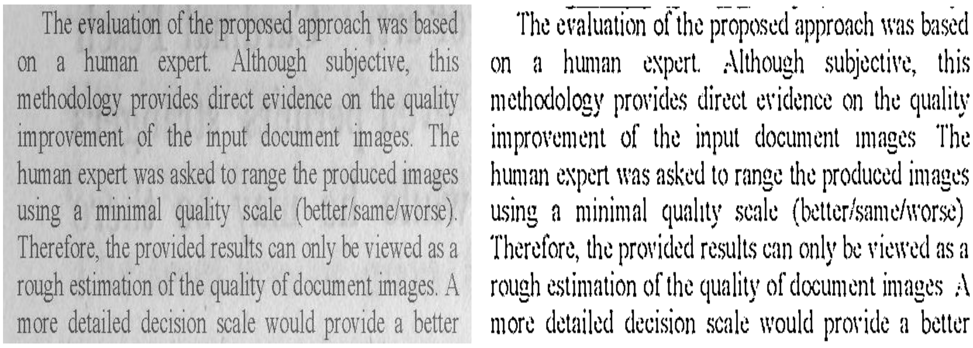
$$\text{Precision (P)} = \frac{TP}{(TP+FP)} \tag{11}$$

$$\text{Recall (R)} = \frac{TP}{(TP+FN)} \tag{12}$$

$$F = \frac{2.P.R}{P+R}.100\% \tag{13}$$

3 Results and Analysis

Figure 2 (b) shows the binarization result from Figure 2 (a). The image sampel uses k =0.09 in local window (n) = 7 in foreground estimation, meanwhile for postprocessing, this research uses k_{sw} = 0,8, k_{sw1} = 0,8 . those k and n value is used because it reaches the highest percentage of F-Measure 100%. For k_{sw}, k_{sw1} in n=7 also are used because they have percentage of F-Measure 100%.


The evaluation of the proposed approach was based on a human expert. Although subjective, this methodology provides direct evidence on the quality improvement of the input document images. The human expert was asked to range the produced images using a minimal quality scale (better/same/worse). Therefore, the provided results can only be viewed as a rough estimation of the quality of document images. A more detailed decision scale would provide a better

The evaluation of the proposed approach was based on a human expert. Although subjective, this methodology provides direct evidence on the quality improvement of the input document images. The human expert was asked to range the produced images using a minimal quality scale (better/same/worse). Therefore, the provided results can only be viewed as a rough estimation of the quality of document images. A more detailed decision scale would provide a better

Figure 2. (a) Image sampel 3.bmp (b) Binarized Image

Here is percentage chart of *k* evaluation on image sampel in local window (*n*) = 3 shows in Figure. 3.

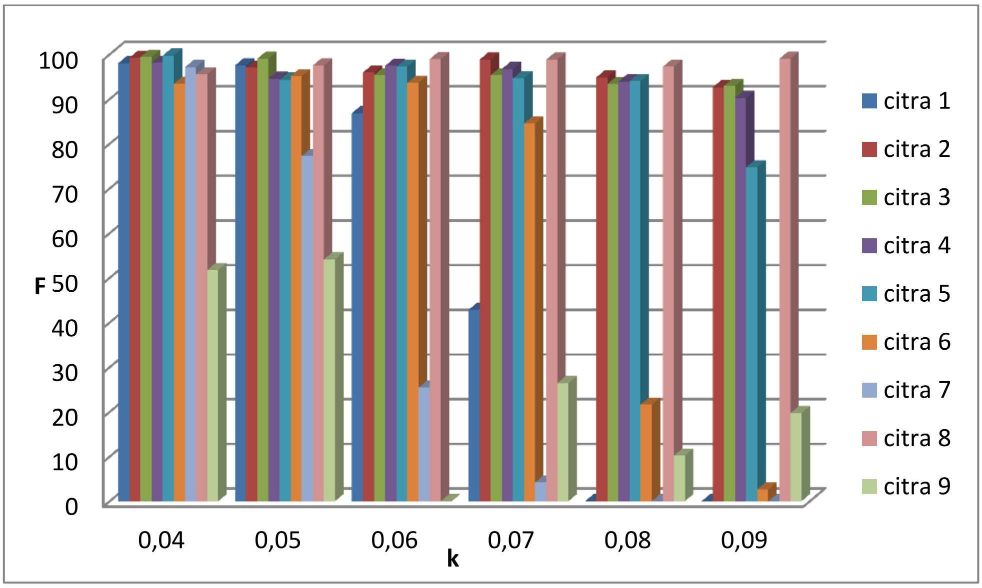


Figure 3. Percentage chart of *k* evaluation on image sampel in local window (*n*) = 3

Each image uses k=0.04 until 0.09 in local window = 3. In image 1, it shows that foreground estimation using k=0.04 obtain a good result, also for image 2,3,4,5, and 7. While for image 6 and 9, it is better to use k=0.05, and for image 8, it is better to use k=0.09.

Here is the percentage chart of *k* evaluation on image sampel in local window (*n*) = 5 shows in Figure 4.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

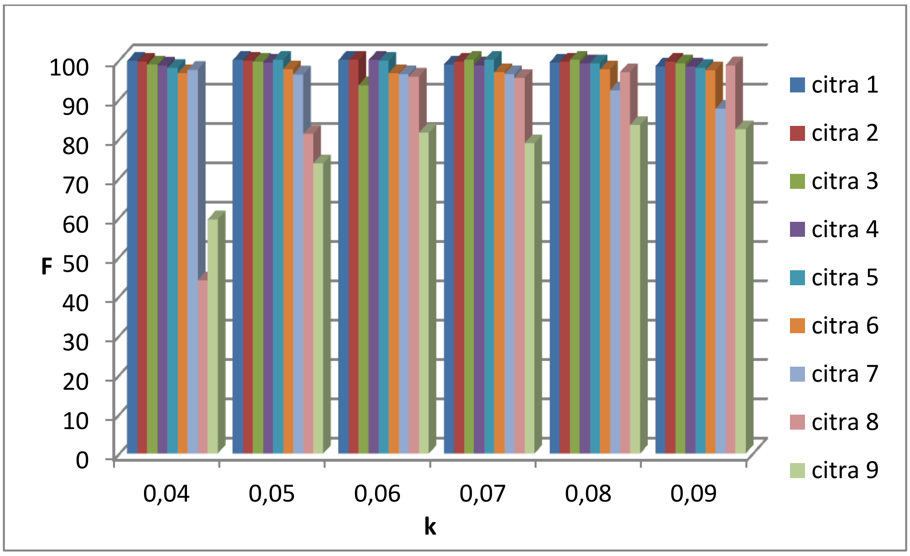


Figure 4. Percentage chart of *k* evaluation on image sampel in local window (*n*) = 5

Each image uses k=0.04 until 0.09 in local window = 5. In image 1, it shows that foreground estimation using k=0.06 obtain a good result, also for image 2 and 4. While for image 5 and 6, it is better to use k=0.05, for image 3, it is better to use k=0.07, for image 7, it is better to use k=0.04, for image 8, it is better to use k=0.09 and for image 9, it is better to use k=0.08.

Here is the percentage chart of *k* evaluation on image sampel in local window (*n*) = 7 shows in Figure 5.

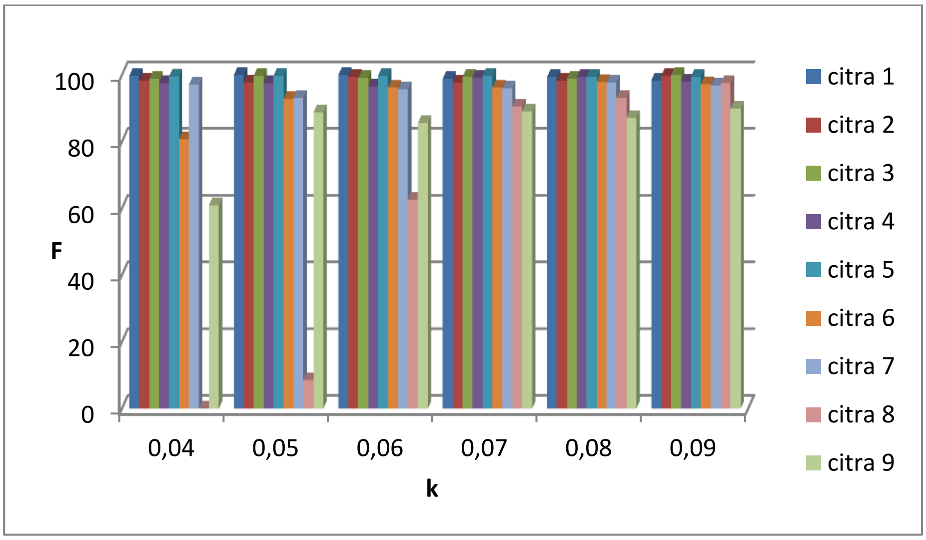


Figure 5. Percentage chart of *k* evaluation on image sampel in local window (*n*) = 7

Each image uses k=0.04 until 0.09 in local window = 7. In image 1, it shows that foreground estimation using k=0.05 obtain a good result, also for image 5. While for image 2,3,8 and 9, it is better to use k=0.09, for image 4 and 7, it is better to use k=0.08, for image 5, it is better to use k=0.05, and for image 6, it is better to use k=0.04.

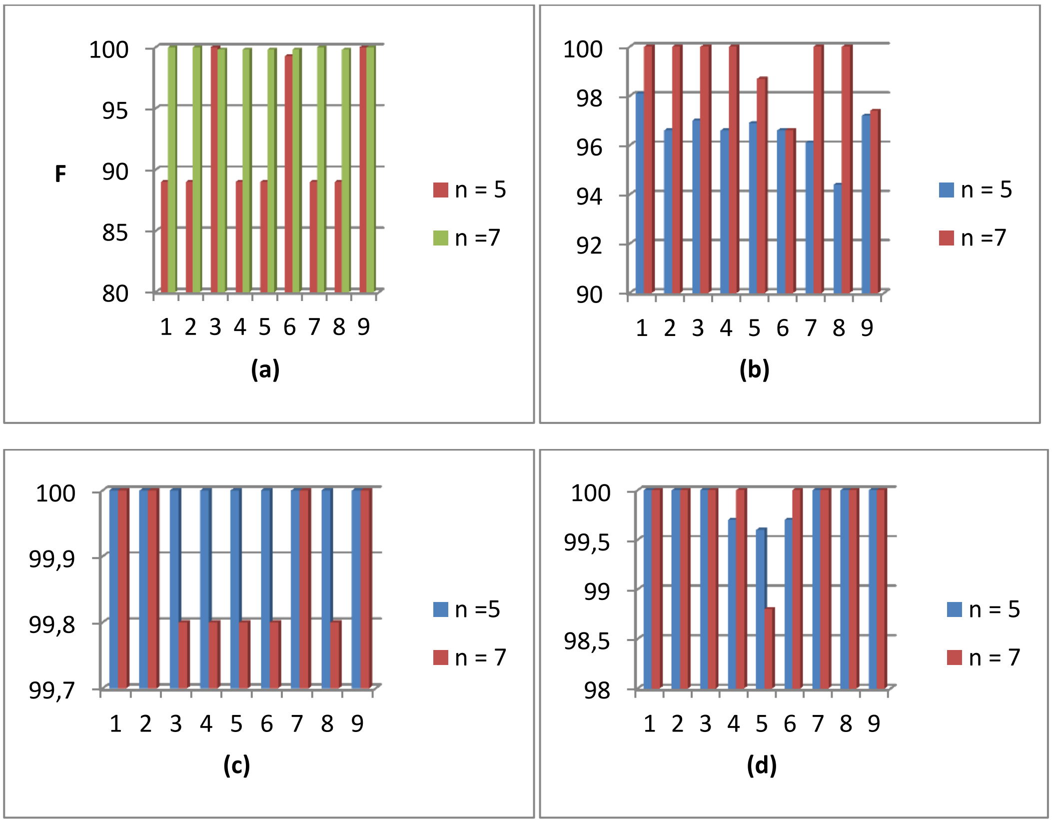


Figure 6. Percentage chart of k_{sw} and k_{sw1} in (a) image 1, (b) image 2, (c) image 3 and (d) image 4

Figure 6 (a), all combination k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 1, using k=0.05 and n=5, meanwhile only combination 3,6 and 9 in local window 5 give best result. Figure 6 (b), all combination k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 2, using k=0.06 and n=5. Figure 6 (c), most of combination k_{sw} and k_{sw1} in local window 5 give best result to final binarized image sampel 3, using k=0.09 and n=7. Figure 6 (d), most of combination k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 4, using k=0.06 and n=5.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

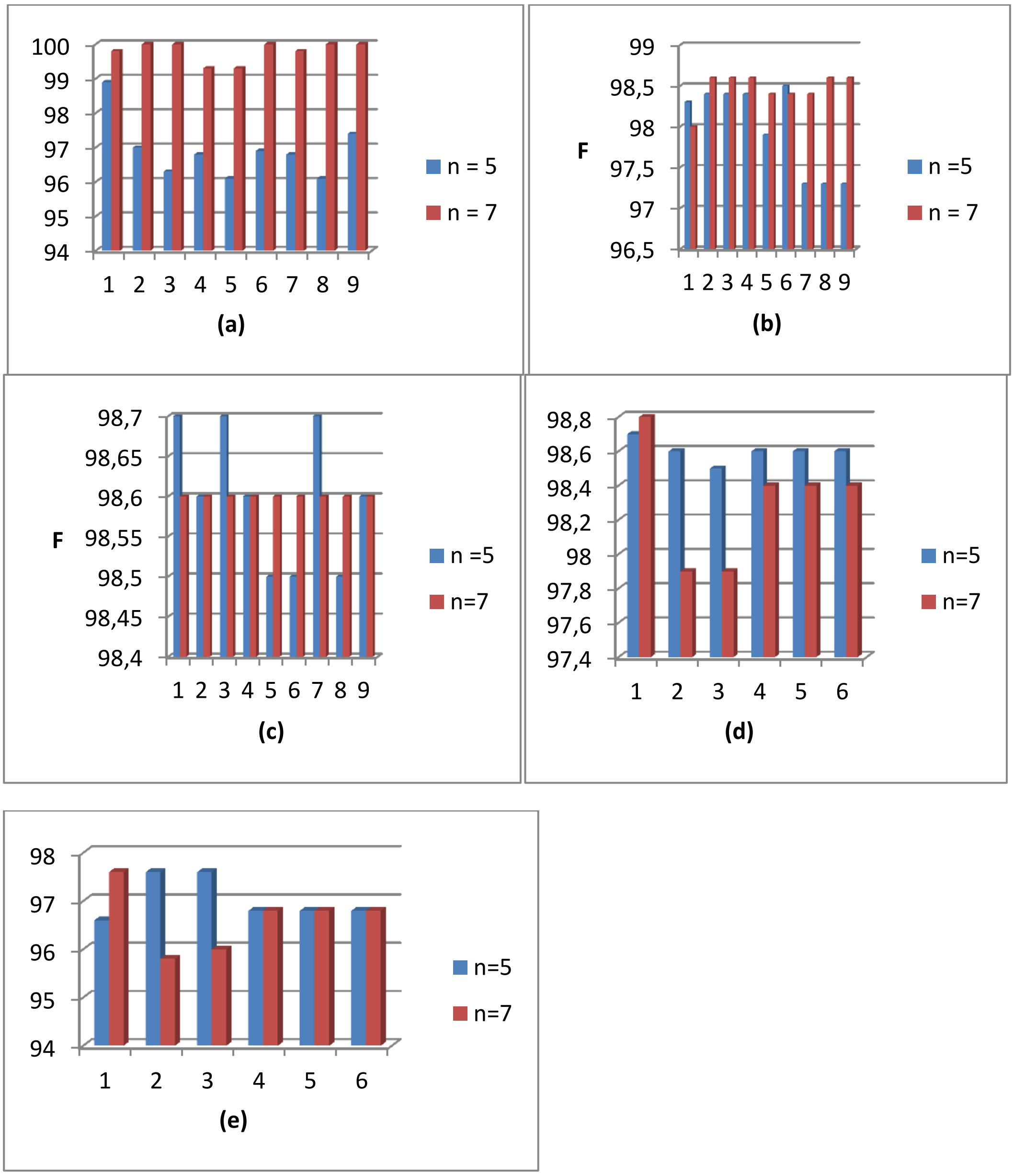


Figure 7. Percentage chart of k_{sw} and k_{sw1} in (a) image 5, (b) image 6, (c) image 7, (d) image 8 and (e) image 9

Figure 7 (a), most of combination k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 5, using k=0.05 and n=5. Figure 7 (b), most of combination k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 6, using k=0.08 and n=7. Figure 7 (c), combination 1,3 and 7 k_{sw} and k_{sw1} in local window 5 give best result to final binarized image sampel 7, using k=0.08 and n=7. Figure 7 (d), combination 1 of k_{sw} and k_{sw1} in local window 7 give best result to final binarized image sampel 8, using k=0.09 and n=3 and Figure 7 (e), combination 1 of k_{sw} and k_{sw1} in local window 7 and combination 2,3 of k_{sw} and k_{sw1} in local window 5 give best result to final binarized image sampel 9, using k=0.04 and n=9.

Table 1: Final Binarized Evaluation based on F-Measure of Different Binarization Method

	F-Measure (%)									
	image 1	image 2	image 3	image 4	image 5	image 6	image 7	image 8	image 9	Avg
Sauvola + Filter Wiener	100	100	100	100	100	97.9	97.8	99	96.8	99.06
Gatos	100	100	100	100	100	98.6	98.7	98.8	98.8	99.43
Sauvola	98.6	98.5	99.5	98.8	98.3	95.6	98	0	74.3	84.62

Based on the Table 1, Gatos's method shows the highest average value of F-measure compare to Sauvola's method using source image and Sauvola's method final thresholding using Wiener filter image result.

4. Conclusion

By using optimal parameter of n, k, k_{sw} and k_{sw1} in Gatos's method, final binarized image has better quality and well recognized by OCR compare to Sauvola's method using original image and Wiener filter image result, based on F-measure value. Gatos's method has the highest average value of F-measure 99.43%, meanwhile Sauvola's method using source image without preprocessing has 99.06% and Sauvola's method using Wiener filter image result has 84.62%. It means that preprocessing also effect the quality of binarized image result.

References

[1] Feng, M.L., and Tan, Y.P. Adaptive Binarization Method for Document Image Analysis. 2004; (1):339-342.
[2] Otsu, N. A threshold selection method from gray-level histograms. *IEEE Transactions Systems, Man and Cybernetics*,1979; 9(1):62–66.
[3] Sauvola, J. and Pietikainen, M.. Adaptive document image binarization. *Pattern Recognition*, 2000; 33(2):225–236.
[4] Gatos, B., Pratikakis, I., and Perantonis, S. J. Adaptive degraded document image binarization. *Pattern Recognition*, 2006; 39(3):317–327.
[5] Seeger, M., and Dance, C. Binarising Camera Images for OCR. *Sixth International Conference on Document Analysis and Recognition (ICDAR'01)*, Seattle, Washington. 2001.
[6] Chen, J.X., and Wang, W.J. An Improved Interpolation of Medical Images. GVIP 5 Conference. Kairo, 19-25 Desember 2005.
[7] Halabi, Y.S., Sa'sa, Z., Hamdan, F., and Yousef, H.K. Modeling Adaptive Degraded Document Image Binarization and Optical Character System. *European Journal of Scientific Research*. 2009; 28(1):14-32.
[8] Sobh, T. Innovations and Advanced Techniques in Computer and Information Sciences and Engineering. Springer, Netherland. 2007.
[9] Gonzalez, R.C., and Woods, R.E.. 2002. Digital Image Processing (2rd Edition). Prentice Hall, New Jersey, USA.
[10] Bukhari, S.S., Shafait, F., and Breuel, T.M. Adaptive Binarization of Unconstrained Hand-Held Camera-Captured Document Images. 2009; (15): 3343-3363.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)