IMPROVED SAUVOLA THRESHOLD FOR BACKGROUND SUBTRACTION ON MOVING OBJECT DETECTION

Elindra Ambar Pambudi¹, Ermadi Satriya Wijaya², Achmad Fauzan³

^{1,2,3} Dept. Informatics Engineering, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia

²Email : ermadi.satriya@ump.ac.id ³Email : mr.achmadfauzan@gmail.com

Correspondence Author : ¹elindraambarpambudi@ump.ac.id

ABSTRACT

Image Segmentation is one essential processing on moving object detection. The one of common segmentation methods is thresholding. In this paper, Thresholding method based on adaptive local technique using local mean and standard deviation is known as 'WAN' method. WAN has been inspired by the Sauvola's binarization method and exhibits its robustness and effectiveness when evaluated on low quality document images. The objective of the WAN to enhance the sauvola method and to get a better binarization result and enhance the accuracy. This research aims to produce output value of WAN algorithm. WAN would be compared to other existing adaptive local method like sauvola and niblack. This research is implemented by using matlab and four videos original from camera. The best result calculation error (MSE,PSNR) of WAN method are (0.0011, 53,6655). Overall, the result of WAN method in this paper is more effective and efficient than the other existing method based on MSE and PSNR.

Keywords: Vision Computing, Adaptive Local Threshold, WAN, Moving Object Detection.

I. INTRODUCTION

Moving object detection is commonly technique and essential research in the case of video processing and computer vision. It is necessary for surveillance applications, monitoring system for smart tracking of moving objects. Utilization computer vision and image processing of our research to support tracking object in the part of segmentation process.

Many researchers commonly used combination background subtraction and image segmentation process in case of moving object. Segmentation process in moving objects is one of the most basic and common tasks in the part of computer vision areas and leaves an active field of scope of video processing. Image segmentation is an essential role in video processing, since image segmentation is often to be the first step which must be successfully taken before the next tasks such as feature extraction, classification, and description (Soeleman et al., 2012). Image segmentation methods could be categorized into region based, edge based, clustering, split and merge, normalized Cuts, and Thresholding (Nikita Sharma, Mahendra Mishra, 2012).

The main steps of moving object are the calculation of background subtraction which is the difference each pixel of the incoming frames between frame's foreground and frame's background is generated at the first time. Next step, the binary processing mask is obtained by using a threshold. All these methods have many difficulties including, the threshold selection which is a crucial operation for the methods described by previous research.

Thresholding is one of the several image segmentation techniques. It used to segment foreground and background an image. This paper will focus on adaptive local threshold selection based on using local mean and standard deviation. Many adaptive local thresholding methods has been related by image segmentation research.

Muhammad Zia ur Rehman et.al explained method to segment lung image from Lung Image Database Consortium (LIDC). The purpose of this research is to obtain acceptable segmentation result and short processing time. Better Lung segmentation method will improve performance CAD system. The intensity based threshold has chosen as an approach segmentation method in this paper (Zia ur Rehman et al., 2016).

M.Chandrakala analyzed quantitavely about the approaches local adaptive thresholding. This paper compared five thresholding techniques which is otsu, niblack, bernsen, feng and sauvola. Five existing techniques examined in real world images and printed text. Result of this research is Niblack better outcome than other threshold method (Chandrakala, 2016).

Sakhare Varsha Dnyandeo reviewed basic and advanced thresholding approaches in the case of automatic recognition for vehicle number plate. This paper presented adaptive thresholding using mean, median, gaussian filter, and integral image. In the part of adaptive thresholding using integral image approach, the author chosen bradely and roth's and wellners method. Based on result image of this paper shown the approach using integral image, gaussian filter and mean are better than median techniques (Dnyandeo & Nipanikar, 2016).

Julian Ghaye presented about exploration adaptive global and local segmentation techniques for a lab-on-chip nutrition monitoring system (NutriChip). They used global threshold (otsu, T-Point) and local threshold (sauvola) that is combined with Top Hat Mathematical Morphology (THMM) filter for localizing sub-resolution fluorescent biomarkers and classifying fluorescence microscopy images, the result his proposed method was proven that sauvola threshold has gotten better result to classify fluorescent images then T-point and Otsu (Ghaye et al., 2013).

Rashmi Saini proposed eight local image binarization methods, concepts, advantage, and drawback. In part of conclusion, there are written adaptive thresholding methods are utilized to address such issues and to preserve textual information for better segmentation result than global thresholding (Saini, 2015).

Senthilkumaran et.al described a locally adaptive thresholding technique that removes background by using local mean and standard deviation. They compare niblack and sauvola thresholding implemented on medical images. The result of this paper is the niblack having better performance than sauvola based on calculation PSNR and jaccard similarity coefficient (Senthilkumaran & Vaithegi, 2016).

Kalaiselvi et.al explained about comparative study thresholding methods in the case of finding a robust thresholding techniques for general images and MRI head scans. In their paper, otsu's thresholding obtained better result when compared to sauvola, Ridler and Calvard, Kittler and Illingworth, and Niblack based on Region Non-Uniformity (RNU) for evaluation performance (Kalaiselvi, 2017).

Wan Azani Mustafa proposed method inspired by sauvola's threshold. The purpose of his research is to surpass the drawback of sauvola failed to segment if the foreground and background have small contrast. The proposed method is known as 'WAN' method. WAN method is used to improve sauvola algorithm and achieved better threshoding result. In this

paper, WAN had good result in term F-measure (72.274) and NRM (0.093) compared Niblack, Feng, Bernsen, and local adaptive. Utilization WAN method is only implemented in image processing, it has been never examined in other case (Mustafa et al., 2018).

Based on related work especially the research in (Ghaye et al., 2013; Kalaiselvi, 2017; Mustafa et al., 2018; Senthilkumaran & Vaithegi, 2016; Saini, 2015), this paper will try to compare adaptive local threshold method based on remove background using local mean and standard deviation for background subtraction in the case of moving object detection. Adaptive local thresholds implemented in this paper by using is sauvola, niblack, WAN. Our contribution of this research is adaptive local thresholds would be combined with background subtraction to optimization threshold value. This paper will be conducted as follow : Section II explains preprocessing, Section IV describes theory background subtraction, Section V discusses proposed method, and then result will be shown in section VI and finally conclusion will be presented in section VII.

II. LOCAL ADAPTIVE THRESHOLDING TECHNIQUES

Thresholding is the process to create image binary from image grayscale with steps changing all pixels which has zero value for below threshold and one value for above threshold. According Senthilkumaran (Senthilkumaran & Vaithegi, 2016) there are two approaches thresholding methods in image processing namely global threshold and local threshold. Global threshold is used one threshold for all pixels in one image, meanwhile local threshold is selecting an individual threshold for each pixel based on the range of intensity values in its local neighbourhood. Adaptive local thresholding is a fundamental and simple algorithm to segmenting foreground from background with non-uniform illumination. For each pixel in the image, a threshold has to be calculated. Whether the pixel value is below the threshold, then it is set to the background value, otherwise, it assumes the foreground value.

2.1. Niblack Method

Niblack's algorithm specifies a threshold value for each pixel-wise by sliding a rectangular window over the gray level image (Subashini & Sridevi, 2011). Calculation niblack's method based on local mean m and standard deviation S of all pixels around window following Eq. (1).

$$Ts_{(i,j)} = m + k * S_{(i,j)}$$
 (1)

Where m is average value of pixel, standard deviation value of the pixels inside the window, m is local mean, and k is a constant parameter which have values -0.2 (Sankur, 2004).

2.2. Sauvola Method

Sauvola thresholding works as the edge detection method based in sliding window. Sauvola targeting objects which has same size with window. Practically, this algorithm provides a threshold value for each pixel based mean value and standart deviation from adjacent pixel intensities (Senthilkumaran & Vaithegi, 2016). Sauvola thresholding emphasizes of increased niblack thresholding method by counting threshold value using dynamic distance standart deviation grayscale. The threshold calculation define in Eq. (2)

Improved sauvola threshold for background subtraction on moving object detection, Indonesia

$$Ts_{(i,j)} = m * \left(1 - k * \left(1 - \frac{S_{(i,j)}}{R} \right) \right)$$
(2)

Where *R* is the maximum value of the standard deviation (R=128 for a grayscale document), *m* is local mean, and *k* is a parameter which takes positive values in the range 0.5 (Singh et al., 2012).

2.3. WAN Method

WAN method is a part of local adaptive thresholding techniques. It is inspired by sauvola's thresholding method. The drawback of sauvola is depends on contrast between foreground and background, if contrast value is small then sauvola method is failed to segment. Therefore WAN method proposed to handle the disadvantage of sauvola. The naming of WAN itself taken from the first name of the author of paper (Mustafa et al., 2018). The full name of author is Wan Azani Mustafa.

III. PREPROCESSING

For every video has a lot of frames and large number of neighboring ones are likely to be similar. Each video recorded at 30 frames per second. The preprocessing step will be implemented after process frame extraction and search background frame.

Noise is sometimes happened when we take videos. It is gonna be problem if we cannot reduce itself. According paper J.Reno (Renno et al., 2006) there are several ways to reduce noise, one of those using preprocessing (filter and image enhancement). The purpose of preprocessing reduces noise using median filtering, resize image, and crop image according our necessary.

IV. BACKGROUND SUBTRACTION

Background Subtraction is the important step in computer vision. It has general approach for object segmentation include background modelling and updating. The purpose of using background subtraction to result sequence frame from video and detecting all object foreground. Fei et al. (Fei et al., 2015) has utilized first frame from video sequence as an initialize background for approach background subtraction. The scene of process object detection is usually represented in a background model. There are four steps of background subtraction algorithm were explained by cheung and kamath in (Cheung & Kamath, 2004), preprocessing, background modelling, foreground detection, and validation data. The first step is simple preprocessing of image processing. The second is background modelling. In this step, the new frame video is used to count and to process update background model. Background model results of processing several image sequence extraction file video. The usage sequence image purposed to get tolerance value against the changing of light having done since take the picture. The third step is foreground detection, In this process, the reduce grayscale of each frame with grayscale background for getting candidate foreground of grayscale image. The formula of background subtraction shown in Eq. (3)

$$\Delta_{t}(x,y) = |B(x,y) - I_{t}(x,y)|$$

$$O_{t}(x,y) = \int_{0}^{1} if \ abs(\Delta_{t}(x,y)) \ge Ts$$

$$O_{t}(x,y) = \int_{0}^{1} Otherwise$$
(3)

B is the frame of background, I_t is the foreground frame of the sequence. Ts is thresholding.

V. PROPOSED METHODS

The kind of our research is experimental research. This research has an aim to implement adaptive local threshold algorithm in the moving object. This method will run in several steps which consist of taking video data, processing video data to be frame, doing thresholding process, running image sequentially.

The Sauvola method is capable to surpass the issue of black noise, it depends on the impact on the standard deviation value by using a range of grey level values in the images (Bataineh et al., 2011; Senthilkumaran & Vaithegi, 2016). Nevertheless, according WAN sauvola has weakness that is failed to segment if the contrast between the foreground and background is small or if the text is in thin pen stroke text. Therefore, the WAN method was suggested to overcome this issue by achieving the maximum threshold value (Mustafa et al., 2018).



Figure 1. Flowchart proposed method on moving object detection

However, implementing this approach involves the following steps :

- 1. Extracting video to be frames : Every video are having 30 frame. This step will extract frames from a video file to sequence of JPG images.
- 2. Initialitation background frames : In second step, every video is taken, it has been set which the first frame as an background image is given symbol B(x,y).
- 3. Preprocessing and Colour Conversion : Each frame produces RGB image, therefore we have to change to grayscale image that simplifying the next move. Every frame utilizes noise reduction and image enhancement as preprocessing step. Reduction noise is using deletion object that have size less than 50000 (video 1), 40000 (video 2), 5600 (video 3), 500.(video 4).
- 4. Subtraction : After background image B(x,y) was obtained, next step is calculation the difference between image of the scene 's static background B(x,y) and current frame It (x,y) using eq. (3). Utilizing pixel difference, if the pixel is bigger than threshold value Thres and then decides the pixels occur in the moving object, otherwise, set the

pixels occur as background pixels (Zhang & Liang, 2010). Next step, WAN is performed to produce the threshold (Thres) of background subtraction.

5. Improved Sauvola Thresholding (WAN): The subtracted frame is processed by the WAN threshold. According paper in (Mustafa et al., 2018) the explained method aimed to enhance the threshold value to segment the information in the low contrast region and to reduce noise. WAN method will improve original sauvola algorithm eq. (2) Original mean (m) of sauvola will replaced by maximum mean (m_{max}) that shown as follows :

$$m_{max} = \frac{max_{(a,b)} + mean}{2} \tag{4}$$

Therefore, final formula WAN changed becomes,

$$Ts_{(i,j)} = \frac{max_{(a,b)} + mean}{2} * \left(1 - k * \left(1 - \frac{S_{(i,j)}}{R}\right)\right)$$
(5)

Additional formula $max_{(a,b)}$ is the maximum intensity of the input image. Formula WAN have similar with sauvola's method. But, It has little bit different on mean calculation.

- 6. Utilization one of morphology operations is dilation by using structure element 'square' and R = 3. Dilation process in this step is process to add pixel in boundary object on image, therefore the result of dilation image have higher size than original image.
- 7. Creating mask from image the result of binarization using WAN, sauvola and niblack. Created to Every frame biner is annotated with a pixel level object mask.
- 8. Running Frame Sequentially : This step running frame by frame result masking of WAN Threshold images sequentially.

VI. RESULTS AND ANALYSIS

Four datasets video was taken by mobile phone on laboratorium faculty engineering and science of university. This research recorded amount 4 videos with 30 fps frame rate, each video has duration time 5-7 second with mp4 format. All project examined in MATLAB from lenovo notebook Intel Core I3 2GB RAM. Three videos captured in indoor location and the other captured in outdoor.

In this section we present the results of image thresholding by using WAN, sauvola, niblack on different images. We explain the results, comment on them, and compare WAN with other existing adaptive local threshold. Sample results thresholding, background subtraction and original image are shown in figure 2-4. Sample Figure 2 shows background subtraction and threshold using WAN, sauvola, and niblack in 10th frame which is followed 16th frame in Fig. 3 and 21st frame in Fig. 4.



Figure 2. Example thresholding result frame 10th : a. original image, b. subtraction, c. niblack, d. sauvola, e. WAN



Figure 3. Example thresholding result frame 16th : a. original image, b. subtraction, c. niblack, d. sauvola, e. WAN



Figure 4. Example thresholding result frame 21th : a. original image, b. subtraction, c. niblack, d. sauvola, e. WAN

The result masked process is shown in Table 1. Overall, comparative result using other exist threshold is looking good. However, there is one thing that focused on frame tenth. The masking of frame tenth doesn't match size. The size of masking object is exceeding object size.

	Table 1	. Comparative	e sample resu	lt video
--	---------	---------------	---------------	----------





For evaluating performance of the tresholding results that influence quality of the moving objects, we use calculation error rate using MSE and PSNR and comparison average processing time. Formula MSE and PSNR is shown in Eq. (6) and Eq. (7).

$$MSE = \frac{\sum_{a=0}^{x-1} \sum_{b=0}^{y-1} (M_{(a,b)} - N_{(a,b)})}{x * y}$$
(6)

MSE is defined the some sort of average of sum (or integral) of square error between image that have been processed and reference image.

$$PSNR = 10\log_{10}\frac{Max_i^2}{MSE}$$
(7)

PSNR (Peak Signal to Noise Ratio) is mathematical measure for image quality assessment between image that have been processed and reference image. Here Max_i is maximum fluctuation in input image data.According paper (Pinki, 2016) which relation between MSE and PSNR, if the value PSNR decreases the corresponding value of MSE increases. Highly distorted image gives high value of mean square error, less value of peak signal noise ratio, if PSNR value is ∞ then corresponding value of MSE becomes zero and the result value of structural similarity becomes highest.

MSE / PSNR (/ frame)									
Video		Niblack			Sauvola			WAN	
Frame	10	16	22	10	16	22	10	16	22
Video 1	0,0032 /	0,0053 /	0,0011 /	0,0011 /	0,0011 /	0,0011 /	0,0011 /	0,0011 /	0,0011 /
	48,9228	46,7910	53,3225	53,3225	53,3225	53,3198	53,3225	53,3225	53,3225
Video 2	0,0369 /	0,0442 /	0,0427 /	0,0434 /	0,1708 /	0,1292 /	0,0983/	0,4362 /	0,3053 /
	37,7164	37,4794	37,5336	38,2614	31,7397	32,9504	34,1385	27,6678	29,2175

Table 2. Sample MSE & PSNR values for existing methods

Video 3	0,2985 /	0,2327 /	0,2287 /	0,4793 /	0,4272 /	0,5901 /	0,2932 /	0,2247 /	0,2295 /
	29,3151	30,3961	30,4722	27,2584	27,7585	26,3559	29,3936	30,5489	30,4565
Video 4	2,0409 / 20,9671	0,7857 / 25,1127	0,4691 / 27,3526	2,2215 / 20,5989	0,9381 / 24,3429	0,6303 / 26,0698	1,6412 / 14.5351	0,5672 / 14,3749	0,2342 / 14,4803

Table 3. Average MSE & PSNR for existing methods

Sample Videos	Niblack	Sauvola	WAN
Video 1	0,0030 / 50,4065	0,0011 / 53,6653	0,0011 / 53,6655
Video 2	0,0249 / 37,3557	0,0855 / 33,3564	1,3508 / 27,3605
Video 3	0,1494 / 30,3105	1,2796 / 25,3933	0,1478/ 30,3479
Video 4	0,7192 / 25, 1700	0,8156 / 24,3525	0,5315 / 27,6409

Comparative analysis of 10th, 16th and 21st frame in sample videos based on MSE, PSNR are tabulated in Table 2. Average results of MSE and PSNR all frames of WAN and other existing local adaptive method are shown in Table 3. There are two things can be observed based on MSE and PSNR.

- Video 4 is having more brightness than other existing videos, hopefully it can be best result in calculation error. Unfortunalety, video 1 is better than video 4.
- WAN almost dominates best result for all existing methods based on MSE and PSNR, but surprisingly WAN's method is not giving best result in video 2.

		Time Process (<i>Seconds</i>)	
Sample Videos	Niblack	Sauvola	WAN
Video 1	176,815057	196,390661	258,879259
Video 2	235,348578	225,723478	377,164505
Video 3	64,653565	100,013028	83,811543
Video 4	108.792491	125,119201	186,447878

Table 4	Processing	Time for	existing	methods
ruore n	Trocessing	I IIIIC I OI	CAISting	methous

Based on processing time in table 4, generally WAN method has long time to execute each video. WAN focused on enhancement complexity of calculation sauvola method. Therefore more complex the caculation will take a lot of time. But surprisingly, WAN gives better time than sauvola in video 3. For all videos, the third video has shortest duration to all existing videos.

Overall, Figure 5 is shown the error rate for different dataset videos, and the error rate for niblack and sauvola is larger than WAN algorithm except for the third video. Figure 6 shows PSNR value of niblack and sauvola smaller than WAN method.



Figure 5. Block diagram : a. average MSE, b. Average PSNR



Figure 6. Block Diagram Average Processing Time

VII. CONCLUSIONS

This paper investigated the different results adaptive thresholding and background subtraction which is implemented in moving object detection. To extract the detected object is carried out using three different adaptive local threshold techniques are implemented on background subtraction image sequence.

Based on performance measure of MSE and PSNR, WAN technique is enough effective in terms of quality threshold of moving object detection. Overall, These thing has proven that MSE and PSNR value of WAN better than other existing local adaptive method. The best result value of WAN is 0,0011 (MSE) and 53,6655 (PNSR). Conceptually WAN is the

improvement of sauvola method which is just focusing to reduce error value, therefore the processing time will increase automatically.

ACKNOWLEDGEMENT

We would like to express our gratitude to *Lembaga Penelitian dan Pengabdian Masyarakat* (*LPPM*) Universitas Muhammadiyah Purwokerto for supporting our research funding, and my research partners Mr. Ermadi and Mr. Achmad Fauzan for valuable suggestion. This paper would not have been possible without the advice and support of many people.

REFERENCES

- Bataineh, B., Abdullah, S. N. H. S., & Omar, K. (2011). An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. *Pattern Recognition Letters*, 32(14), 1805–1813. https://doi.org/10.1016/j.patrec.2011.08.001
- Chandrakala, M. (2016). Quantitative Analysis of Local Adaptive Thresholding Techniques. International Journal of Innovative Research in Computer and Communication Engineering, 4(5), 8432–8439. https://doi.org/10.15680/IJIRCCE.2016
- Cheung, S. S., & Kamath, C. (2004). Robust techniques for background subtraction in urban traffic video. *Pattern Recognition Letters*, *5308*, 881. https://doi.org/10.1117/12.526886
- Dnyandeo, S. V., & Nipanikar, P. M. R. S. (2016). A Review of Adaptive Thresholding Techniques for Vehicle Number Plate Recognition. *International Journal of Advanced Research in Computer and Communication Engineering*, 5(4), 944–946. https://doi.org/10.17148/IJARCCE.2016.54232
- Fei, M., Li, J., & Liu, H. (2015). Visual tracking based on improved foreground detection and perceptual hashing. *Neurocomputing*, 152, 413–428. https://doi.org/10.1016/j.neucom.2014.09.060
- Ghaye, J., Kamat, M. A., Corbino-Giunta, L., Silacci, P., Vergères, G., De Micheli, G., & Carrara, S. (2013). Image thresholding techniques for localization of sub-resolution fluorescent biomarkers. *Cytometry Part A*, 83(11), 1001–1016. https://doi.org/10.1002/cyto.a.22345
- Kalaiselvi, T. (2017). a Comparative Study on Thresholding Techniques for Gray Image Binarization. *International Journal of Advanced Research in Computer Science*, 8(7), 1168–1172. https://doi.org/10.26483/ijarcs.v8i7.4510
- Mustafa, W. A., Yazid, H., & Jaafar, M. (2018). An improved sauvola approach on document images binarization. *Journal of Telecommunication, Electronic and Computer Engineering*, *10*(2), 43–50.
- Senthilkumaran N., & Vaithegi S. (2016). Image Segmentation By Using Thresholding Techniques For Medical Images. *Computer Science & Engineering: An International Journal*, 6(1), 1–13. https://doi.org/10.5121/cseij.2016.6101
- Nikita Sharma, Mahendra Mishra, M. S. (2012). Colour Image Segmentation Techniques and Issues: an Approach. *International Journal of Scientific & Technology Research*, 1(4), 9–12.
- Pinki, R. M. (2016). Estimation of the Image Quality under Different Distortions. International Journal Of Engineering And Computer Science, 5(17291), 17291–17296. https://doi.org/10.18535/ijecs/v5i7.20
- Renno, J., Lazarevic-McManus, N., Makris, D., & Jones, G. (2006). Evaluating motion detection algorithms: issues and results. *IEEE International Workshop on Visual*

Surveillance, July, 97–104.

- Saini, R. (2015). Document Image Binarization Techniques, Developments and Related Issues: A Review. International Journal of Computer Applications. https://doi.org/10.5120/20352-2541
- Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, *13*(1), 146. https://doi.org/10.1117/1.1631315
- Singh, T. R., Roy, S., Singh, O. I., Sinam, T., & Singh, K. M. (2012). A New Local Adaptive Thresholding Technique in Binarization. *International Journal of Computer Science Issues*, 8(6), 271–277.
- Soeleman, M. A., Hariadi, M., Purnomo, M. H., & Ng, K. K. (2012). Adaptive Threshold for Background Subtraction in Moving Object Detection using Fuzzy C-Means Clustering -Soeleman - Unknown.pdf. TENCON 2012 IEEE Region 10 Conference, 1–5. https://doi.org/10.1109/TENCON.2012.6412265
- Subashini, P., & Sridevi, N. (2011). An optimal binarization algorithm based on Particle Swarm Optimization. *International Journal on Soft Computing and Engineering*, 1(4), 32–36.
- Zhang, L., & Liang, Y. (2010). Motion human detection based on background subtraction. 2nd International Workshop on Education Technology and Computer Science, ETCS 2010, 1, 284–287. https://doi.org/10.1109/ETCS.2010.440
- Zia ur Rehman, M., Gilani, S. O., Shah, S. I. A., Jamil, M., Irfanullah, & Butt, S. I. (2016). Adaptive thresholding technique for segmentation and juxtapleural nodules inclusion in lung segments. *International Journal of Bio-Science and Bio-Technology*, 8(5), 105– 114. https://doi.org/10.14257/ijbsbt.2016.8.5.10