Combining background subtraction and three-frame difference to detect moving object from underwater video

Jialun Dai, Ruchen Wang, Haiyong Zheng*, Bing Zheng, Guangrong Ji
Department of Electronic Engineering
College of Information Science and Engineering
Ocean University of China
Qingdao, China 266100
*Corresponding author: zhenghaiyong@ouc.edu.cn

Abstract—**Moving object detection is the basis of object tracking and classification in computer vision, which can be applied to underwater robots to perform underwater missions and do marine ecological research. The complicated scene and poor lighting condition in underwater environment, however, usually make moving object detection very hard. To solve above problems, we propose an approach combining background subtraction and three-frame difference to detect moving object from underwater video, in which background subtraction provides the object information to supplement the incomplete information detected from three-frame difference. Firstly, moving object pixels are detected by background subtraction and three-frame difference perspectively. Next, perform "AND" operation on the results of background subtraction and three-frame difference. Lastly, morphology processing is performed on the result to remove noise caused by non-static objects in background. Experimental results show that an effective and reliable detecting performance for moving object from underwater video is achieved.**

I. Introduction

Moving object detection in video with the aid of visual anysis has drawn increasing attention. Detecting the moving objects such as fishes from underwater video has important potential applications in Remotely Operated Vehicles (ROVs) or Autonomous Underwater Vehicles (AUVs). With the advanced technique of underwater object detection, it can accelerate the development of ROVs or AUVs and make them perform the high-difficulty and demanding missions, such as underwater survey mission, biological observation and marine ecological research.

At present, methods utilized in moving object detection are mainly the background subtraction method, the frame difference method and the optical flow method[1]. Background subtraction is a widely used approach for detecting moving objects in video from static cameras[2]. The rationale in the approach is to detect moving objects from the difference between the current frame and a reference frame, called as "background model". This method needs to consider the representation and initialization of background model, and the updating of background model[3]. And a robust background subtraction also should be able to handle lighting changes, repetitive motions from clutter and long-term scene changes[4]. Frame differencing is a technique where the computer checks the difference between two consecutive video frames. If the pixels have changed, there apparently was something changing in the later frame[5]. Owing to the advantages of small calculations and high performance of real-time, this method is commonly used in object detection. Optical flow method makes use of the instantaneous velocity which is generated in the continuous movements of the pixels in the moving objects. In the case of unknown scene information, the method can detect the moving objects. Optical flow method has poor anti-noise performance, the calculation is quite complicated, and the hardware requirements are also very high, so it is not suitable for real time processing[6]. However, there are still several challenges in detecting moving object from underwater video, including scene noise, due to a poor quality image source, small movements of non-static objects and the variable lighting conditions in the scene. And conventional methods of moving object detection can not solve the above problems very well.

Therefore, a method combining background subtraction and three-frame difference is proposed to detect moving object from underwater video efficiently. Background subtraction generates a difference image to indicate object pixels, but also information of some movements of non-static objects. Three-frame difference provides an effective result, which is insensitive to the noise caused by background scene and movements of non-static objects, but may not contains complete object information. We utilize the object information obtained from background subtraction to supplement the incomplete information computed from three-frame difference, which can remove the noise from background scene and overcome the difficulties imposed by small movements of non-static objects. And the experimental results show the method we proposed can achieve accurate performance in detecting moving object from underwater video.

II. Methods

Figure 1. explains the steps of our method for moving object detection from underwater video. The method we proposed comprises five steps: *(i)* Capture one frame randomly from the underwater video and the underwater video is shot in the lab where we build a pool to simulate underwater environment. From the underwater video, the light is poor and the noise caused by tiny particles is very big. *(ii)* According to the underwater video, we can establish a background

978-0-933957-43-5 ©2015 MTS

This is a DRAFT. As such it may not be cited in other works.
The citable Proceedings of the Conference will be published in
IEEE Xplore shortly after the conclusion of the conference.

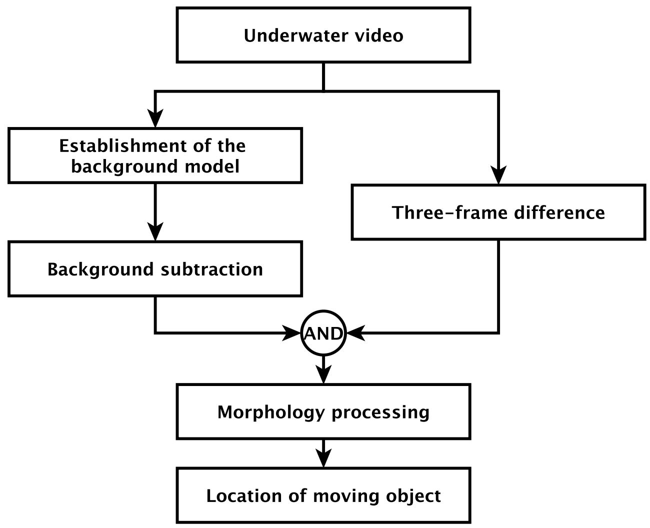

Fig. 1. The flow chart of our method for moving object detection from underwater video by combining background subtraction and three-frame difference.

model without objects. Under the background model, use the background subtraction to compare the frame captured with the background model. We can get some pixels that indicates possible moving objects or some noise. *(iii)* Three-frame difference is adopted to determine the possible moving object pixels from the frame captured. We adjust the frame interval considering the underwater video. After three-frame difference, morphology operation is performed to improve accuracy. *(iv)* Perform "AND" operation on results of background subtraction and three-frame difference to determine moving object from underwater video. Background subtraction generates the objects pixels result, but it is sensitive to noise. And three-frame difference provides an effective way to overcome the effect of noise, but it may cause the loss of outline information. In order to solve above problem, we combine these two methods to detect moving object from underwater video in this step. *(v)* To eliminate the effect of background noise and camera shake, the fundamental methods of mathematical morphology, erosion and dilation, are utilized to remove noise and fill the empty of target to get a better result in the end.

A. Background subtraction

Background subtraction method is to use the difference method of the current frame and background frame to detect moving objects, with simple algorithm, but very sensitive to the changes in the external environment and has poor anti-interference ability[1]. The basic idea of the background subtraction is to do differential between current frame and background frame. And according to the threshold to judge whether it belongs to the moving objects or not.

The video sequence is collected during a certain period of time. $f_k(x, y)$ represents the current frame, and $B(x, y)$ represents the background model which is obtained by averaging several frames from the beginning of underwater video.

$$B(x, y) = \frac{1}{N}\sum_{k=1}^{N} f_k(x, y) \tag{1}$$

$D_k(x, y)$ represents the difference between the current frame and background model.

$$D_k(x, y) = f_k(x, y) - B(x, y) \tag{2}$$

$$D_k(x, y) = \begin{cases} 1 & D_k(x, y) \geq T \\ 0 & \text{else} \end{cases} \tag{3}$$

In the background training process, the reference background image and some parameters associated with normalization are computed over a number of static background frame. The background is modeled statistically on a pixel by pixel basis[7].

In background subtraction, the most critical part is to establish and update background model background. In our method, the background is computed by the the average background model. The average background model, which is more representative and robust, is selected to replace the normal background model and it is established with 15 frames from the beginning of underwater video where there are no object. After that, we just need to compare the current frame with the average model to detect moving object pixels.

B. Three-frame difference

Moving object detection from a sequence of frames from a static camera is widely performed by frame difference method. The objective of the approach is to detect moving objects from the difference between the existing frame and the reference frame. The frame difference method is the common method of motion detection. The method adopts pixel-based difference to find the moving object[1].

For the video sequence which is shooting by a stationary camera, we make smooth denoising for frame k and frame $k+1$ first, and then treat them by frame difference, i.e. frame k subtracts frame $k+1$, and we can get a binary image $D(x, y)$, which is expressed as follows:

$$D(x, y) = \begin{cases} 1 & \mid f_k(x, y) - f_{k-1}(x, y) \mid \geq T \\ 0 & \mid f_k(x, y) - f_{k-1}(x, y) \mid < T \end{cases} \tag{4}$$

From the formula of frame difference, it can be seen that, if the moving object is in uniform motion, the moving object detected by the method of frame difference would be the same one. If the moving object accelerates, the detection result gives more or less numbers than the reality. In addition, the moving object detected by the method of frame difference contains information of two frames. In order to overcome the disadvantage of the method of frame difference, three-frame difference has been proposed[8].

The first step of three-frame difference is to make smooth denoising for three consecutive frames, and then process them by the method of frame difference respectively, i.e. frame k subtracts frame and frame $k-1$, and we can get a binary image $D_1(x, y)$, frame $k+1$ subtracts frame k, and we can get a binary image $D_2(x, y)$, and the final step is to make an and operation of $D_1(x,y)$ and $D_2(x,y)$, the result is three-frame difference image $D(x,y)$, which can be expressed as follows:

$$D_1(x,y) = \begin{cases} 1 & \mid f_k(x,y) - f_{k-1}(x,y) \mid \geq T \\ 0 & \mid f_k(x,y) - f_{k-1}(x,y) \mid < T \end{cases} \quad (5)$$

$$D_2(x,y) = \begin{cases} 1 & \mid f_{k+1}(x,y) - f_k(x,y) \mid \geq T \\ 0 & \mid f_{k+1}(x,y) - f_k(x,y) \mid < T \end{cases} \quad (6)$$

$$D(x,y) = \begin{cases} 1 & D_1(i,j) \cap D_2(i,j) = 1 \\ 0 & D_1(i,j) \cap D_2(i,j) = 0 \end{cases} \quad (7)$$

From underwater video, we can see that many motions are in scene, but actually the moving object, such as fish, is what we want to detect and the motion of some tiny particles in background is noise. Moving object is in uniform motion and object moves a longer distance than non-objects in the same period of time. The distances of object and non-static objects moving are approximate in a short time, so it is not appropriate to use adjacent three frames to distinguish moving object pixels from non-static objects pixels by normal three-frame difference. To solve this problem, we increase the interval of frames from 1 to 4. If we use adjacent three frames, the difference image $D(x,y)$ may contains moving object pixels and noise pixels and it is difficult to distinguish these pixels because they are all small regions. Now, the result of three-frame difference image $D(x,y)$ can be expressed as follows:

$$D_1(x,y) = \begin{cases} 1 & \mid f_k(x,y) - f_{k-4}(x,y) \mid \geq T \\ 0 & \mid f_k(x,y) - f_{k-4}(x,y) \mid < T \end{cases} \quad (8)$$

$$D_2(x,y) = \begin{cases} 1 & \mid f_{k+4}(x,y) - f_k(x,y) \mid \geq T \\ 0 & \mid f_{k+4}(x,y) - f_k(x,y) \mid < T \end{cases} \quad (9)$$

$$D(x,y) = \begin{cases} 1 & D_1(i,j) \cap D_2(i,j) = 1 \\ 0 & D_1(i,j) \cap D_2(i,j) = 0 \end{cases} \quad (10)$$

If we increase the interval of frames from 1 to 4, we can see some big regions and small regions from the result of difference image $D(x,y)$. The regions of moving object in difference image may be bigger than the regions of noise if the frame interval is bigger due to the longer distance of object moving. After that, we just need to perform some morphology processing to eliminate the effect of noise. However, the outline of object may be incomplete and it may leave some empty in object in result.

C. Combining background subtraction and three-frame difference

Through background subtraction and three-frame difference, we have obtained that some possible moving object pixels from frames. But these two results may have been affected by noise of underwater scene. If we just use one of these results to determine moving object pixels, the detecting result will be very bad. Therefore, we can combine the background subtraction and three-frame difference.

In background subtraction, the outline and shape may be conversed quite well. But due to the simplicity, the result may detect moving object pixels efficiently, but is affected by the noise of underwater environment heavily to result in containing lots of noise pixels. In three-frame difference, we adjust the frame interval according to the poor condition of underwater video and obtain a result which is affected by the noise little, but contains incomplete moving object information.

Considering all of the above, we utilize background subtraction and three-frame difference perspectively first and then perform "AND" operation on these results of two methods, not only conserving the most information and outline of moving objects, but also obtaining the position and region of objects and eliminate the noise.

D. Morphology processing

The result after combining background subtraction and three-frame difference indicates the most possible pixels of moving objects from underwater video. In order to improve the accuracy, we use morphology processing, erosion and dilation, to process the combining result. It can remove some noise caused by non-static objects or in background by the erosion. It can keep the completion of objects and fill the empty of target by the dilation operation. According to the final result, we can locate the moving object accurately from underwater video.

III. Experimental results

The section demonstrates the performance of the proposed method on one frame of underwater video.

Fig. 2. One frame is captured randomly from underwater video and it is used to detect moving object in this frame

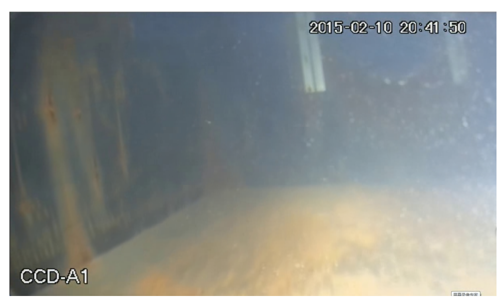

Fig. 3. Use 15 frames from the beginning of underwater video to establish an average background. There is no objects in the average background.

IV. Conclusion

We propose a method combining background subtraction and three-frame difference to ensure a reliable moving object detection from underwater video with complicated environment and low contrast. Background subtraction establishes an average background, which is more accurate and simple to perform. And three-frame difference is developed to overcome

Fig. 4. Compare the current frame with the average background by background subtraction. The result indicates that the white pixels are possible moving object pixels. But due to the poor image quality and noise caused by non-static objects, these pixels may cause errors.

Fig. 5. Through the three-frame difference, we can obtain a more accurate moving objects pixels with less noise pixels. Big regions may be more possible objects.

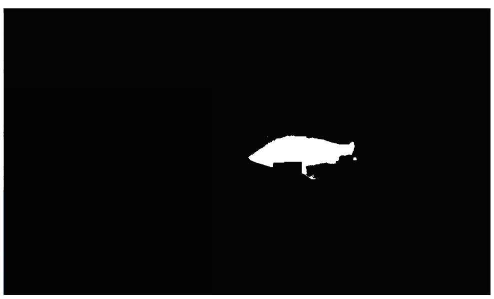

Fig. 6. Perform "AND" operation on results of background subtraction and three-frame difference. The morphology processing can eliminate the effect of noise and improve the accuracy.

Fig. 7. Draw a green rectangle to label the moving object. From the figure, we can see that the fish, which is what we want to detect, is labeled exactly.

the challenges posed by unstable illumination and ubiquitous noise in underwater environment. The experiments show that the proposed method has good performance and efficiency in detecting moving object from underwater video. Future work includes multiple and real-time moving objects detection in real and more complicated underwater environment from video.

Acknowledgment

This work was supported by National Natural Science Foundation of China (Grant No. 61301240 and No. 61271406), Natural Science Foundation of Shandong Province (Grant No. ZR2014FQ027) and Basic Research Program of Qingdao Municipal Scientific Plan (Grant No. 13-1-4-223-jch).

References

[1] N. Singla, "Motion detection based on frame difference method," *International Journal of Information & Computation Technology., ISSN*, pp. 0974–2239, 2014.

[2] M. Piccardi, "Background subtraction techniques: a review," vol. 4. IEEE, 2004, pp. 3099–3104.

[3] Y. Zhang, X. Wang, and B. Qu, "Three-frame difference algorithm research based on mathematical morphology," *Procedia Engineering*, vol. 29, pp. 2705–2709, 2012.

[4] B. Tamersoy, "Background subtraction–lecture notes," *University of Texas at Austin*, 2009.

[5] N. Prabhakar, V. Vaithiyanathan, A. P. Sharma, A. Singh, and P. Singhal, "Object tracking using frame differencing and template matching," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 24, pp. 5497–5501, 2012.

[6] M. Li, J. Fan, Y. Zhang, and R. Zhang, "Moving object detection and tracking algorithm," *TELKOMNIKA*, vol. 11, no. 10, pp. 5539–5544, Oct 2013.

[7] T. Horprasert, D. Harwood, and L. S. Davis, "A robust background subtraction and shadow detection," 2000, pp. 983–988.

[8] W.-h. HE, P. LI, Y.-m. WEN, and B. YE, "Algorithm of extracting moving object silhouette based on frame fusion under complex background [j]," *Journal of Computer Applications*, vol. 1, p. 037, 2006.