Adaptive Background Subtraction for Monitoring System

Afit Miranto
Department of Electrical Engineering
Faculty of Engineering
Universitas Lampung
Bandar Lampung, Indonesia
afit.miranto@gmail.com

Sri Ratna Sulistiyanti
Department of Electrical Engineering
Faculty of Engineering
Universitas Lampung
Bandar Lampung, Indonesia
sr_sulistiyanti@eng.unila.ac.id

F.X. Arinto Setyawan
Department of Electrical Engineering
Faculty of Engineering
Universitas Lampung
Bandar Lampung, Indonesia
fx.arinto@eng.unila.ac.id

Abstract— Security is one of the most important things for human life at this time. Complex activities often cause homes to be left unattended. One of the actions taken by most people to guard their homes while traveling is to use CCTV (Closed Circuit Television) cameras. This conventional CCTV is less effective because the camera is only recording without analyzing objects. From the shortcomings, the camera is made so that it can monitor the activities of the changes in the movement of objects seen by the camera, in this case the object detected is human movement. The monitoring system using this camera can detect passing objects. This paper proposes the adaptive background subtraction method needed to adapt to frame changes. The background frame will always be updated against the previous background intensity inference. Then it will analyze the effectiveness of the method. The effectiveness of the method used is then evaluated by comparing the results of object extraction with ground truth. The best success rate in object detection from object detection method is measured by calculating recall precision and F-measure values. The experimental results show satisfactory performance from the proposed method.

Keywords—Adaptive Background subtraction, Monitoring System, Object Detection

I. INTRODUCTION

Security is one of the most important things, especially in today's modern age. Often homes are left unattended because of someone's busy life. Theft at home due to lack of supervision caused many homeowners to lose their valuable items when leaving their homes. Therefore a good security system is needed so that theft can be detected quickly or even prevent theft.

According to data from the *Badan Pusat Statistik* (BPS) in 2018 [1], Lampung province ranks 3rd under DKI Jakarta and South Sumatra for villages where there was a violent crime with a percentage of 8.78%. Percentage of theft in villages in Lampung was 70.65%. with a level of concern when leaving home in an uninhabited state of 65.83%.

At present most people guard their homes by installing security systems based on CCTV (Closed Circuit Television). This security system is less effective because conventional CCTV only records events but does not explain what happened during the recording. The disadvantages of this security system can be overcome by adding a human presence detection system by modeling adaptive background subtraction. With this method the system will work by detecting the movement of an object in this case which is analyzed is human movement. Detecting this object is done in a real-time situation and applied to monitor a room/environment.

The problem with object detection is that there are many disturbances because the dynamic nature of the background has disturbances such as changing light intensity, and the movement of small objects that should not be considered as objects. This disorder can affect the results of identification so background modeling needs to be adapted to these disorders. So that it can separate the background and objects that will be detected correctly. Adaptive background works by taking into account the pixel values of all frames in sequence. However, it will be less effective if the detected input is too much and the actual interference is in the background, so a Gaussian Mixture Models (GMM) method is needed.

There are several methods used for object detection in previous studies. Among these methods are detection of optical flow detection, pattern matching, background subtraction and so on. Research on the detection of objects using the adaptive background subtraction method has been done before. Some previous studies that have been done include; Active Background Modeling Using the Gaussian Mixture Models [2]. Application of image processing using the adaptive motion detection algorithm method on security camera systems with push notification to android smartphones [3]. Implementation of Visual Object Tracking by Adaptive Background Subtraction on FPGA [4]. Object Tracking and Detecting Based on Adaptive Background Subtraction [5]. An Adaptive Background Subtraction Method Based on Kernel Density Estimation. Sensors [6]. A Novel Approach on Object Detection and Tracking Using Adaptive Background Subtraction Method [7]. Background Subtraction Algorithm Based Human Motion [8]. And, Moving detection employing objects iterative updates of the background [9].

Based on these previous studies, this study focuses on the detection system of moving objects that can be done in real time from the movement of objects. The method used is adaptive background subtraction with the Gaussian mixture model. this model is used to be able to update the background that can adapt to dynamic changes in conditions, namely the intensity of light or the movement of small objects that should be able to reduce or ignore. Then the adaptive system will always update the background model at any time. The proposed method can be applied to indoor and

outdoor automatic surveillance systems, where large dynamic changes can occur.

II. THE PROPOSED METHOD

Background subtraction is used to detect foreground objects by comparing two different frames and will find differences. Basically, this compares the difference value with the threshold value. The threshold value is not predetermined, but the threshold value is calculated using a number of the initial frames given. So if the difference is greater than the threshold value than it will be marked as an object, the opposite will be considered as a background image.

To determine the object with the background subtraction method, you can write with Equation 1 as follows:

$$x_t(s) = 1 \, if \, d(I_{s,t}, B_s) > \tau \, otherwise \, 0 \qquad (1)$$

Where τ is the threshold, xt is the field of motion labeling at time t (also called the motion mask), d is the distance between Is, t, color at time t and pixel s, and Bs is the background model of pixel s [8].

In order for the background model to adapt to changes in light better, the background needs to be updated every time as in Figure 1. so that it can extract moving objects more accurately. Background update algorithm according to Paul is written with the following equation 2 [3]:

$$B(x,y,t) = \alpha, f(x,y,t) - (1-\alpha)B(x,y,t-1) \qquad (2)$$

where $\alpha \in (0,1)$ is the update coefficient, f (x, y, t) is the gray value of pixels in the current frame. B (x, y, t) and B (x, y, t-1) are the background values of the current frame and the previous frame. The background model can remain relatively stable for a long time. Using this method can effectively avoid unexpected symptoms from the background, such as the appearance of something sudden in the background that is not included in the original background. In addition, by updating the gray pixel values of the background, the effects brought by light, weather and other changes in the external environment can be effectively adjusted. In detecting a moving object, pixels that are valued as belonging to a moving object retain the original background gray value, not updated [4].

In this study the video was taken using a raspberry pi night vision camera that can be used in conditions both day and night. In a security camera system designed, in this system shooting and determining objects and backgrounds will be difficult due to changes in the intensity of light that occurs. So it is less effective when using conventional background subtraction methods for object detection. The right method is needed to model the background adaptively to changes in pixel intensity. So that object extraction will get better. The algorithm flow diagram in this study is as shown in Figure 1.
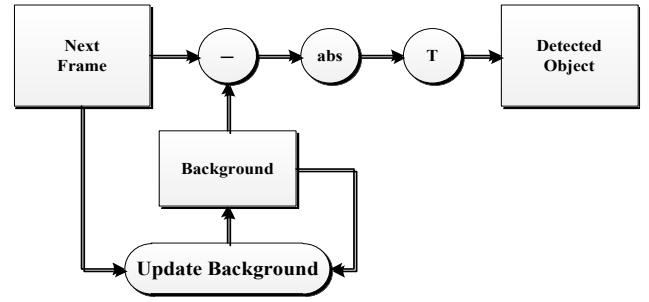

Figure 1 Block Diagram object detection using adaptive background subtraction

From Figure 1 it can be explained that the initial frame is used as a comparison for the next frame and will then be subtracted against the new frame.

The initial processing of object detection is to convert RGB images taken previously to gray scale images using the following Equation 3 [9].

$$I = 0.2989R + 0.5870G + 0.1141B \qquad (3)$$

Where I is the intensity of gray scale, R, G and B is the intensity of red, green and blue. The video image is taken, namely during the daytime condition.

The purpose of conversion to gray scale is to reduce the computational burden so that the time needed to execute one frame is faster. RGB images consist of 8bit each channel (3 channels / 24bit) while gray scale images only have 1 channel (8 bits).

After the image has been converted to gray scale. The next step the image frame will be extracted so that it will separate the object from the background. Each pixel in the frame will be specified as an object or as a background using Equation 4 below. This equation is used to determine the condition of each pixel in the latest frame [12].

$$f(x,y) = \begin{cases} background & \dfrac{|f(x,y) - \mu(x,y)|}{\sigma} \le T \\ object & \dfrac{|f(x,y) - \mu(x,y)|}{\sigma} > T \end{cases} \qquad (4)$$

f (x, y) is the intensity of the current frame in the pixel position (x, y). μ (x, y) is the intensity of the background frame at the pixel position (x, y) and T is the threshold value. In the initial frame all pixels are expressed as background so that f (x, y) - μ (x, y) = 0. The background pixel is black and the pixels of the detected object are white.

This background model is obtained from the average value of renewal. The mean and variant values can be stated when there is a new frame entered. Renewing the mean is obtained by using Equation 5 below [9],[10].

$$\mu_{T+1_{(x,y)}} = \begin{cases} \alpha * f_{t+1_{(x,y)}} + (1-\alpha) * \mu_{T+1_{(x,y)}} \\ \quad if \, f_{T+1_{(x,y)}} = background \\ (1-\beta) * f_{t+1_{(x,y)}} + \beta * \mu_{T+1_{(x,y)}} \\ \quad if \, f_{t+1_{(x,y)}} = not \, background \end{cases}$$

$$(5)$$

Whereas the variant update uses Equation 6 below.

$$\sigma^2_{T+1_{(x,y)}} = \begin{cases} \alpha * \left(f_{T+1_{(x,y)}} - \mu_{T+1_{(x,y)}}\right)^2 + (1-\alpha) * \sigma^2_{T+1_{(x,y)}}, \\ \qquad if \ f_{T+1_{(x,y)}} = background \\ (1-\beta) * \left(f_{T+1_{(x,y)}} - \mu_{T+1_{(x,y)}}\right)^2 + \beta * \sigma^2_{T+1_{(x,y)}}, \\ \qquad if \ f_{t+1_{(x,y)}} = not \ background \end{cases} \quad (6)$$

Where μ is the average pixel intensity, $\sigma2$ is a variant or a specified threshold value. α is the constant update background. And β is the renewal constant of the threshold value. In this study the values of α and β are constant values with certain values. Values α and β are included in several values to find the best conditions for each condition that will be used in the object detection process [10],[11],[12].

III. EXPERIMENTAL RESULT

In this study video data has been taken using a statically mounted camera. Specifications used are: raspberry pi 3 B mini PC Broadcom Quad Core 1.2GHz BCM2837 64bit CPU, 1GB RAM, using Camera: 5MP Night Vision camera with dual IR LED.

Tests carried out are in bright daylight conditions. Video is captured with a size of 640x480 pixels at a speed of 25fps. the video is captured for 10 seconds so that it will produce an image of 250 frames per 10 seconds of video. then the video is extracted into an image to a smaller size of 300x228 pixels. Next, the initial video on object detection is to convert RGB images into gray scale images using Equation (3) and produce an image as shown in Figure 1 below. (a) is the original RBG image before the change is made while the right (b) is the converted image from the color image to the gray scale image.

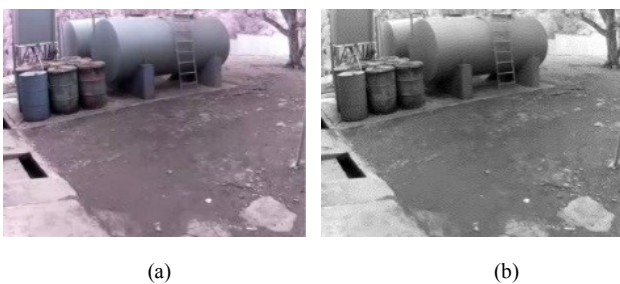(a)   (b)

Figure 2 Real color picture and gray scale picture

The next step is to extract the object against the background. Each pixel in the frame will be specified as an object or as a background. In the initial frame all pixels are declared as background. The background pixel is black and the pixels of the detected object are white. Background and object models obtained by using equation 5 and equation 6 by considering the threshold value given.

Threshold is very influential on the results of testing because the threshold will relate to sensitivity in detecting objects. The lower the threshold value the higher the sensitivity, so that only a slight change in the background will take effect. And the lower the threshold value, the lower the sensitivity of the camera to detect objects.

To evaluate the effectiveness of the algorithms used in this study. By comparing the results of extraction of objects with ground truth from manually created videos. This comparison will result in true positive (TP), false positive (FP) and false negative (FN) regions.

TP is an area that is detected as an object using the method used, and in reality it is true as the desired object. TN is an area that is detected as a background which in fact is indeed a background. FP is an area that is detected as an object but in reality it is background. While FN is an area that should be an object but is not detected as an object. A good algorithm is to have small FP and FN values.

The success in object detection is very dependent on the level of precision of the object detection method used so that measured in this study there are important parameters to analyze, namely using parameter recall or so-called sensitivity expressed in Equation 7, precision or positive estimates expressed in Equations 8 and the F-measure stated in Equation 9 below [13].

$$recall = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$precision = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

$$F = 2 \frac{recall \times precision}{(recall + precision)} \quad (9)$$

From the equation TP is the number of pixels in the true positive area, FN is the number of false negative pixels, and FP is the number of false positive pixels [9].

The following are the results of detection of videos of people walking. In figure 3 (a) is the original image, (b) is the result of the detection object and (c) is the result of the evaluation presented in different color gradations. For red is true positive (TP), white is true negative (TN), blue is false negative (FN) and green is false positive (FP). From the picture, it can be seen that the object detected is quite good even though there is still a little noise that occurs, there is still a background detected as an object, so the precision is slightly low.

Frame 26   Frame 26   Frame 26

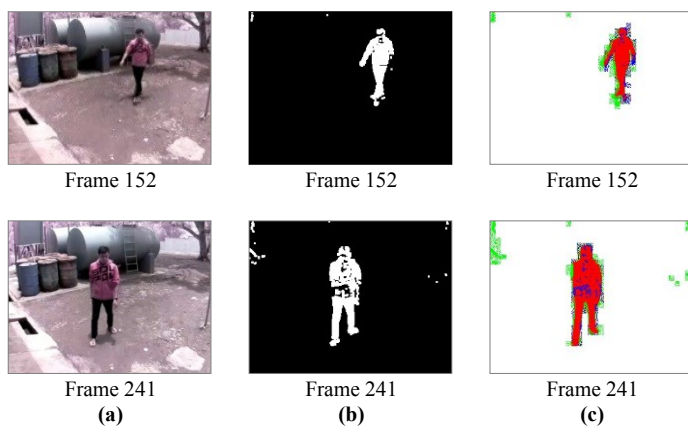Frame 72   Frame 72   Frame 72
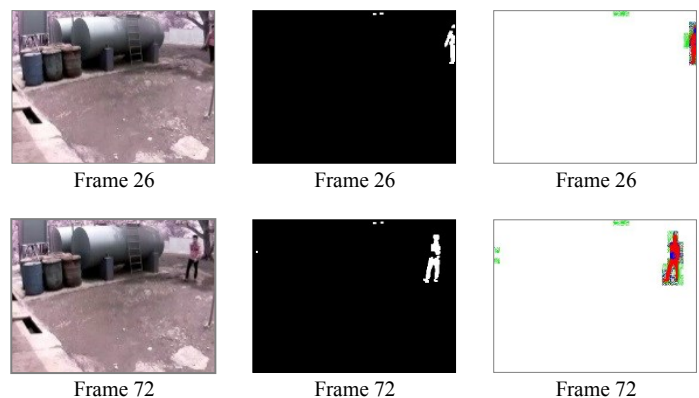
Figure 3. Experiment result (a) original image, (b) detected moving object (c) result of evaluation

From Figure 3 the experiment result. From the tests carried out based on frame 26 using Equations 7, 8 and 9, the Recall value is 73.89%, Precision 56.22% and F 63.84. In testing frame 72 the value of Recall was 70.48%, Precision 61.56% and F value 66.72. While in frame 152 testing, the value of Recall was 81.05%, Precision 68.05% and F value of 72.42. while in frame 241 testing, the value of Recall was 85.25%, Precision 71.95% and F value 78.04. For other frames using the equation done for other frames in one video the same can be explained in the following graph.
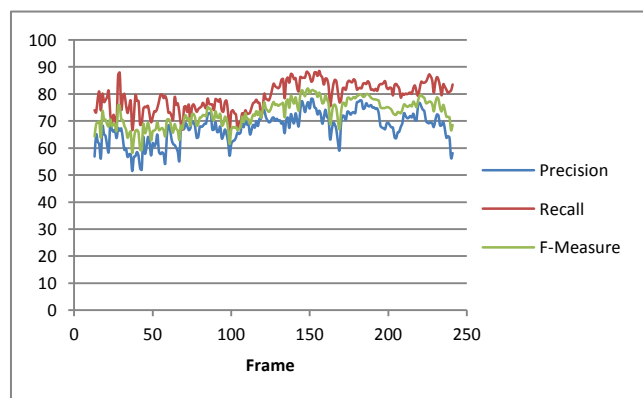

Figure 4 graphics recall, precision and F score all frame

From Figure 4 above it can be seen that for each frame in 1 video it is quite good, above 60% for each recall, precision and f measure values. This indicates that the method used is effective enough to detect objects.

IV. Conclusion

Based on the results of the analysis and discussion it can be concluded that the proposed detection method is based on the reduction of background adaptive. Able to detect moving objects well and can adapt to gradual changes. In modeling the background, the proposed method can adapt the background changes well during the day which changes the intensity of the light.

The renewal effectiveness of the method used for the experiment is also quite good, which is shown in Figure 4, the average recall value, precision and f measure above 60%. Adaptive background reduction method is very good for detecting moving objects. Applications can be used for monitoring systems.

Further improvements need to be made to improve the accuracy of object detection that can be developed to be able to detect objects in different conditions, for example during heavy rain conditions or when the weather changes or when there are sudden changes in lighting such as lights. a brief outage. Need to add other algorithms to improve the effectiveness of detection results such as filter morphology and so on.

References

[1] BPS. "*Statistik Kriminal 2018*". Badan Pusat Statistik. 2018.

[2] Nurul, Hudayani., Sulistiyanti, Sri Ratna,. Setyawan, F.X. Arinto. "Pemodelan Latar Belakang Adaptif Menggunakan Metode Gaussian Mixture Model pada Video Dalam Air*". Electrician. Volume 12, No.2. 2018

[3] Antonius, Alvin. Dedi Triyanto, Ikhwan Ruslianto. "Penerapan pengolahan citra dengan metode adaptive motion detection algorithm pada system kamera keamanan dengan push notification ke smartphone android". Jurnal Coding Sistem Komputer Untan. Volume 03, No. 2. 2015.

[4] Paul, Denna dan Ranjini Surendran. "Implementation of Visual Object Tracking by Adaptive Background Subtraction on FPGA". IJERT. 2015.

[5] Lee, Jeisung dan Mignon Park. "An Adaptive Background Subtraction Method Based on Kernel Density Estimation". Sensors. 2012.

[6] Angelo, Kandavalli Michael. "A Novel Approach on Object Detection and Tracking Using Adaptive Background Subtraction Method". IEEE Xplore. 2018.

[7] Rupali S.Rakibe, dan Bharati, D.Patil. "Background Subtraction Algorithm Based Human Motion Detection. International Journal of Scientific and Research Publications". Volume 3, Issue 5, May 2013

[8] Desai, Helly M dan Vaibhav Gandhi. "A Survey: Background Subtraction Techniques". International Journal of Scientific & Engineering Research. 2014.

[9] Setyawan, F.X. Arinto,. Tan, Joo Kooi,. Kim, Hyoungseop,. Ishikawa, Seiji.. "Moving objects detection employing iterative update of the background". Springer. 2016

[10] Setyawan, F.X. Arinto,. Sulistiyanti, Sri Ratna. "Extraction of Moving Objects on Underwater Video Using Method of Subtraction the Background Modeling Results". Electrician. Volume 11, No.2. 2017.

[11] Sulistiyanti, S.R., Setyawan, FX Arinto., Komarudin, Muhamad. 2016. Pengolahan Citra Dasar dan Contoh Penerapannya, Teknosain.

[12] Prabowo, MR,. Hudayani, N., Purwiyanti, S., Sulistiyanti,S.R., Setyawan, F.X.Arinto. "A Moving Object Detection In Underwater Video Using Substraction Of The Background Model". Proc. EECSI. 2017.

[13] R, Baeza-Yates., B Ribeiro-Neto. "Modern Information Retrieval".. ACM Press Addison Wesley. 1999.