TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 18, No. 2, April 2020, pp. 1122~1129
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v18i2.14878

Adaptive threshold for moving objects detection using gaussian mixture model

Moch Arief Soeleman, Aris Nurhindarto, Muslih, Karis W., Muljono, Farikh Al Zami, R. Anggi Pramunendar
Faculty of Computer Science Dian Nuswantoro University, Indonesia

Article Info

Article history:

Received Aug 14, 2019
Revised Dec 30, 2019
Accepted Feb 20, 2020

Keywords:

Detection
Gaussian mixture model
Moving object
Otsu
Threshold

ABSTRACT

Moving object detection becomes the important task in the video surveilance system. Defining the threshold automatically is challenging to differentiate the moving object from the background within a video. This study proposes gaussian mixture model (GMM) as a threshold strategy in moving object detection. The performance of the proposed method is compared to the Otsu algorithm and gray threshold as the baseline method using mean square error (MSE) and Peak Signal Noise Ratio (PSNR). The performance comparison of the methods is evaluated on human video dataset. The average result of MSE value GMM is 257.18, Otsu is 595.36 and Gray is 645.39, so the MSE value is lower than Otsu and Gray threshold. The average result of PSNR value GMM is 24.71, Otsu is 20.66 and Gray is 19.35, so the PSNR value is higher than Otsu and Gray threshold. The performance of the proposed method outperforms the baseline method in term of error detection.

Corresponding Author:

Moch Arief Soeleman,
Faculty of Computer Science,
Dian Nuswantoro University,
207 Imam Bonjol St., Indonesia,
Email: arief22208@gmail.com

1. INTRODUCTION

Research in video processing has become one of the exciting fields in computer vision research, as the use of methods continues to be developed to produce quality results in the field of computer vision, especially detection of moving objects based on surveillance cameras. The increasing CTTV technology, especially in the use of surveillance cameras in monitoring human activities or activities in daily life both at home, offices, and in open spaces. In the central part of monitoring activities of moving objects, especially humans, the process of detection of moving objects is an essential step carried out at an early stage to be able to carry out more specific operations in object analysis such as extraction, classification, and identification. Moving object detection has some basic formation, mathematically the exposure of moving objects results from modeling the background and foreground [1, 2]. This modeling aims to get pixels that will become foreground-background.

Some approaches in modeling this static background include using statistical component methods such as running average [3, 4], and histogram analysis [5]. Otsu thresholding [6] is a threshold method in segmentation techniques, the application of the Otsu method makes it easier to do homogeneous division of parts based on similarity criteria to recognize objects. Before observing the segmented image, the process must first go through the image input process so that it is easy to proceed to the next process.

The next process is to add brightness to the image to improve image quality. After that, the image segmentation process is done with the Otsu thresholding method and through discriminant analysis approach so that it can maximize these variables so that objects with a background can separate automatically. In a previous study, Soeleman, Mauridhi Hery, Hariadi in [7] used the adaptive threshold approach to get the background of moving objects using fuzzy c-means clustering, and the fuzzy method produced better performance than the Otsu threshold. In another study, Zeng, Jia and Chen in [5] used fuzzy to obtain foreground objects with a histogram approach to partition the foreground from the background video sequence.

In another study, Yesong, Xiaoping Li, Na Fu and Qiongxin Liu in [8] used the gaussian mixture model approach to detect moving objects based on background subtraction. Vivek Maik, Hyungtae Kim, Daedhee Kim, Eunjung Chae and Joonki Paik in [9] modified the gaussian mixture model to overcome sudden changes in the object's background. The same method in the gaussian mixture model is also used in [10-12]. From some of these studies it can be concluded that detecting moving objects is a difficult challenge in the detection method. Kittipop Peuwnuan, Kuntpong Woraratpanya and Kitsuc in [13] using integral image for thresholding adaptive in image, authors proposed two stage for process in low and high intensity for thresholding and segmentation. Senthilkumaran and Vaithegi in [14] using thresholding technique for segmentation medical image, the proposed method have fast performance to extract foreground from background. In different method many researcher using background modelling [15-17] to detect moving objects in video, but background modelling have a problem when it suddenly occurs an intensity change and dynamic background, so the method have disadvantage to detect moving object. In other method Nur Ayuni Mohamed and Mohd Asyraf Zulkifley in [18] using optical flow to detect the motion of human fall-dawn, that evaluation using intersection over union for measuring.

In this paper, the gaussian mixture model method is proposed to detect moving objects with adaptive threshold approach to obtain the foreground from the dynamic background of a frame sequence. As threshold strategic gaussian mixture model were compared against gray threshold and Otsu algorithm to evaluate the achievement. The rest of paper is arranged as follows. Passage 2 discussed our research method in detection moving objects. Passage 3 shows the results and analysis of the experimental our proposed and discussion. Finally, conclusion are given in passage 4.

2. RESEARCH METHOD

The proposed method be composed of several steps to threshold strategic in the moving object detection, as illustrated in Figure 1. In the initial stage, the frame breakdown begins in the video dataset, then adaptive threshold using gaussian mixture model, Otsu and gray threshold in parallel process to get the foreground from video dataset, the next step the binary moving objects in the frame is continued by the morphology operation until the moving object is detected, finally we evaluate all performance of adaptive threshold method.

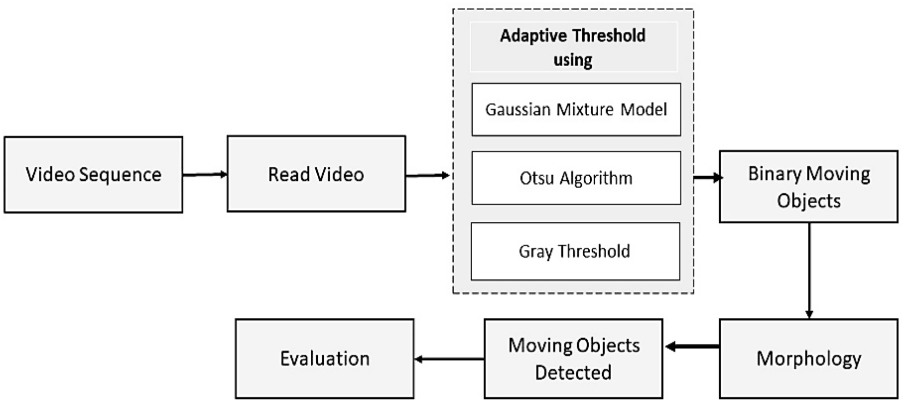

Figure 1. Flow method in threshold for object detection

2.1. Gaussian mixture model

Stauffer et.al in [19] proposed the gaussian mixture model (GMM) method as a type density model consisting of gaussian function components. The component of this function consists of different weights to produce multi-model density. The gaussian mixture model method proposed by Grimson is very efficient in separating the foreground and background from the input frame sequence

The number of GMM in [8, 17] used affects the number of background models. The greater the number of GMM models used, the more background models a pixel has. There are several stages of the process for this method, namely the stage of matching input to the distribution and the stage of selecting the distribution that reflects the background. In GMM each pixel In the matching stage, there is a parameter update stage. The GMM equation model [8] is as (1).

$$P(X_t) = \sum_{i=1}^{K} \omega_{it} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

The parameter K in (1) is the numeral of distributions, while μ is the average value of gaussian at term unit t, and ∑ is the covariance matrix at I threshold on gaussian and ω is the weight. The following equation is a gaussian probability density function.

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} \Sigma^{1/2}} exp^{\frac{1}{2}(x_t-\mu_t)^T} \Sigma^{-1(x_t-\mu_t)} \quad (2)$$

$$\omega_{i,t} = (1-\alpha)\omega_{i,t-1} + \alpha M_{(i,t)} \quad (3)$$

The parameter n is the gaussian distribution dimension, the value of n = 1 when the background model is a gray image, ∑I,t is a color image, and the amount of n = 3 is an RGB image. When the parameter M is 1, then it matches, and vice versa, the parameter M is 0. The parameter values μ and α are updated with the following equation.

$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho X_{i,t} \quad (4)$$

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(X_t - \mu_{i,t})^2 \quad (5)$$

$$\rho = \alpha\eta(X_t, \mu_{i,t}, \sigma_{i,t}) \quad (6)$$

$$B = \arg min_b(\Sigma_{k=1}^{b} w_k > c_f) \quad (7)$$

Using (3) in the parameter update stage, the values of GMM parameters are used to process the next input. Updated values consist of weights, averages, and variants. The weight value is updated every time. After the weight values have been normalized, the total weight of all distributions is not more than 1. Then the mean value of a distribution is updated every time there is a pixel value that matches the distribution value. Using (5) and (6) the standard deviation value of a distribution is updated every time there is a pixel value that matches that distribution. The next step by (7) the pixels which are probabilistic matches the gaussian model, if they are against the background, they will be classified as background pixels and their opponents as the foreground.

2.2. Otsu algorithm

The Otsu method [6] aims to split the grey image histogram upon two different regions automatically without any help from the user to enter the threshold value. Otsu as the threshold method is set to the mark using *z*. The scope value of *z* is in the midst of 1 and R, wherever $R = 255$. The likelihood of every pixel in the i^{th} degree could be resolved by employed (8).

$$p_i = n_i/N \quad (8)$$

Where in n_i is the amount of pixel in the i^{th} degree and N is the unqualified of amount of pixels. The mean grey degree of an image frame use (9).

$$\mu_T = \sum_{i=1}^{R-1} i \times p_i \quad (9)$$

Towards single threshold, Otsu split the pixels inside two group $C_1 = \{0,1,\ldots,z\}$ and $C_2 = \{z+1, z+2, \ldots, R-1\}$. The likelihood of group can be calculated by applying (10).

$$\omega_1(z) = \sum_{i=1}^{z} p_i \qquad \omega_2(z) = \sum_{i=1}^{R-1} p_i \tag{10}$$

$$\mu_1(z) = \sum_{i=1}^{z} \frac{i \times p_i}{\omega_1(z)} \qquad \mu_2(z) = \sum_{i=1}^{R-1} \frac{i \times p_i}{\omega_2(z)} \tag{11}$$

The amount of z can be calculated applying (12) and (13).

$$z' = \max_{1 \le k < L} \alpha_B^2(z) \tag{12}$$

$$\alpha_B^2(z) = \omega_1(z)(\mu_1(z) - \mu_T)^2 + \omega_2(z)(\mu_2(z) - \mu_T)^2 \tag{13}$$

2.3. Gray threshold

A binary image [20] construct by thresholding through a grey scale or color image by surrounding pixel grade to 1 or 0 rely on whether they are above or below the threshold value. This frequently usual to distinct or portion a section object within the image based upon the pixel values. The basic operation thresholding operates on an image as follows:

For pixel $Z\ (r, s)$ scope the image Z

$if\ Z(r, s) > threshold$

$Z(r, s) = 1$

$else$

$Z(r, s) = 0$

end

Towards scale of gray image frame, that pixels have substance a single intensity value, a single threshold must be picked, and for colour image, a detached threshold can be specified for each narrow.

2.4. Binary frame moving object

The result of the thresholding process with the adaptive threshold with gaussian mixture model, Otsu and gray threshold that produces binary images [21] zero and one where the value 1 represents as the object or foreground and the amount 0 as the background.

2.5. Morphology

Morphological filter results to noise from a foreground distribution by using an erosion filter which aims to reduce the dimensions of the vehicle's dimensions to be close to the actual size, while reducing small movements if it does not need to be segmented. Although the results do not look completely clean, this condition is good enough for the detection of moving objects.

Morphology operations in the form of erosion and dilation were done to get better object results based on the shape of the object [22]. Dilation is the process of adding pixels to the boundary of an object in a digital input image, while erosion is the process of moving/reducing pixels at the border of an object Morphological operations include erosion $F' = f \ominus s$ where F' is new binary image from image f have structuring element s, the dilation notation of image f with structuring element s is $F' = f \oplus s$, opening, and closing [23]. At this stage, the disposal process is carried out [24] imperfect pixel objects from the detection result.

3. RESULTS AND ANALYSIS

In this section, the results of the experimental algorithm that have been applied to the threshold of moving objects for object detection are presented in Figure 2. The dataset used consists of 200 frame sequences. In order to judgement our adaptive threshold practice, we use a dataset in realistic with permission to Le2i "Laboratorie Eletronique, Informatique et Image" video surveillance state by using an unaccompanied camera [25]. The video has frame rate is 25 per second, and with resolution on 320 x 240 pixels. The video data explained with principal adversity of reasonable frame sequences which can come across at a human as well as in a straightforward office room. The frame sequences comprise fluctuating of illumination, and distinctive hardship as though occlusions or confound and textured background. The actors bring about certain normal habitual activities. The dataset comprises number of frames. For judgement goal, with extra information representing the ground-truth of the detection position in the image sequence. Then, each frame of each video is labelled: the localization of the body is automatic assigned with bounding boxes. The experiment uses MATLAB 2016b and runs on a CPU with an i5-8265U processor with RAM LPDDR3 8 GB.

For the detailed explanation, we applied our approach in Figure 2, whereas we compared gaussian mixture model with Otsu and gray threshold. For sake the simplicity, we take example as follows: in the frame number 90, the subject arrived at room, thus the bounding box start to detect the subject, then followed by frame number 102 and frame number 168. In the frame number 168 and 179, we can see that gaussian mixture model able to maintain the bounding box to the subject, where the other method is not successful. Thus, from visual evaluation, we can conclude that the gaussian mixture model is robust than other methods. The following were shown the resulting frame of the experiment:

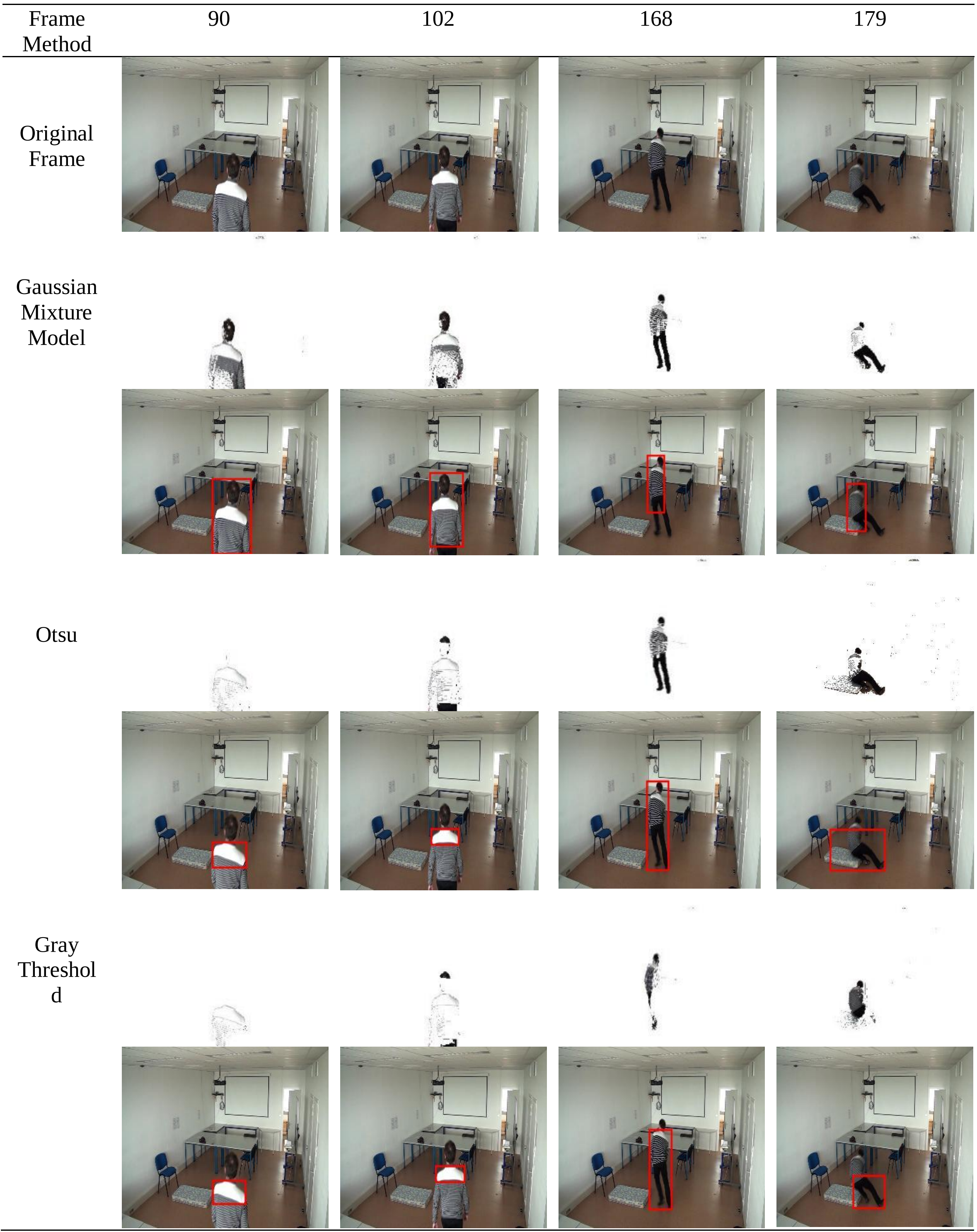

Figure 2. Result of the threshold and detection by gaussian mixture model, Otsu and gray threshold

3.1. Evaluation

In the experiments carried out to measure the performance of each adaptive threshold method that has been applied for the detection of moving objects, including the gaussian mixture model, the Otsu algorithm and the gray threshold, MSE and PSNR are used. Evaluation of the threshold results in the detection method is used peak signal to noise ratio (PSNR) [26] which is a method of comparing the maximal value of the stated signal with the quantity of noise intensity on the signal. Before determining the PSNR value, first, calculate the value [27]. Whereas mean square error (MSE) [28] is the average square error value between the original image and the improved image, which is formulated in the (14) dan (15).

$$MSE(F,X) = \frac{1}{KQ}\sum_{a=1}^{K}\sum_{b=1}^{Q}[F(a,b) - X(a,b)] \quad (14)$$

$$SNR(F,X) = 10.log_{10}\left(\frac{max^2}{MSE(F,X)}\right) \quad (15)$$

The variable F indicates that the frame ground truth, while the parameter X is a frame sequence that has a size of K x Q, the max value indicates the maximum amount of the pixel frame. The results of the mean square error and PSNR values of the gaussian mixture model and the Otsu technique can be seen in Table 1. From Table 1 shows that the MSE GMM value is lower than Otsu and gray, and the PSNR GMM shows higher than the Otsu and Gray Algorithm.

Table 1. Average performance evaluation adaptive threshold with MSE and PSNR (dB)

Evaluation	GMM Threshold	Otsu Threshold	Gray Threshold
MSE	257.18	595.36	645.39
PSNR	24.71	20.66	19.35

3.2. Experimental results

The result of thresholding in human moving in video dataset, Figure 3 and Figure 4 show the MSE and PSNR of adaptive threshold in human detection, successively. The diagram of result show the best accomplishment of GMM (MSE = 89.11, PSNR = 28.63) at frame 33. Averagely, MSE of GMM is 257.18, Otsu is 595.36 and gray threshold is 645,38. Also, PSNR of GMM is 24.71, Otsu is 20.66 and gray threshold is 19.36**.**

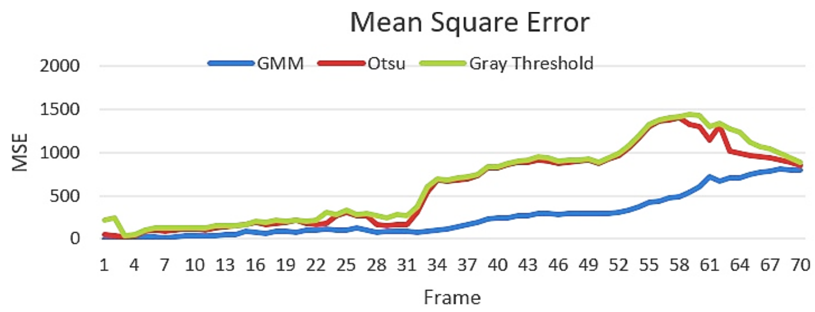

Figure 3. Performance MSE (dB) of adaptive threshold

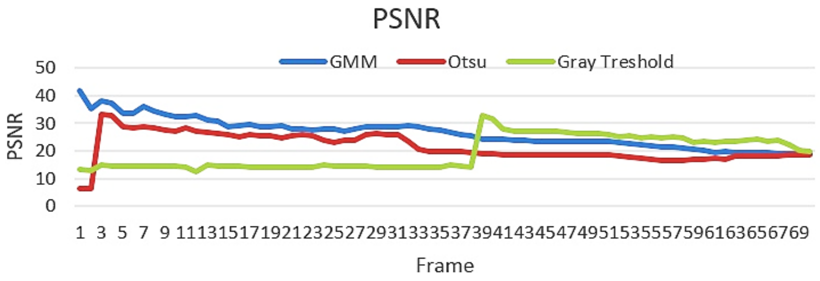

Figure 4. Performance PSNR (dB) of adaptive threshold

Based on the test results that the lower the value of mean square error (MSE), the thresholding process of finding objects in moving on the video frame the better the results or have a low error rate, whereas if the mark of the peak signal to noise ratio (PSNR) is higher the better the quality threshold moving objects in the video frame. Then it can be shown in Figure 3 graph that shows the success rate of the moving object threshold process.

3.3. Discussion

The possible reason Otsu threshold method is underperformed compared with GMM are the Otsu does not use spatial coherence and any object structure. Also, the Otsu threshold still use assumption in using binary class, which partition the grayscale histogram value into binary class. Due to dataset which can be seen at Figure 3, it shows that the video contains more than two segment class. Thus, in frame number 179, the Otsu threshold segment the bed and person into one due to color of bed and color of clothes have same gray color.

The possible reason gray threshold is underperformed compared with GMM are the Gray threshold failed to detect non-edge pixels. As we can see at Figure 3, the Gray only segment the pants due to it contains more value than the clothes which more same like the foreground. Although GMM possess same characteristic like Otsu threshold, which uses histogram-based techniques, but the GMM itself assumes the histogram distribution is represented by two gaussian curves, which able to finding a threshold and describe the two regions obtained.

4. CONCLUSION

This paper provided an adaptive threshold technology on assign the threshold of object moving for especially in human detection. The outcome showed that the achievement of an object moving in video detection using a gaussian mixture model is better than with standard otsu algorithm and gray threshold. The average of MSE values is 257.18 dB smaller than the other and the PSNR value is 24.71 dB higher than Otsu and gray threshold. Future work is an arrangement to evaluate other threshold techniques, to improve performance in threshold strategy.

ACKNOWLEDGEMENTS

This research was funded by a PDUPT Research Grant number 026/L6/AK/SP2H/RESEARCH/2019, from the Ministry of Research, Technology, and Higher Education, Republic of Indonesia.

REFERENCES

[1] M. A. Kandavalli, "A novel approach on object detection and tracking using adaptive background subtraction method," in *Second International Conference on Computing Methodologies and Communication (ICCMC)*, 2018.
[2] Xiaofeng Lu and Caidi Xu, "Novel Gaussian mixture model background subtraction method for detecting moving object," *International Conference of Safety produce Information (IICSPI)*, 2018.
[3] A. Rahman, M. Ahmed and Hosian, "An adaptive background modeling based on modified running Gaussian average method," *International Conference on Electrical, Computer and Communication Engineering (ECCE), Bangladesh*, 2017.
[4] Syed Tariq Ali, Kalpana Goyal and Jyoti Singhai, "Moving Object Detection Using Self Adaptive Gaussian Mixture Model for Real Time Applications," *Proceeding International conference on Recent Innovations is Signal Processing and Embedded Systems (RISE-2017)*, 2017.
[5] Z. Zeng, J. Jia and Y. Chen, "Pixel modeling using histograms based on fuzzy partitions for dynamic background subtraction," *IEEE Transaction In Fuzzy System*, vol. 25, no. 3, 2016.
[6] Otsu N, "A threshold selection method from gray-level histogram," *IEEE Transact. Man and Cybernet*, vol. 9, no. 1, pp. 62-66, 1979.
[7] M. Soeleman, Mauridhy Hery and Hariadi M, "Adaptive Threshold for Background Subtraction in Moving Object Detection using Fuzzy C-Means Clustering," *IEEE TENCON 2012*, Cebu, Philliphine, 2012.
[8] Yesong, XiaopingLi, Na Fu and Qiongxin Liu, "Fast Moving Object Detection Using Improved Gaussian Mixture Models," *ICALIP 2014*, 2014.
[9] Vivek Maik, Hyungtae Kim, Daedhee Kim, Eunjung Chae and Joonki Paik, "Robust Background Generation Using a Modified Mixture of Gaussian Model for Object Detection," *IEEE ISCE*, 2014.
[10] Xuehua Song, Jingzhu Chen, Chong He and Xiang Zhou, "A Robust Moving Objects Detection Based on Improved Gaussian Mixture Model," *International Conference on Artificial Intelligence and Computational Intelligence*, 2010.
[11] Ying Shi, Shu Cheng, Jie Chen and Di Chen, "Moving Objects Detection by Gaussian Mixture Model : a comparative analysis," 2011.
[12] Yi Zheng and Phing Zheng, "Hand Segmentation based on Improved Gaussian Mixture Model," *2015 International Conference on Computer Science and Applications*, 2015.

[13] Kittipop Peuwnuan , Kuntpong Woraratpanya and Kitsuc, "Modified adaptive thresholding using integral image," *13th International Joint Conference on Computer Science and Software Engineering (JCSSE),* 2016.
[14] Senthilkumaran and Vaithegi S, "Image Segmentation By Using Thresholding Techniques for Medical Images," *Computer Science & Engineering: An International Journal (CSEIJ),* vol. 6, no. 1, pp. 1-12, 2016.
[15] K. Srinivasan, K. Porkumaran and G. Sainarayanan, "Improved Background Subtraction Techniques for Security in Video Applications," *Proc. of 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication,* 2009.
[16] Marwa abd el Azeem Marzouk, "Modified background subtraction algorithm for motion detection in surveillance systems," *Journal of American Arabic Academy for Science and Technology*, vol. 1, no. 2, pp. 112-123, 2010.
[17] Y. Z and F. Liangzhong, "Moving object detection based on running average background and temporal difference," Hangzhou, 2010.
[18] M. A. Nur and A. Z. Mohd, "Moving object detection via TV-L1 optical flow in fall-down videos," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 3, p. 839-846, 2019.
[19] C. Stauffer and W. E. Grimson, "Adaptive background mixture models for real-time tracking," *IEEE Computer Society Conference on Computer Vision and Patern Recognition*, 2007.
[20] QIAO Baoming, Jing Liu and YANG Yufan, "An Adaptive Algorithm for Grey Image Edge Detection Based on Grey Correlation Analysis," *12th International Conference on Computational Intelligence and Security IEEE*, 2016.
[21] S. O. Imocha and S. Tejmani, "Local Contrast and Mean based Thresholding Technique in Image Binarization," *International Journal of Computer Applications*, vol. 51, no. 6, 2012.
[22] S. C. Nagaraju, G. Nagamani and P. G. Rakesh, "Morphological Edge Detection Algorithm Based on Multi-Structure Elements of Different Directions," *International Journal of Information and Communication Technology Research*, vol. 1, no. 1, 2012.
[23] A. Amer, "New binary morphological operations for efective low-cost boundary detection," *International Journal of Pattern Recognition and Artitcial Intelligence*, vol. 17, no. 2, 2002.
[24] K. Mahesh and K. Kuppusamy, "A New Hybrid Video Segmentation Algorithm using Fuzzy C Means Clustering," *International Journal of Computer Science*, vol. 2, p. 9, 2012.
[25] Charfi, J. MitÈran, J. Dubois, M. Atri, R. Tourki, "Optimised spatio-temporal descriptors for real-time fall detection: comparison of SVM and Adaboost based classificationî," *Journal of Electronic Imaging (JEI)*, vol. 22, no. 4, 2013.
[26] A. Z. Seyed Mohammad, Jun Zhou, Xin Liu and K. Tony Tae-Hyoung, "An Area- and Energy-Efficient FIFO Design Using Error-Reduced Data Compression and Near-Threshold Operation for Image/Video Applications," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 11, pp. 2408 - 2416, 2015.
[27] K. Bhoyar and O. Kakde, "Color image segmentation based on jnd color histogram," *Internationa Journal of Image Processing (IJIP)*, vol. 3, no. 6, pp. 282-293, 2010.
[28] Wang, Z, Bovik, A. C., Sheikh, H. R and Simocelli, "Image Quality Assessment: From Error Visibility to Structural Visibility," *IEEE Transactions on Image Processing*, vol. 4, pp. 600-612, 2004.