Motion detect application with frame difference method on a surveillance camera

A M Husein, Calvin, David Halim, Raymond Leo, William
Faculty of Technology and Computer Science, Universitas Prima Indonesia, Medan, Sumatera Utara 20111, Indonesia

* amirmahmud@unprimdn.ac.id

Abstract. Security becomes one of the major necessities in our lives nowadays however criminal activities are still at large with criminals unable to be persecuted without eligible proofs of their misdeeds. Surveillance Camera is one of the better solutions to these problems in which they can be positioned at every corner of a building even streets and alleys. Their functions can be enhanced by adding algorithms that can identify objects. Frame Differences method is an algorithm to identify an object’s motion. Using this algorithm, we could differentiate an object moving in the environment. Background subtraction is one of the methods suitable to further improve frame differences thus increasing its effectiveness and precision. After implementing the method on a camera, the luminosity was founded to influence the threshold value significantly, the threshold value of 35 is the optimal value.

1. Introduction
Surveillance camera or better known as CCTV (Close-circuit television) becomes one of the essential equipment for security in a home or working environment especially on a big company. CCTV provides surveillance and monitoring every event that happened in their field of view as a video recording/live cam feed. CCTV on average has the effective field of view that ranges up to 30 ~ 50 meters top in which a person or a car license plate can still be identified with. These range depends on the lens that the camera was equipped with, the resolution in which the video is recorded and the transmission speed of the camera that is measured with fps (frames per second). Another way to improve the effectiveness of a surveillance camera is to add some features like a face recognizing app, object segmentation and even a motion detection program. On the other side, tracking systems has been developed and further improved to monitor restricted areas or endangered species on a specific nature conservation area. However, these systems were designed to monitor on a limited range or an open space like a savannah [1].

Video surveillance has received a great attention as active application-oriented research areas in image processing computer vision, artificial intelligence. The process of video surveillance aims at analyzing video sequences. Video surveillance activities can be manual, semi-autonomous or fully-autonomous. Manual video surveillance involves analysis of the video content by a human. Semi-autonomous video surveillance involves some form of video processing but with significant human intervention. In a fully-autonomous system only input
is the video sequence taken at the scene where surveillance is performed. In such system there is intervention of human and the system does both tasks, like motion detection and tracking [2], and decision making tasks like abnormal event detection and gesture recognition. The surveillance system starts with motion and object detection.

When detecting a motion on a video feed, two frames of the video will be compared to find the differences in the pixel values, this differences in value will determine the object’s motion. When this motion values can be found then the program can interpret these values to draw out the object’s shape. There are a few relevant conditions that became an important factor to determine the results of this experiment which are;

a. Image noise, due to the camera’s quality source.

b. Gradual changes in the background’s lighting conditions such as flickering street lights.

c. Dynamic movements of static objects such as rustling trees and bushes.

d. Shadows projected by foreground objects that can be detected as moving objects.

These factors above affect the pixel values spontaneously on the video thus potentially change the experiments result and the program to falsely recognize said changes as the motion of an unidentified object. Nowadays, present research on image processing focus solely on developing an algorithm that can detect moving object on specified range for object tracking application [3]. Many research on motion detection uses inter-frame difference, optical flow and background subtraction as the additional method to improve and accompany the frame differences method [4]. However, methods like background subtraction has a few limitations and weakness such as losing their effectiveness when used on an environment with unstable lighting and shadow projected by foreground objects [5]. Sum of Absolute Difference (SAD) can be used as an addition to get the pixel values from the frame that was compared [6]. Frame difference method can also be improved using correlation coefficient while classifying the background in groups of pixel blocks [7]. By utilizing the combination of frame differences and background subtraction to do motion detect, a device mounted with RF camera can be made into a Video Motion Detection Security Systems (VDMSS) that proved to be cheaper than the average VDMSS market price [8]. The compression rate of videos that uses a background subtraction features can be improved by utilizing DirectDraw technology [9]. Human detection also has a few restrictions in their solution such as moving objects, simple background and higher image resolution [10]. The accuracy of detecting moving object is also affected by the presence of noise created by detection methods that was used [11]. Background subtraction needs robust initial background [12], a selection of detection threshold [13], and a set of foreground, background and prior frame likelihood [14] to function effectively.

Frame differences is a method that is commonly used to detect an object through its motions. Therefore, the main objective of this paper is to propose a motion detection method that could be used with a live cam input from a video camera. The method is used to improve the effectiveness of security cameras like those used on Smarthome and CCTVs. The method that will be used are frame differences with the addition of background subtraction. These methods were chosen due to their high precision by comparing the amounts of pixels on each frame. The result of this research is hoped to be useful as a further reference on readers on the topic of motion detection.

2. Methodology

2.1. Frame Difference
Frame difference method is used to detect every motion that an object make that was captured by the camera. The frame difference algorithm takes every pixel within 2 frames to be compared sequentially and adds their differences on that block. This difference then was intended to be shown as a “motion” that resulted from a moving object that was caught by the camera [1]. On the first equation (1) shows the differential of pixel values, with $\Delta n$ as the differential value on the $n^{th}$ frame and $I_n$ as the pixel intensity on the $n^{th}$ frame.

$$\Delta n = |I_n - I_{n-1}| \tag{1}$$

After the value of $\Delta n$ is obtained, the motion of the object can be calculated by comparing the value of $\Delta n$ with a threshold that has been stated. The value of the threshold is usually within 15% of the range that was used as the observed pixel intensity. So, if the range consists within 0-255 then, the threshold that will be used is rounded up to 40 [15]. Motion depicted as $(Mn)$ can then be calculated by doing the operation per pixel with this equation:

$$Mn = \begin{cases} 1 & \text{if } \Delta n \geq T \\ 0 & \text{if } \Delta n < T \end{cases} \tag{2}$$

This method can be improved by using more than two frames to compare or by using a threshold value that is adaptive. The frame difference method can be simplified into a few steps where video inputs is firstly collected and converted to acquire frames that will be compared, a binarization process was later performed in which the algorithm does to create pixel values that represents the motion detected on each frame. The result of this binarization would be a black and white picture frames with white pixels depicting the motion that was captured.

2.2. Background Subtraction

Background subtraction method models the background thoroughly from pixels by using different filters such as running average / approximate median filter or temporal median filter. The main objective of this method is to make a model of the background to be a reference on foreground detection [15]. If the foreground on the $t^{th}$ frame is shown as $F_t$, and pixel intensity and background values is depicted as $I_t$ and $B_t$, then the value of foreground can be calculated using this equation:

$$F_t = \begin{cases} 1, & |I_t - B_t| > T \\ 0, & \text{else} \end{cases} \tag{3}$$

Background subtraction is usually used to detect motion on a static camera. There are a few advantages on using this method which is: easy to be implemented; quick; and precise. While it is quick and precise, background subtraction is keen to be affected from background changes that includes the lighting condition dynamic movements of objects in the background like wind blowing to trees and waving flags in the distance [16]. The background subtraction method can be simplified into few steps.

Firstly, background subtraction method takes the first frame captured as a background image. This background image will be used as a reference to be compared with incoming frames captured by the camera. If the difference in value exceeds the bound threshold then that pixel will be treated as a part of the pixels on the moving object or a background pixel. The threshold value is important because if the value is too small, then it will produce a lot of false change points and if the value is too large, it will decrease the scope of changes in movement [16].
2.3. Process Analysis
The process of the application will be as follows; (1) the application will firstly take the first frame as a background reference; (2) it will then do a comparison between the frame that is captured with the background reference; (3) If the difference in the pixel value exceeded the threshold then the application will display it as a motion. This process can be viewed as a flowchart as shown in Figure 1 below.

![Application flowchart](image)

**Figure 1.** Application flowchart

2.4. Diagram
The designed application is modeled with an activity diagram. In the activity diagram, it is explained that (1) Initial Node, an object that begins; (2) Action state that shows the action that is performed such as: process, detect motion, and shows output; (3) 1 Decision Node, that select the decision available. In this case, when the camera detected a motion, it will display a white binary output on the pixels of that object and shows black binary output when there’s no motion to detect; (4) 1 Final State, as the object to be terminated. As shown as a use case diagram in Figure 2. below.
2.5. Motion Testing

The experiment is performed with a human as the object in three scenarios with the range between the object and the camera lens are 1.5 meters apart. During its experiment, the camera that was used in this experiment is an external webcam with the display resolution of 640 x 480 pixel with the transmission speed of 10 frames per second. The application will display a video with the resolution of 640 x 480 and displayed RGB real time video with the motion detect frame besides it. The experiment is performed with these options:

- The first test is carried by doing a background shot with nothing in front of the camera’s field of view.
- The second test is performed by doing a recording of the object standing motion less in front of the camera.
- The third test is performed where the object would move across the camera’s field of view.

The application can detect the object’s motion quite well. The result was displayed in a binarized image that shows the movement outline of the object. While not perfect, the application is able to outlines the object’s motion. In the second test at a brighter location, the application is able to outline the object even though it’s motionless. This was caused by the lighting position behind the object that causes the shadow to interfere with the application thus recognizing the object as moving as shown in Figure 3 and Figure 4 as shown below.
3. Implementation

There are few factors that concern the motion detect result from surveillance camera. The factors are resolution, determine picture reference technique, colour components and threshold value. In this
section, the analysis had been done based on picture resolution with frame differences method. Extensions that used are RGB, Grayscale and YCbCr. The second analysis is make a decision about the colour components that caught by surveillance camera. The analysis equals to the best and adaptive to any condition to determinate thresholds.

3.1. Effect of Threshold and Resolution
This invention is to compare the corelation coefficient between threshold with various resolution. Standard picture caught by surveillance camera is 640×480 pixels. Comparison with 5 pictures resolution which is 640×480 pixels, 512×384 pixels, 384×288 pixels, 256×192 pixels and 128×96 pixels. The value of threshold used is between 5 – 80 with multiply by 5. This corelation testing done by using analysis corelation product moment pearson. Because data type that used is interval. The result of this invention can be see in Table 1.

| Coefficient | Effect          |
|-------------|----------------|
| -1.0 to -0.5 or 0.5 to 1.0 | Significant |
| -0.5 to -0.3 or 0.3 to 0.5  | Quite Significant |
| -0.3 to -0.1 or 0.1 to 0.3  | Low           |
| -0.1 to 0.1               | Very Low      |

Tabel 2 describes that coefficient values between 0.75 and 0.78. The value means threshold have the significant effect to comparison picture for various sizes. If the value of coefficient corelation negative means if the value of the threshold become smaller then the difference pixels that caught will become larger. If the difference of pixel value caught by camera become larger, then the noise from comparison become larger too. Beside that, according to corelation coefficient can be concluded that using picture resolution doesn’t effect the result of picture comparison. That means comparison can be done with using various picture resolution. Smaller size of the picture more better than the larger one, because the process of iteration equation that done will become fewer. But the ideal picture that used must be noticed. Because it relate to information of picture that used, therefore hopefully the lower resolution of the picture can be used by user to monitoring the field.

| No. | Picture Size (pixel) | Picture Type  | Coefficient Correlation RGB | Coefficient Correlation Grayscale |
|-----|----------------------|---------------|-----------------------------|----------------------------------|
| 1   | 640×480              |               | -0.7786                     | -0.7876                          |
| 2   | 512×384              |               | -0.7798                     | -0.7713                          |
| 3   | 384×288              |               | -0.7779                     | -0.7695                          |
| 4   | 256×192              |               | -0.7732                     | -0.7652                          |
| 5   | 128×96               |               | -0.7653                     | -0.7577                          |
3.2. Detection with RGB

After done testing, picture resolution used are 256×192 pixels. Doing testing to compare detection using picture with RGB type, average RGB. This invention using 10 data samples that caught in different location and condition. Reference picture that used is picture that caught at time t-1. Other than that, also set adaptive threshold with all condition that caught by surveillance camera. Comparison done with analyze percentage of comparison results from 10 data samples. Comparison results analyzed to give information about the result status. Status is detected or not. Result of the comparing detection using RGB can be see in table 2.

According to the comparison, various better threshold got for any method in frame differences using RGB. The best threshold can be see in table 3. According to the table, the lower percentage cause by threshold 5 happen cause by higher noise. Therefore sensitivity detection become higher. Sometimes algorithm to detect the motion in the picture, but reality it doesn’t happen. This could happen because lightning effect that caught in surveillance camera.

| Table 3. Detection using various threshold |
|------------------------------------------|
| Threshold  | Detected percentage |
|            | RGB (%) | RGB Mean (%) |
| 5          | 0       | 6            |
| 10         | 28      | 64           |
| 15         | 72      | 86           |
| 20         | 86      | 94           |
| 25         | 90      | 94           |
| 30         | 92      | 94           |
| 35         | 96      | 98           |
| 40         | 92      | 90           |
| 45         | 90      | 92           |
| 50         | 88      | 88           |
| 55         | 88      | 88           |
| 60         | 88      | 86           |
| 65         | 84      | 86           |
| 70         | 80      | 78           |
| 75         | 80      | 78           |
| 80         | 72      | 70           |

4. Conclusion

Frame differences is one of the most common method to be used to detect an object’s motion. This method is flexible as in capable to be modified and adjusted to match the system requirements. Background subtraction method can reduce a lot of noise created by static non-moving objects in the background. Nevertheless, the noise that was shown in Figure 3. And Figure 4. Was caused by the position of the camera and the traits of background subtraction method itself that is quite sensitive.
towards changes in illumination of the background. Suggestions upon improving the effectiveness of these motion detecting cameras are to add a function that allows the camera to perform on a low-light condition. Based on the tests that has been performed, the threshold value of 35 is the optimal value with 96% on RGB and 98% on RGB Mean

References

[1] Son. Byung-rak, Shin. Seung-chan, Kim. Jung-gyu, and Her. Yong-sork 2007 Implementation of the Real-Time People Counting System using Wireless Sensor Networks International Journal of Multimedia and Ubiquitous Engineering Volume 2 No.3.

[2] Mishra K.Sapana, K.S.Bhagat 2015 A Survey on Human Motion Detection and Surveillance International Journal of Advanced Research in Electronics and Communication Engineering Volume 4 No.4

[3] N. Singla 2014 Motion Detection Based on Frame Difference Method International Journal of Information & Computation Technology Volume 4 Number 15

[4] J. Guo, J. Wang, R. Bang, Y. Zhang, Y. Li 2017 A New Moving Object Detection Method Based on Frame-difference and Background Subtraction IOP Conference Series: Materials Science and Engineering Volume 242

[5] K. Kavitha, A.Tejaswini 2012 Background Detection and Subtraction for Image Sequences in Video International Journal of Computer Science and Information Technologies Volume 3 Issue 5 pp.53223-5226

[6] H. H. Kenchannavar, G. S. Patkar and U. P. Kulkarni 2010 Simulink Model for Frame Difference and Background Subtraction Comparison in Visual Sensor Network The 3rd International Conference on Machine Vision (ICMV 2010)

[7] Y. G. Wu, C. Z. Qiang, G. Jian, X. Dan, L. R. Jian, H. J. Yuan 2014 A Moving Object Detection Algorithm Based on a Combination of Optical Flow and Three-Frame Difference. Information Technology Journal Volume 13 (11) : 1863-67

[8] S. Mishra, P. Mishra, N. K. Chaudhary, P. Asthana April 2011 A Novel Comprehensive Method for Real Time Video Motion Detection Surveillance. International Journal of Scientific & Engineering Research Volume 2, Issue 4

[9] Z. Wang, Y. Zhao, J. Zhang, Y Guo October 2010 Research on Motion Detection of Video Surveillance System International Congress on Image and Signal Processing Volume 1 pp. 13-197

[10] Hou. Ya-Li and K.H.Pang. Grantham 2011 People Counting and Human Detection in a Challenging Situation IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans Volume 4 No.1

[11] Yanzu Zhang, Xiaoyan Wang, Biao Qu December 2012 Three-Frame Difference Algorithm Research Based on Mathematical Morphology Procedia Engineering Volume 29 pp.2705-09

[12] Jun Yang, Tusheng Lin, Bi Li 2011 Dual Frame Differences Based Background Extraction Algorithm ICCP2011 Proceedings pp. 44-47

[13] J. Mike McHugh, Janusz Konrad, Ventakesh Saligrama, Pierre-Marc Jodoin June 2009 Foreground-Adaptive Background Subtraction IEEE Signal Processing Letters Volume 16 No.5 pp.390-93
[14] Manjunath Narayana, Allen Hanson, Erik G. Learned-Miller July 2013 Background Subtraction – Separating the Modeling and the Interference *Machine Vision and Applications* Volume 25 No.5

[15] Luqman A. Mushawwir, Iping Supriana 2015 Deteksi dan tracking Objek untuk Sistem Pengawasan Citra Bergerak *Konferensi Nasional Informatika*

[16] D. Stalin Alex, Dr. Amitabh Wahi 28th February 2014 BFSD : Background Subtraction Frame Difference Algorithm for Moving Object Detection and Extraction *Journal of Theoretical and Applied Information Technology* Volume 60 No. 3

[17] N. K. Patil, R. M. Yadahali and J. Pujari, “Comparison between HSV and YCbCr Color Model Color-Texture based Classification of the Food Grains,” *International Journal of Computer Applications* (0975 – 8887) Volume 34– No.4, 2011