❏   264

Improvement of binarization performance using local otsu thresholding

Khairun Saddami[1], Khairul Munadi[2], Yuwaldi Away[3], Fitri Arnia[4]
[1]Postgraduate Program in Engineering, Syiah Kuala University, Indonesia
[2,3,4]Department of Electrical and Computer Engineering, Syiah Kuala University, Indonesia

Article Info

ABSTRACT

Ancients document usually contains multiple noises such as uneven-background, show-through, water-spilling, spots, and blur text. The noise will affect the binarization process. Binarization is an extremely important process in image processing, especially for character recognition. This paper presents an improvement to Nina binarization technique. Improvements were achieved by reducing processing steps and replacing median filtering by Wiener filtering. First, the document background was approximated by using a Wiener filter, and then image subtraction was applied. Furthermore, the manuscript contrast was adjusted by mapping intensity of image value using the intensity transformation method. Next, the local Otsu thresholding was applied. For removing spotting noise, we applied labelled connected component. The proposed method had been testing on H-DIBCO 2014 and degraded Jawi handwritten ancient documents. It performed better regarding recall and precision values, as compared to Otsu, Niblack, Sauvola, Lu, Su, and Nina, especially in the documents with show-through, water-spilling and combination noises.

Corresponding Author:

Fitri Arnia,
Department of Electrical and Computer Engineering,
Faculty of Engineering, Syiah Kuala University,
Jl. Syech Abdurrauf No.7, Banda Aceh, Indonesia.
Email: f.arnia@unsyiah.ac.id

1. INTRODUCTION

The consequence of interaction among nations centuries ago leaving various types of heritage, one of those is ancient documents. This document contains information on history, science, war tactics and stories of inspiration. South East Asia is a region which had many types of heritages such as the ancient document, which was written in the Malay language by using Jawi alphabets. Jawi characters are a variance of the Arabic alphabet that has six additional characters to cover the sound of the Malay language. Jawi alphabets had been used for centuries in South East Asia. Jawi ancient documents are saved in some countries like Indonesia, Malaysia, Germany, The Netherlands, and USA [1-3]. Some of the ancient documents were saved in digital image format.

To preserve information and knowledge in Jawi ancient document, an application that can recognize text in image format is required. The application that transforms the image into a digital text is called Optical Character Recognition (OCR). In building OCR, several steps such as pre-processing, segmentation, feature extraction and pattern recognition are required.

The pre-processing procedure is important because it will establish the OCR become more accurate. Some of the pre-processing steps that play the important role in OCR application are binarization and denoising (noise removal). Many degraded documents had to be enhanced by using denoising and binarization process thus make the manuscript easier to segment [4].

Generally, there are two types of digital document: noisy and noiseless document and binarizing a noisy document is more challenging than that noiseless ones. In ancient documents, the noise not only caused by the digitalization process but also due to the age of the documents. Examples of noise that contained in an ancient document are uneven-background, blurring-text, text-fading or the combination of them [5]. Others caused by water spilling, spot, fox, text-fading, and show-through.

Previously, several methods had been developed to transform an ancient document image into a binary image. Some of the techniques are Otsu [6], Niblack [7], Nick [8], Sauvola [9] and some latest techniques such as Nina [10], Lu [11] and Su [12]. In earlier research, Otsu method was benchmark methods for binarizing document by determining global thresholding. All of the methods were tested on DIBCO databases. However, DIBCO has less noise in the document than Jawi ancient document. So that, these techniques had to be retested in ancient Jawi manuscript that contains miscellaneous noise. The noise came from the variety condition.

Several researchers tested the method for Jawi ancient document [13-16]. The result of this testing informed the existing method did not perform well for Jawi ancient document. To improve denoising method, DCT was applied ahead existing techniques [13]. In this paper, DCT has collaborated with Otsu, Niblack, Nick, and Sauvola binarization technique. The result showed the DCT makes the method perform better than without applying the DCT.

This paper proposes an improvement technique of Nina binarization [10]. Nina introduced six steps of document binarization technique which shown in Figure 1. Nina proposed median filtering for background estimation, contrast compensation, bilateral filtering, recursive Otsu and despeckling. Therefore, this paper suggests a modification Nina binarization method by replacing median filtering with Wiener filtering for background estimation, improving contrast variation with mapping high saturation method and displaced recursive Otsu by local Otsu. Furthermore, bilateral filtering step and despeckling algorithm are removed.
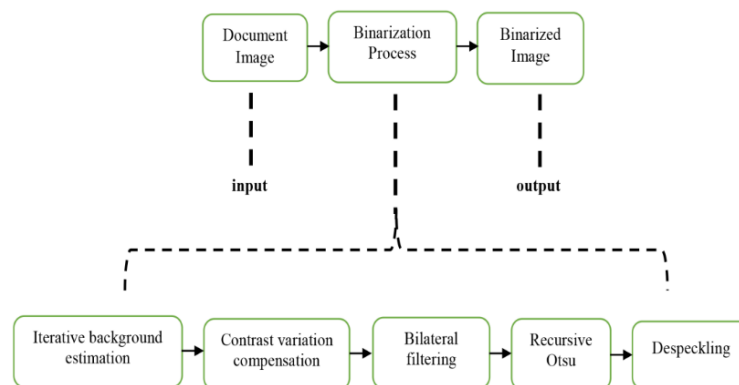

Figure 1. Nina binarization method

2. THE PROPOSED METHOD

In this section, we describe the proposed method of document binarization. Figure 2 shows the proposed technique procedure for ancient manuscript binarization. Overviewed the proposed method as shown in Figure 3.
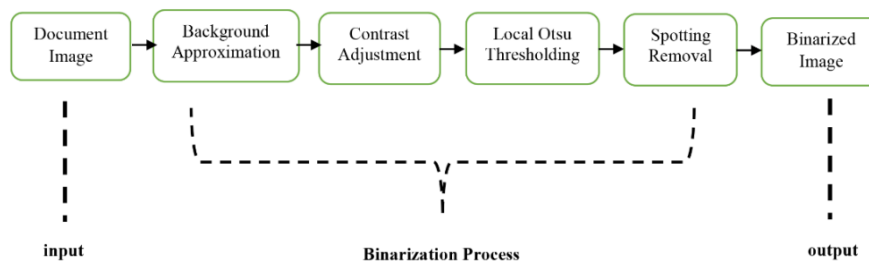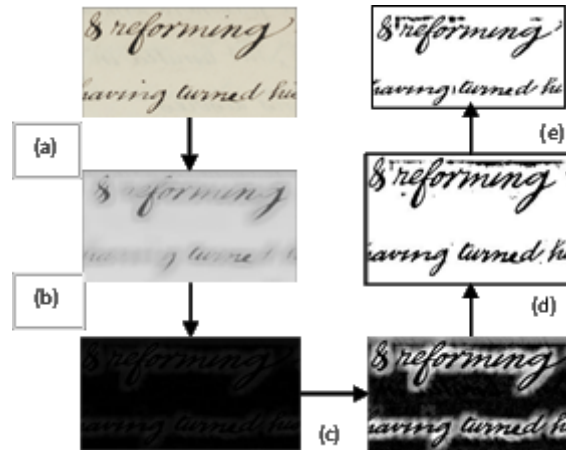

Figure 2. The proposed method

Figure 3. Overviewed the proposed method

2.1. Background approximation and subtraction

The first step of binarization in cleaning degraded ancient document is to approximate the background of the document. Background approximation step was introduced by Hutchison [17]. In this technique, we used Wiener filter [18] to estimate the background. The main idea of Wiener filtering for approximate the image background is to remove the entire noise or others component that showed in a non-text area. The size of a window of Wiener filtering will affect the background approximation performance. In this experiment, we used 47x47 window size. The area of foreground (text) was determined by bold region. Removing the background was performed by subtracting the approximating background from the image. The result of estimated background image was given by:

$$I_n = I - I_{EB} \tag{1}$$

I_n is the new image after background estimation performed and I refers to the acquisition image while I_{EB} is the background approximation image by using Wiener filter.

2.2. Contrast adjustment

After we get the result of the background estimation procedure, the input image was enhanced the image contrast by mapping the intensity of grayscale image value to new image value by using intensity transformation method [19]. Image contrast was adjusted to increase the difference between text and noise. This difference will raise the threshold performance efficiently. Figure 3c showed image contrast adjustment processing and the result gets an image in high contrast.

2.3. Local otsu thresholding

The next stage in this technique is local Otsu thresholding. Once we get the result of the contrast adjustment we use the local Otsu thresholding. This thresholding is local thresholding of Otsu binarization. We set the window of block processing manually but we use this window use for each image. Before performing local Otsu thresholding, we applied Wiener filtering for removing the noise.

The main idea of local Otsu thresholding is to process the image by applying Otsu technique on a local window. In this case, Otsu method will get different threshold on each window size, it depends on image condition. Otsu technique has proven as an effective algorithm in binarizing a document. Otsu method was described as:

$$\sigma^2 B\ (T) = \ \omega_1(T)\omega_2(T)[\mu_1(T) - \mu_2(T)]^2 \tag{2}$$

where ωi represent pixels in the class and μi represent mean of the class.

2.4. Spotting removal

Spotting removal is a process to remove unnecessary spot noise from the binary image. In the proposed method, spotting removal is to remove the unnecessary small shape that occurred in the image. This procedure concept is all object less than 50 pixels will remove from the image. Figure 3e shows the result of

noise removal. The deficiency of this performance might be the dot from a character will be removed from the image, but it will remove the image noise effectively without using any filtering.

3.   RESEARCH METHOD

The proposed binarization technique was tested on H-DIBCO database and five types of degraded ancient Jawi document. The noisy document content five kinds of degraded Jawi document: spot, show-through, text-fading, water spilling, and the combination of noise. Figure 4 shows noises on Jawi ancient document.

(a)   (b)   (c)

(d)   (e)

Figure 4. Examples of the noise in Jawi ancient document (a. text-fading noise; b. show-through noise; c. spot noise; d. water spilling noise; e. the combination of noise)

This binarization technique was evaluated by using recall and precision methods [20].

$$recall = \frac{CD}{CBD} \tag{3}$$

$$precision = \frac{CD}{GT} \tag{4}$$

Recall is a number of characters correctly detected per total number of characters detected and precision is the number of characters correctly detected per total number of characters in a document. CD is total numbers of character in a document that recognized correctly. GT is the total number of ground-truth in the document. CBC is the total number of detected including correctly and broken character. This method used to evaluate how good the technique extracted the ground truth from a noisy image. The result represented by six parameters: Ground truth character (GT), correct detected character (CD), broken character (BR), missing character (M), precision and recall. F-measure is the harmonic mean of recall and precision.

4.   RESULTS AND ANALYSIS

In this section, we presented and discussed the result of research. The result is seperated into two sub section: HDIBCO 2014 dataset and ancient Jawi dataset.

4.1. HDIBCO 2014 dataset

The proposed method was tested on H-DIBCO 2014 database and compared it to Otsu, Niblack, Sauvola, Lu, Su and Nina binarization technique. The results are shown in Table 1. The recall and precision value is the average value of recognition rate. The proposed method had recognition rate 76.60% of recall and 79.52% of precision. In second placed, Otsu method had accuracy 66.54% of precision and recall. Otherwise, Niblack is the lowest recognition rate with accuracy is 0% of recall and precision. Figure 5 showed binarization result on H-DIBCO 2014 database.

Table 1. Result of H-DIBCO 2014 (Average Result)

Methods	GT	BR	CD	M	Precision	Recall
OTSU	517	173	344	0	0.6654	0.6654
NIBLACK	517	425	0	92	0.0000	0.0000
SAUVOLA	517	438	1	73	0.0020	0.0023
LU	517	261	228	28	0.4401	0.4663
SU	517	263	252	11	0.4874	0.4893
NINA	517	190	315	1	0.6093	0.6238
PROPOSED	517	102	396	19	0.7660	0.7952

The result showed the proposed method had the highest recognition rate. Otsu method is in second place of accuracy. On the database, local binarization such as Niblack and Sauvola had a worse result. Most of the broken characters result caused by reducing shape that performed by binarization technique such as Lu and Su. Figure 5 shows examples of the result H-DIBCO 2014 binarization by using Otsu, Niblack, Sauvola, Lu, Su, Nina and the proposed method.
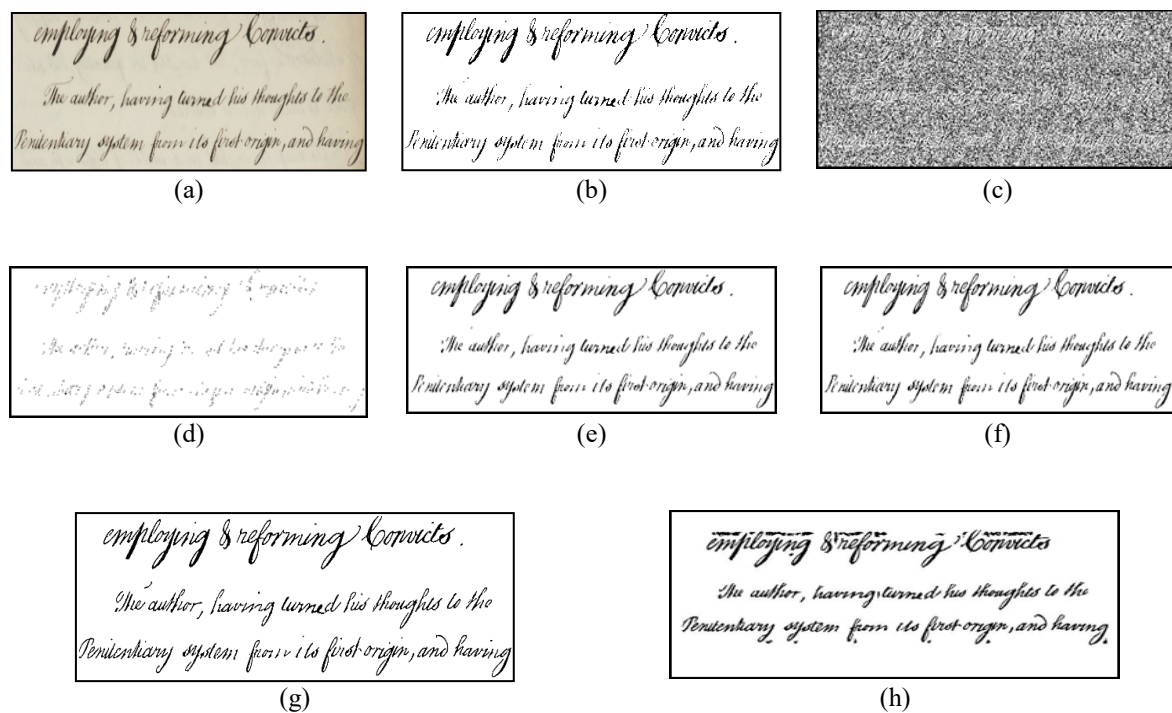

Figure 5. The result of H-DIBCO binarization performance: (a) original image; (b) Otsu method; (c) Niblack method; (d) Sauvola method; (e) Lu method; (f) Su method; (g) Nina method; (h) the proposed method

4.2. Ancient jawi dataset

Figure 6 provides a comparison of the image that binarized by using Otsu, Niblack, Sauvola, Lu, Su, Nina and the proposed techniques. The result of the experiment method shows in Table 2 to Table 6 Table 2 shows the result of binarizing documents which contained combination noise. The result shows that the proposed method has the best result of recall and precision. The proposed technique has 98.4% of precision

and recall. Sauvola technique is in the second place that has 97.58% of precision and recall, while Su in third place with 97.17% of precision and recall. The lowest accuracy of precision and recall is in Niblack technique. Niblack has 33.1% of precision and recall.

Table 2. Result of Combination Noise Binarization

Methods	GT	BR	CD	M	Precision	Recall
OTSU	248	9	233	6	0.9628	0.9395
NIBLACK	248	166	82	0	0.3310	0.3310
SAUVOLA	248	6	242	0	0.9758	0.9758
LU	248	16	232	0	0.9354	0.9354
SU	248	7	241	0	0.9717	0.9717
NINA	248	12	236	0	0.9516	0.9516
PROPOSED	248	4	244	0	0.9840	0.9840

Figure 6. Examples of binarized Jawi ancient document (top: from left to right (original image, otsu, niblack, and sauvola techniques), bottom: from left to right (lu, su, nina, and proposed techniques))

Table 3 informs the result of binarizing documents which contained show-through noise. The result indicates that the proposed technique has the highest result of recall and precision. The proposed technique has 98.34% of precision and recall. Sauvola technique was in the second place that has 95.04% of precision and recall, while Su was in third place with 91.73% of precision and recall. The lowest accuracy of precision and recall has in Niblack technique. Niblack has 1.41% of precision and 8.12% of recall.

Table 3. Result of Show-through Noise Binarization

Methods	GT	BR	CD	M	Precision	Recall
OTSU	121	78	20	23	0.2041	0.1652
NIBLACK	121	70	1	50	0.0141	0.0082
SAUVOLA	121	6	115	0	0.9504	0.9504
LU	121	14	105	2	0.8824	0.8677
SU	121	10	111	0	0.9173	0.9173
NINA	121	9	112	0	0.9256	0.9252
PROPOSED	121	2	119	0	0.9834	0.9834

Table 4 provides the result of binarizing documents which contained spot noise. The result refers that the proposed method has the highest result of recall with accuracy 92.4%. The highest result of precision was in the ancient documents that binarized by using Su technique. It has 93.24% of accuracy while the proposed technique has 92.4 % of accuracy. The result informs, although Su technique has better performance in the precision result, Su technique is worse in the recall. By using Su method, the binarizing document of spot noise has miss three characters, but the proposed method can recognize all of the

characters. The lowest accuracy of precision and recall has in Niblack technique. Niblack has 8.54% of precision and 7.6% of recall.

Table 4. Result of Spot Noise Binarization

Methods	GT	BR	CD	M	Precision	Recall
OTSU	92	42	46	4	0.5227	0.5
NIBLACK	92	75	7	10	0.0854	0.0760
SAUVOLA	92	15	76	1	0.8352	0.8260
LU	92	44	45	3	0.5056	0.4891
SU	92	6	83	3	0.9324	0.9022
NINA	92	10	82	0	0.8913	0.8913
PROPOSED	92	7	85	0	0.9240	0.9240

Table 5 provides the result of the binarizing manuscript which contained text-fading noise. The result shows that the proposed technique has the highest result of recall with accuracy of 44.14%. The highest result of precision was in the document that binarized by using Nina technique which 59.26% of accuracy, while the proposed method has 48.52% of accuracy. The result informs, although Nina technique has better performance in the precision result, Nina technique is worse than the proposed method in the recall. By using Nina method, the binarizing manuscript of text-fading noise image has missed 60 characters while the proposed method only missed 20 characters. The lowest accuracy of precision and recall are in Niblack technique. Niblack has 5.2% of precision and 3.6% of recall.

Table 5. Result of Text-fading Noise Binarization

Methods	GT	BR	CD	M	Precision	Recall
OTSU	222	157	47	18	0.2304	0.2117
NIBLACK	222	146	8	68	0.0520	0.0360
SAUVOLA	222	119	48	55	0.2874	0.2162
LU	222	96	50	76	0.3425	0.2252
SU	222	86	96	40	0.5275	0.4324
NINA	222	66	96	60	0.5926	0.4324
PROPOSED	222	104	98	20	0.4852	0.4414

Table 6 shows the result of binarizing documents which contained extremely noise due to water spilling. The result indicates that the proposed method has the highest result of recall and precision. The proposed technique has 61.9% of precision and 60.71% of recall. Nina technique was in the second place that has 61.54% of precision and 58.93% of recall, while Sauvola in third place with 44.9% of precision and 39.28% in recall. The lowest accuracy of precision and recall has in Otsu technique. Otsu has 7.14% of precision and 1.78% of recall.

Table 6. Result of Water Spilling Noise Binarization

Methods	GT	BR	CD	M	Precision	Recall
OTSU	56	13	1	42	0.0714	0.0178
NIBLACK	56	47	8	1	0.1454	0.1428
SAUVOLA	56	27	22	7	0.4490	0.3928
LU	56	35	20	1	0.3636	0.3571
SU	56	51	5	0	0.0893	0.0893
NINA	56	20	32	4	0.6154	0.5893
PROPOSED	56	21	34	1	0.6190	0.6071

Generally, the proposed method has better performance of segmented Jawi character from a noisy background. Comparing to Otsu, Niblack, Sauvola, Lu, Su, and Nina methods, the proposed technique is the highest for recall and precision the character from the document, except in documents with spot and text-fading noises. However, Su missed more characters in spot noise and Nina in text-fading noise comparing to the proposed method. Lu and Su were performed well in denoising character document but performed worse for bold character or too thin character because Lu and Su techniques reduce shape of the characters.

5.   CONCLUSION

This paper presented an improvement of binarization technique for binarizing degraded Jawi ancient document. This technique combined Wiener filtering for background estimation, image contrast adjustment, and local Otsu thresholding to extract the text from the background and spot noise removal. Five noise types were tested in the experiments: spot, text-fading, show-through, water spilling and the combination of noises. Meanwhile, this technique also tested on H-DIBCO 2014 database.

The proposed technique was compared to Otsu, Niblack, Sauvola, Lu, Su and Nina binarization techniques. The result showed that the proposed technique got the highest value of recall and precision, especially in documents with show-through, water-spilling and combination noises. Furthermore, the proposed method also had the highest recall in spot and text-fading noises.

REFERENCES

[1]   A. Sariyan, "Standardization of Malay Language: Philosophy, history, and its aspects (in Bahasa)," *Dewan Bahasa, Bil. S*, Vol 33, pp. 600-611, 1989.
[2]   K. Saddami, *et al*., "Moment Invariant-Based Features for Jawi Character Recognition," *International Journal on Electrical and Computer Engineering*, under-review.
[3]   K. Saddami, *et al*., "A Database of Printed Jawi Character Image," *2015 IEEE Conference on Image Information Processing (ICIIP)*, pp. 56-59, 2015.
[4]   L.F. Wong and M.Y. Idris, "ATLAS: Adaptive Text Localization Algorithm in High Color Similarity Background," *Telecommunication Computing Electronics and Control* (*TELKOMNIKA*), vol. 13, pp, 963-975, 2015.
[5]   F. Arnia, et al., "Noise Characterization in Ancient Document Image Based on DCT Coefficient Distribution*," 2015 IEEE 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 971-975, 2015.
[6]   N. Otsu, "A Threshold Selection Method From Gray-Level Histogram," *IEEE Transactions on System, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 62-66, 1979.
[7]   W. Niblack., "Introduction to Digital Image Processing," *Prentice-Hall*, pp 115-116, 1986.
[8]   K. Khurshid,et al., "Comparison of Niblack inspired binarization methods for ancient documents," *SPIE Proceedings, 16th Document Recognition and Retrieval Conference, DRR-09*, col 7247, pp. 1-10, 2006.
[9]   J. Sauvola and M. Pietikainen, "Adaptive Document Image Binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225-236, 2000.
[10] S. Lu, et al., "Document Image Binarization Using Background Estimation and Stroke Edges," *International Journal on Document Analysis and Recognition*, vol. 13, 2010.
[11] O. Nina, et al., "A recursive Otsu Thresholding Method for Scanned Document Binarization," *IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 307-314, 2011.
[12] B. Su, et al., "Robust Document Image Binarization Technique for Degraded Document Images," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1408 – 1417, 2013.
[13] F. Arnia, et al, "Improvement of Binarization Performance by Applying DCT as Pre-processing Procedure*," in Internasional Conference on,* ISCCSP, pp. 128-132, 2014.
[14] F. Arnia and K. Munadi, "Binarization of Ancient Document Images based on Multipeak Histogram Assumption," *Telkomnika*. vol. 15(3), pp, 1317-1327, 2017.
[15] S. Muchallil. *et al.*, "Performance Comparison of Denoising Methods for Historical Documents," *Jurnal Teknologi*, vol.77, pp. 143-137, 2015.
[16] Fardian, *et al.*, "Identification of Most Suitable Binarisation Method For Acehnese Ancient Manuscripts Restoration Software User Guide," *Jurnal Teknologi*, vol. 77, no. 22, pp. 95-102, 2015.
[17] L.A.D Hustchison and W.A. Barrett, "Fast Registration of Tabular Document Image Using The Fourier-Mellin Transform," 2004 Proc. *IEEE International Conference on Document Image Analysis for Libraries*, pp. 253-267, 2004.
[18] A.D. Poularikas and Z.M. Ramadan, "Adaptive Filtering Primer with MATLAB," *CRC Taylor and Francis: London*, 2006.
[19] M. Cheriet, *et al.*, "A Recursive Thresholding Technique for Image Segmentation," *IEEE Transactions on Image Processing*, Vol. 7, No. 6, pp. 918-921, 1998.
[20] Q. Wang and C.L. Tan, "Matching of Double-Sided Document Image to Remove Interference," *2001 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. I-1084-I-1089, 2001.

BIOGRAPHIES OF AUTHORS

Khairun Saddami received the B.Eng. degree in electrical engineering from Syiah Kuala University, Banda Aceh, Indonesia, in 2015. He is currently a PhD candidate in Postgraduate program of Electrical and Computer Engineering at Syiah Kuala University. He is also research assistant in Multimedia and Signal Processing research group (Musig), Electrical and Computer Engineering Department, Syiah Kuala University. His research interests include computer vision and image processing.

Khairul Munadi received the B.E. degree from Sepuluh Nopember Institute of Technology, Surabaya, Indonesia, in 1996, and the M.Eng. and Ph.D. degrees from Tokyo Metropolitan University (TMU), Japan, in 2004 and 2007 respectively, all in electrical engineering. From 1996 to 1999, he was with Alcatel Indonesia as a system engineer. Since 1999, he joined Syiah Kuala University, Banda Aceh, Indonesia, as a lecturer at the Electrical Engineering Department. He was a visiting researcher at the Information and Communication Systems Engineering, Faculty of System Design, TMU, Japan, from March 2007 to March 2008. His research interests include multimedia signal processing and communications.

Yuwaldi Away Yuwaldi Away, he was born in Tapaktuan, Aceh Selatan, Indonesia in 1964. He received his B.Eng degree in Electrical-Computer Engineering from "Sepuluh Nopember" Institute of Technology (ITS) Surabaya, Indonesia, the M.Sc degree from "Bandung" Institute of Technology (ITB) Bandung, Indonesia. He obtained his Ph.D. in Computer Technology from the National University of Malaysia. He joined as teaching staff in Syiah Kuala University start from 1990, and from 2007 until now he as a professor in Electrical Engineering. His current research interest includes the microprocessor-based system, embedded system, FPGA, optimation, and visualization.

Fitri Arnia received B. Eng degree from Universitas Sumatera Utara (USU), Medan in 1997. She finished her master and doctoral degree from Universsity of New South Wales (UNSW), Sydney, Australia and Tokyo Metropolitan University, Japan in 2004 and 2008 respectively. She has been with the Department of Electrical Engineering, Faculty of Engineering, Syiah Kuala University since 1999. Dr. Arnia was a visiting scholar in Tokyo Metropolitan University (TMU), Tokyo, Japan in 2013 and Suleyman Demirel University (SDU), Isparta, Turkey in 2017. She is a member of IEEE and APSIPA. Her research interests are signal, image and multimedia information processing.