# A MOVING OBJECTS DETECTION IN UNDERWATER VIDEO USING SUBTRACTION OF THE BACKGROUND MODEL

M. R. Prabowo, N. Hudayani, S. Purwiyanti, S. R. Sulistiyanti, F. X. A. Setyawan
Department of Electrical Engineering, Faculty of Engineering
University of Lampung
Bandar Lampung, Indonesia
mrprabowo9@gmail.com, nurul.hudayani@students.unila.ac.id, sr_sulistiyanti@eng.unila.ac.id, fx.arinto@eng.unila.ac.id

*Abstract*—This paper proposes a method for detecting moving objects on an underwater video. Video obtained using an underwater camera to capture the environmental conditions of the area. This research is the initial stage of the underwater surveillance system. Underwater surveillance system enables objects passing can be recognized shapes, types, and its behavior. The detection method used in this research is a subtraction between the current frames with the background modeling results. Underwater video retrieval has a high level of difficulty because the background is always changing either due to a change the intensity and the movement of water currents. Therefore, it needs to be made an appropriate background model to address this problem. Modeling of the background on this research using adaptive modeling method, where the intensity of the background pixels is updated based on inference of the background intensity before. If the intensity of the pixels changed drastically beyond the allowed threshold value, the pixel is considered as the pixels of the object and the pixel values of the background model are updated based on this pixel value. The effectiveness of the proposed method is expressed with the value of recall and precision. The average recall value of the two videos is 83% and the value of its precision is 67.5%.

***Keywords—adaptive modeling; background modeling; detection object; underwater surveillance***

## I. INTRODUCTION

One target of the maritime field is increasing the fish production. Efforts are underway to increase fish production is the use of fishing technology. One of the technologies used in fish catching is the utilization of fish caller submersible lamp. This research proposes the use of cameras to help determine the number of fish around a fish caller submersible lamp.

The rapid development of electronic technology can be used to support increased fisheries productions. Underwater camera technology can be mounted on a fish caller lamp in order to replace the human eye in overseeing the object around the lamp. Currently, the technology of fish caller lamp that used by fisherman only to call the fish to come closer and did not use other technologies to see if the fish are coming has been quite a lot. Therefore, it is often the catch of fishermen using fish caller lamp less than optimal.

Much research has been done previously regarding the object detection and use of underwater cameras. The simplest technique to detect an object is using the method of subtraction of the background [1]. In this research is used fixed background and is used on outdoor, while this proposed method using a background modeling and used on the underwater video. The use of a method of the background subtraction is more suitable for a static background. For the dynamic background, the result of a subtraction of the background is not good because of intensity a change in the background is detected as an object. The research on object detection with dynamic backgrounds also has been a lot to do [2,3,4,5].

The research on background modeling has also been done [6,7,8]. The difference with the proposed method is the overall parameter of the modeling using adaptive parameter while in the proposed method, partly parameters using constants in order to reduce the computational load. In the proposed method, updates of the background were done on the gray level image in order to reduce the computational load.

Much the theme of research about the object detection on video of the underwater has been done [9,10,11,12]. In the proposed method, the detection is done by subtracting the current frame with a frame of the background model which has been formed from the previous frames. In this research used a static camera to take underwater video.

## II. THE PROPOSED METHOD

### A. Preprocessing

The video data taken by an action camera that installed on a tripod is put under water. In this research used four video data taken in the morning (video 1 and 3) and afternoon (video 2 and 4) with the water conditions are not a quiet. The image preprocessing step is done by changing the RGB image into a gray level image. This conversion is done by using (1).

$$I = 0.2989R + 0.5870G + 0.1141B \qquad (1)$$

Here, I is the intensity of grayscale, R, G, and B is the intensity of red, green, and blue respectively.

### B. Initialization normal distribution

The next step is to determine the normal distribution model of each pixel in the image. This proposed method uses the first frame of the video as the initial background. The intensity of each pixel becomes an initial mean value and the value of the initial variant specified with a value of 1.

### C. Decision the pixels as a background or objects

After the first background model is obtained, then the next step is to determine each pixel in the next frame whether as the background or object. Equation (2) is used to determine the condition of each pixel on the current frame.

$$f(x,y) = \begin{cases} background & \frac{|f(x,y)-\mu(x,y)|}{\sigma} \le T \\ object & \frac{|f(x,y)-\mu(x,y)|}{\sigma} > T \end{cases} \qquad (2)$$

Here $f(x,y)$ is an intensity of a current frame in $(x,y)$ position pixels, $\mu(x,y)$ is an intensity of a background model frame in $(x,y)$ position pixels and $T$ is a threshold value. On the first frame, all pixels detected as a background is caused due to the first frame is regarded as background so $f(x,y) - \mu(x,y) = 0$. All the pixels in the frame checked to determine as the background or object.

The pixel determination results on frame 28, 35, 42, and 53 from video 3 are shown in Fig. 1. The black pixels indicate pixels as background pixels and white pixels indicate the pixels as part of the object.

### D. Background modeling

Underwater video taken with a static camera has a dynamic background. The dynamic background is a background that changes due to a small movement of the background object (e.g., the movement of plants caused a water flow) and changes in intensity due to the weather so that the intensity of the background pixels changed. The good background modeling has a big influence on the success of the object detection.

The background model was obtained from updating a value of the mean and variance when there is a new incoming frame. Updating a mean value using (3) and a variant value using (4).

$$\mu_{T+1_{(x,y)}} = \begin{cases} \rho * f_{t+1_{(x,y)}} + (1-\rho) * \mu_{T_{(x,y)}} \\ \quad if\ f_{T+1_{(x,y)}} = background \\ (1-\beta) * f_{T+1_{(x,y)}} + \beta * \mu_{T_{(x,y)}} \\ if\ f_{T+1_{(x,y)}} = \ not\ background \end{cases} \qquad (3)$$

$$\sigma^2_{T+1_{(x,y)}} = \begin{cases} \rho * \left(f_{T+1_{(x,y)}} - \mu_{T+1_{(x,y)}}\right)^2 + (1+\rho) * \sigma^2_{T_{(x,y)}}, \\ \quad if f_{T+1_{(x,y)}} = background \\ (1+\beta) * \left(f_{T+1_{(x,y)}} - \mu_{T+1_{(x,y)}}\right)^2 + \beta * \sigma^2_{T_{(x,y)}}, \\ \quad if\ f_{T+1_{(x,y)}} = not\ background \end{cases} \qquad (4)$$

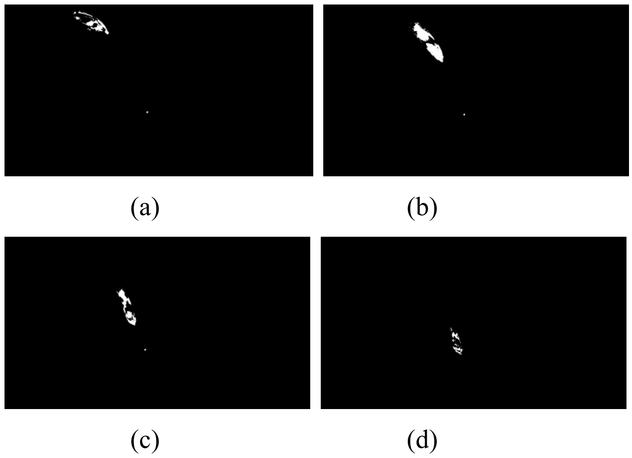



Fig. 1 The results of the pixel determination from video 3 for frame 28, 35, 42, and 53 successively

On the proposed method, the value of $\rho$ and $\beta$ are variable learning rate, which its value is determined. The background modeling results are shown in Fig. 2 and Fig. 3.

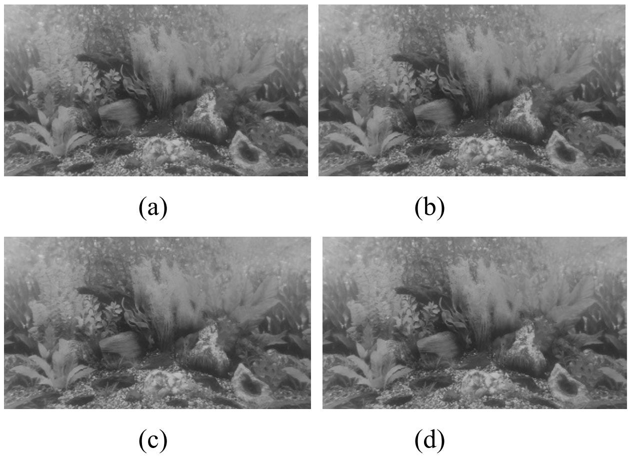



Fig. 2. The results of the background modeling from video 1 for frame 2, 30, 60, and 90 successively.

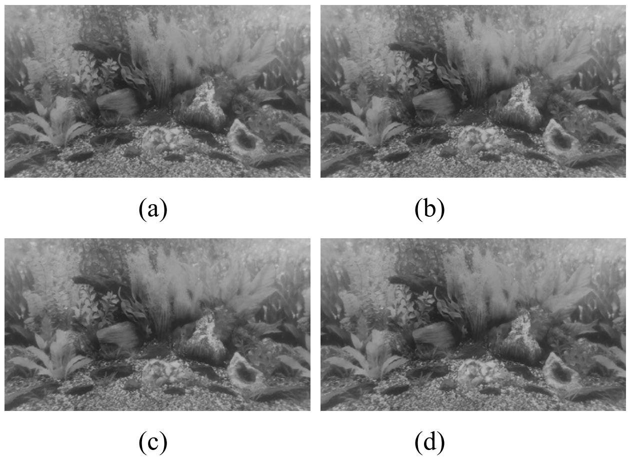



Fig. 3. The results of the background modeling from video 2 for frame 2, 30, 60, and 90 successively.

### *E. The object detection*

The detection of the object is done by subtracting the pixels in the current frame with the pixels in the previous frame background model. Fig. 1 shows the result of a subtraction of the pixel so that it looks the object represented by white pixels. After the object is known, it needs to be given a bounding box that states the location of the object. Done first morphological processes from the background subtraction result, so that the depiction of the bounding box works well. The morphological process used is the process of closing. Fig. 4 shows an example of the object detection using the proposed method.

## III. EXPERIMENTAL RESULTS

For experiment, we use 4 video scenes. Two videos are used for modeling the background (video 1 and video 2). Video 1 and Video 2 do not contain an object, only the dynamic background. The other two videos are used for object detection (video 3, and video 4). Video 3 and 4 contain the object to be detected.

The video frame rate and the size of an image are 30 fps and 1920x1080 pixels, respectively. The experimental environment is as follows: Operating system is Windows 8, processor is Intel® core™ i5, 4GB RAM, and the used software is Ms Visual Studio 2010 and OpenCV 2.4.13.

The background modeling results measured using a PSNR parameter. Fig. 5 and 6 shows the average of PSNR for video 1 (morning) and video 2 (afternoon) with $\beta$ and $\rho$ values was modified. From Fig. 5 and 6 is seen that the value of $\beta$ = 0.2 and $\rho$ = 0.2 provide the largest PSNR value. The larger of the PSNR value then the result of the background modeling getting better because can reduce the intensity change or a small change in the background. Equation (5) used to obtain the PSNR value of the image.

$$PSNR = 20log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \quad (5)$$

$$MSE = \frac{1}{M\times N}\sum_{i=1}^{N}\sum_{j=1}^{M}[f(i,j) - \mu(i,j)]^2 \quad (6)$$

Here, $M$ and $N$ is the size of the image, while i and j are the pixel positions.

Fig. 4. The results of the detection object using the proposed method.

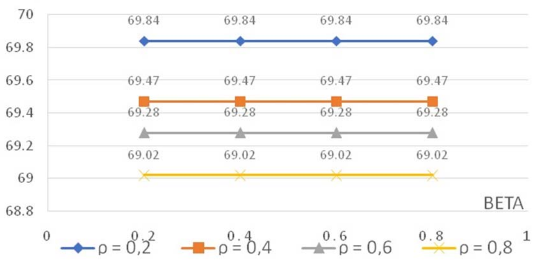



Fig. 5. The average of PSNR for video 1.

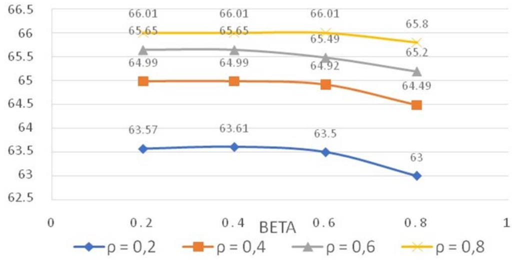
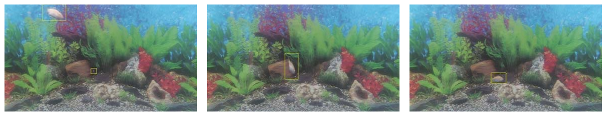



Fig. 6. The average of PSNR for video 2.

The results of object detection using the proposed method are shown in Fig. 7 and Fig. 8. Fig. 7 shows the results of object detection on frame 25, 50, and 75 respectively on the third video which taken in the morning. While the results of object detection on the frame 25, 50, and 75 in the fourth video taken in the afternoon are shown in Fig. 8.

(a)

(b)



(c)

Fig. 7. Result of object detection using the proposed method for video 3. Time elapses from up to down: (a) Original image, (b) the result of the pixel determination, and (c) the object detected.

(a)

(b)

(c)

Fig. 8. Result of object detection using the proposed method for video 4. Time elapses from up to down: (a) Original image, (b) the result of the pixel determination, and (c) the object detected.

The effectiveness of the proposed object detection method is evaluated using the parameters of recall, precision, and F-measure. Equation (7), (8), and (9) are used to determine the value of the recall, precision, and F-measure.

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

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

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

Here $N_{TP}$ stating the number of pixels that are detected as objects and indeed the object. $N_{FN}$ is the number of pixels of the object that's not detected by the proposed detection method. $N_{FP}$ is the number of background pixels that detected as an object by the proposed detection method. To determine the value of $N_{TP}$, $N_{FN}$, and $N_{FP}$, then the image of the determination pixels results compared with the ground truth are created manually. Table 1 shows the results of the evaluation of the proposed method.

Table 1. The result of evaluation of the proposed method.

| Video | Evaluation values | | |
|---|---|---|---|
| | Recall | Precision | F-measure |
| Video 3 | 95% | 57% | 70 |
| Video 4 | 72% | 78% | 74 |

## IV. Conclusion

This paper proposes a technique to detect moving objects on underwater video using background modeling. The performance of the proposed method provides satisfactory value because a recall value well above 50% and also a precision well above 50%, are shown in Table 1. This technique was tested on two pieces of an underwater video taken at different times. The third video was taken in the morning while the fourth video was taken in the afternoon.

As future work, this proposed method needs to be tested in the condition of night by using additional lighting. This is consistent with the need to help to catch fish using underwater submersible lights. To improve the effectiveness of the proposed method needs to be done the processing on a model RGB image. In addition, necessary to do research to modify the proposed methods for videos captured using the moving camera.

## Acknowledgment

Thanks to LPPM UNILA for providing financial support through DIPA BLU Research, number of a research assignment letter (from Chairman LPPM Unila) 550/UN26/8/LPPM/2016.

## References

[1] Amandeep, E.M.Goyal, “Moving object detection using background subtraction techniques,” International Journal Of Engineering Sciences & Research Technology, vol. 5, pp. 1-9, 2016.

[2] J. V. Turgude, S.S. Shetkar,"A survey of moving target detection using dynamic background in video surveillance system," International Journal of Advance Engineering and Research Development, Vol. 3, pp. 126-133, 2016.

[3] Y. Yu, C. Zhou, L. Huang, Z. Yu, “A moving target detection algorithm based on the dynamic background,” International Conference on Computational Intelligence and Software Engineering, Wuhan – China, pp. 1-5, Dec. 2009.

[4] Y. K. Jain, S. Gupta, "A highly adaptive method for moving target detection in dynamic background with a simplified manner," International Journal of Computer Applications, Vol. 102, No. 10, pp. 20-26, Sept. 2014.

[5] D.S. Pham, O. Arandjelovi´c, S. Venkatesh, "Detection of dynamic background due to swaying movements from motion features," IEEE Transactions on Image Processing, Vol. 24, No. 1, pp 332-344, Dec. 2014.

[6] F. X. A. Setyawan, J. K. Tan, H. Kim, S. Ishikawa, “Detecting moving objects from a video taken by a moving camera using sequential inference of background images,” Artif life Robotics, Vol. 19, pp. 291-298, 2014.

[7] F. X. A. Setyawan, J. K. Tan, H. Kim, S. Ishikawa, “Detecting foreground objects by sequential background inference in a video captured by a moving camera,” Proceedings of the SICE Annual Conference, Nagoya - Japan, pp. 1699-1702, September 2013.

[8] F. X. A. Setyawan, J. K. Tan, H. Kim, S. Ishikawa, “Moving objects detection employing iterative update of the background,” Artif life Robotics, Vol. 22, pp. 168-174, 2017 doi:10.1007/s10015-016-0347-9

[9] M. S. Srividya, R. Hemavathy, G. Shobha, "Underwater video processing for detecting and tracking moving object," International Journal Of Engineering And Computer Science, Vol. 3, No. 5, Pp 5843-5847, May 2014.

[10] C. Spampinato, Y. H. Chen-Burger, G. Nadarajan, and R. Fisher, “Detecting, tracking and counting fish in low quality unconstrained underwater videos,” Proc. 3rd Int. Conf. on Computer Vision Theory and Applications (VISAPP), vol. 2, pp. 514-519, 2008.

[11] O.Ancha, "Speed algorithm for underwater moving objects detection based on image sequence," International Journal of Computer Science and Information Technologies, Vol. 7, No. 3, pp. 1032-1035, 2016.

[12] M. R. Shortis, E. S. Harvey, "Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations," International Archives Photogrammetry and Remote Sensing, Vol. 32, No. 5, pp. 792-799, 1998.