

Underwater Video Processing for Detecting and Tracking Moving Object

Srividya M. S., Hemavathy R., Shobha G.

Abstract: In this paper, we present a vision system capable of analyzing underwater videos for detecting and tracking moving object. The video processing system consists of three subsystems, the video texture analysis, object detection and tracking modules. Moving object detection is based on adaptive Gaussian mixture model. The tracking was carried out by the application of the Kalman algorithm that enables the tracking of objects. Unlike existing method, our approach provides a reliable method in which the moving object is detected in unconstrained environments and under several scenarios (murky water, algae on camera lens, moving plants, low contrast, etc.). The proposed approach was tested with 20 underwater videos, achieving an overall accuracy as high as 85%.

Keywords: Video Processing, Detection, Tracking, Gaussian Mixture Model, Kalman Filtering

I. INTRODUCTION

Traditionally, marine biologists determine the existence and quantities of different types of fish using several methods, including casting nets in the ocean for collecting and examining fish, human underwater observation and photography, combined net casting and, more recently, human hand-held video filming. Each of such methods has their drawbacks. For instance, although the net-casting method is accurate, it kills the collected fish, damages their habitat and costs much time and resources. Human manned photography and video-making but do not damage observed fish or their habitat, the collected samples are scarce or limited and is intrusive to the observed environment therefore do not capture normal fish behaviors. This paper presents an alternative approach by using an automated Video Processing (VP) system that analyses videos to identify interesting features. These videos are taken automatically and continuously by underwater video-surveillance cameras. The proposed automated video processing system is able to handle large amount of videos automatically and speedily gives those un-watched video clips a chance to be analyzed. The system's performance is promising, as shall be reported later on in this paper. The next session introduces our Image Processing (IP) Tasks, which is a description of our texture and color analysis subsystem. The fish detection and tracking systems are then described followed by an analysis of their performance.

Manuscript Received on July, 2014

Srividya M. S., Student, Information Technology, G.H.Raisoni Institute of Engineering & Technology, Pune, India.

Hemavathy R., Student, Information Technology, G.H. Raisoni Institute of Engineering & Technology, Pune, India.

Shobha G., Student, Information Technology, G.H. Raisoni Institute of Engineering & Technology, Pune, India.

II. IMAGE PROCESSING TASKS

Currently, marine biologists manually analyze underwater videos to find useful information. This procedure requires a lot of time and human concentration, as an operational camera will generate huge amount of imagery data. For one minute's video, it will take a human about 15 minutes for classification and annotation. To fully analyze existing videos alone, generated by the three underwater cameras over the past 1 year, will take human approximately 100s of years. The proposed system therefore aims to support the non-Video processing trained users to leverage Video and Image Processing software to help their normal line of work.

2.1 Object Detection

For detecting the moving object Adaptive Gaussian Mixture Model was used. The real-time segmentation of moving regions in image sequences involves background subtraction. The numerous approaches to this problem differ in the type of background model used and the procedure used to update the model. Here we model each pixel as a mixture of Gaussians. The Gaussian distributions of the adaptive mixture model are then evaluated to determine which are most likely to result from a background process. Each pixel is classified based on whether the Gaussian distribution which represents it most effectively is considered part of the background model. This results in a stable, real-time detector which reliably deals with lighting changes, repetitive motions from clutter, and long-term scene changes. The usage of Gaussian mixture models for video segmentation has been widely adopted. However, the main difficulty arises in choosing the best model complexity. High complex models can describe the scene accurately, but they come with a high computational requirement, too. Low complex models promote segmentation speed, with the drawback of a less exhaustive description. The System compares color or grayscale video frame to a background model to determine whether individual pixels are part of the background or the foreground. It then computes a foreground mask. By using background subtraction, you can detect foreground objects in an image taken from a stationary camera. Based on the persistence and the variance of each of the Gaussians of the mixture, we determine which Gaussians may correspond to background colors. Pixel values that do not fit the background distributions are considered foreground until there is a Gaussian that includes them with sufficient, consistent evidence supporting it. If only lighting changed over time, a single, adaptive Gaussian per pixel would be sufficient. In practice,

multiple surfaces often appear in the view frustum of a particular pixel and the lighting conditions change. Thus, multiple, adaptive Gaussians are necessary. We use a mixture of adaptive Gaussians to approximate this process. Each time the parameters of the Gaussians are updated, the Gaussians are evaluated using a simple heuristic to hypothesize which are most likely to be part of the “background process.” Gaussians are grouped using connected components. The distribution of recently observed values of each pixel in the scene is characterized by a mixture of Gaussians. A new pixel value will, in general, be represented by one of the major components of the mixture model and used to update the model. Because there is a mixture model for every pixel in the image, we implement K-means approximation. Every new pixel value, X_t , is checked against the existing K Gaussian distributions, until a match is found. The threshold can be perturbed with little effect on performance. This is effectively a per pixel/per distribution threshold.

2.2 Background Model Estimation

As the parameters of the mixture model of each pixel change, we would like to determine which of the Gaussians of the mixture are most likely produced by background processes. Heuristically, we are interested in the Gaussian distributions which have the most supporting evidence and the least variance. To understand this choice, consider the accumulation of supporting evidence and the relatively low variance for the “background” distributions when a static, persistent object is visible. In contrast, when a new object occludes the background object, it will not, in general, match one of the existing distributions which will result in either the creation of a new distribution or the increase in the variance of an existing distribution. Also, the variance of the moving object is expected to remain larger than a background pixel until the moving object stops. To model this, we need a method for deciding what portion of the mixture model best represents background processes. After re-estimating the parameters of the mixture, it is sufficient to sort from the matched distribution towards the most probable background distribution, because only the matched models relative value will have changed. This ordering of the model is effectively an ordered, open-ended list, where the most likely background distributions remain on top and the less probable transient background distributions gravitate towards the bottom and are eventually replaced by new distributions.

2.1 Connected Components

The method described above allows us to identify foreground pixels in each new frame while updating the description of each pixel’s process. These labeled foreground pixels can then be segmented into regions by a two-pass, connected components algorithm. Because this procedure is effective in determining the whole moving object, moving regions can be characterized not only by

their position, but size, moments, and other shape information. Not only can these characteristics be useful for later processing and classification, but they can aid in the tracking process. Establishing correspondence of connected components between frames is accomplished using a linearly predictive multiple hypotheses tracking algorithm which incorporates both position and size. We have implemented a method for seeding and maintaining sets of Kalman filters. At each frame, we have an available pool of Kalman models and a new available pool of connected components that they could explain. First, the models are probabilistically matched to the connected regions that they could explain. Second, the connected regions which could not be sufficiently explained are checked to find new Kalman models. Finally, models whose fitness (as determined by the inverse of the variance of its prediction error) falls below a threshold are removed. Matching the models to the connected components involves checking each existing model against the available pool of connected components which are larger than a pixel or two. All matches are used to update the corresponding model. If the updated model has sufficient fitness, it will be used in the following frame. If no match is found a “null” match can be hypothesized which propagates the model as expected and decreases its fitness by a constant factor. The unmatched models from the current frame and the previous two frames are then used to hypothesize new models. Using pairs of unmatched connected components from the previous two frames, a model is hypothesized. If the current frame contains a match with sufficient fitness, the updated model is added to the existing models. To avoid possible combinatorial explosions in noisy situations, it may be desirable to limit the maximum number of existing models by removing the least probable models when excessive models exist.

III. EXPERIMENTAL RESULTS

This section presents results of applying the proposed methods onto different videos with different backgrounds. The proposed system was implemented in the Matlab environment. From the Footage Search website, representing the NatureFootage, OceanFootage, and AdventureFootage collections, more than 50 underwater videos were downloaded and tested. In the output, we have captured 6 different frames. They are,

- a. The actual input video’s current frame,
- b. AGMM generated Background frame
- c. The difference frame by subtracting the current frame with the background frame
- d. AGMM generated foreground frame
- e. Foreground Mask frame
- f. Masked current frame



Figure 1: Snapshot of Video with the Movement of Jelly Fish

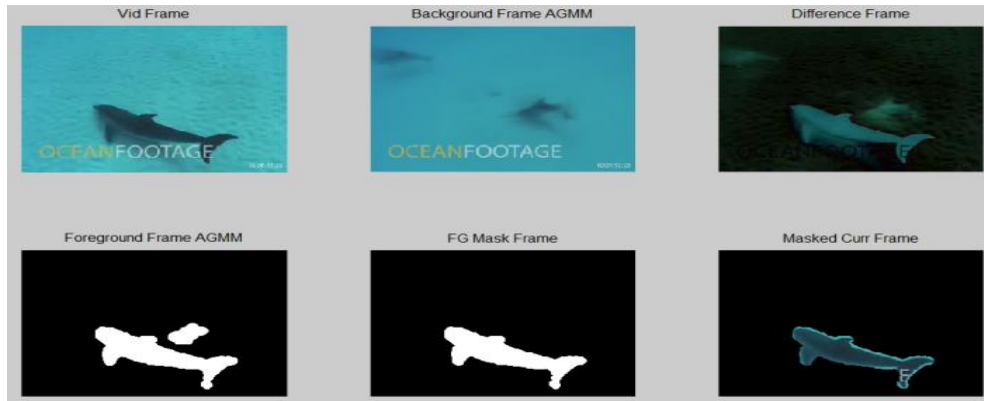


Figure 2: Another Snapshot of Video with the Movement of Jelly Fish

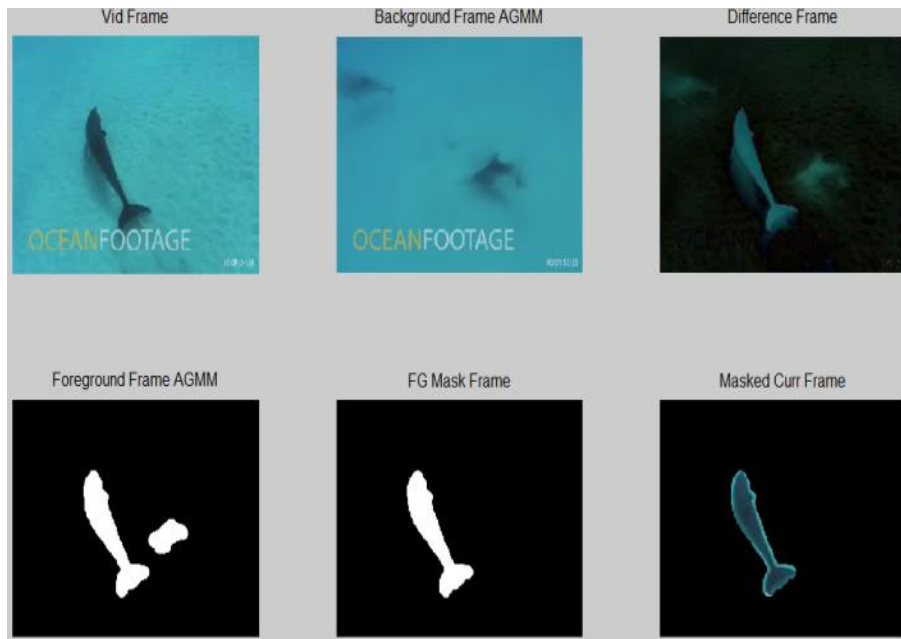


Figure 3: Snapshot of Video with the Movement of Fish

Underwater Video Processing for Detecting and Tracking Moving Object

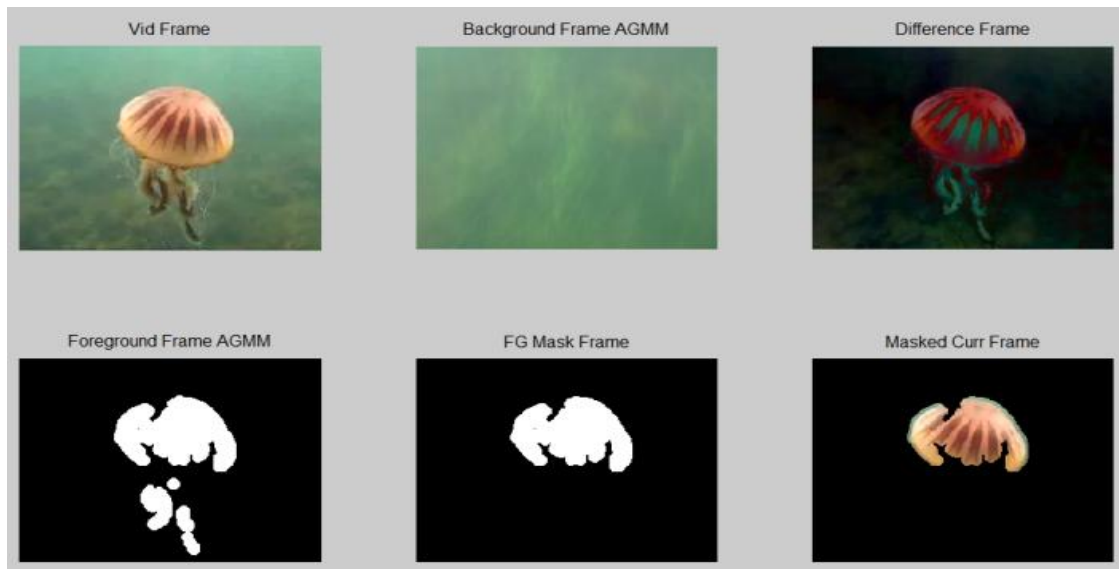


Figure 4: Another Snapshot of Video with the Movement of Fish

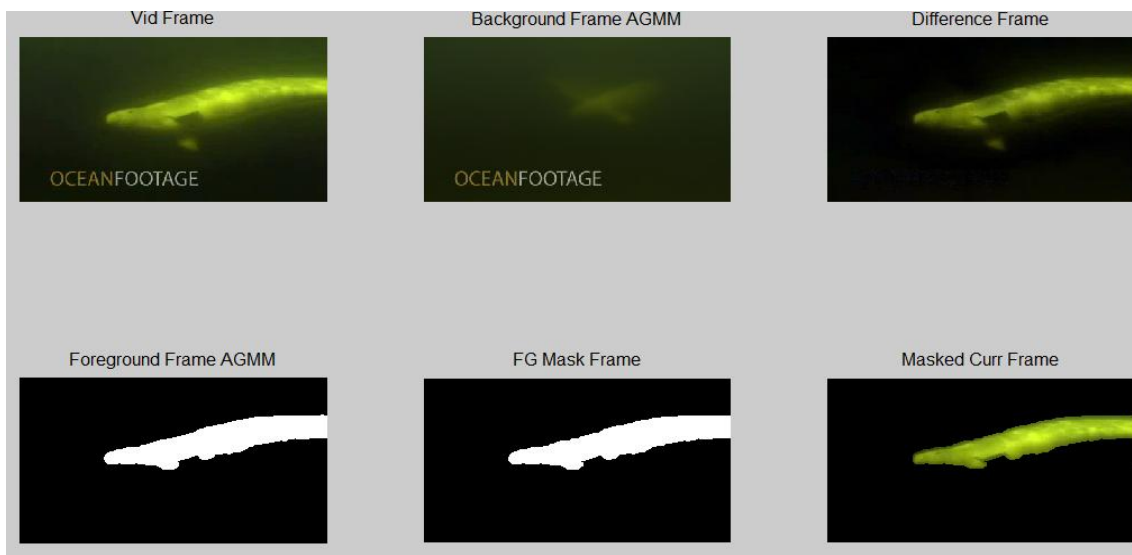


Figure 5: Snapshot of Video with Movement of Whale

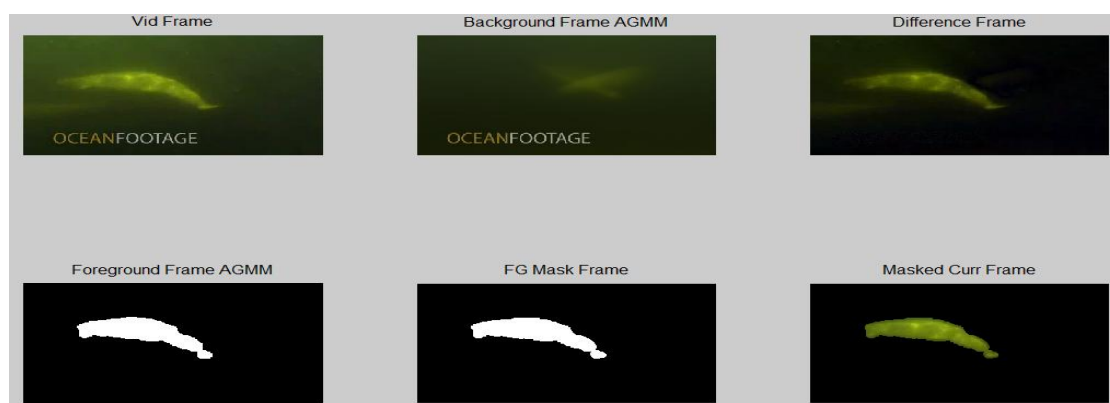


Figure 6: Another Snapshot of Video with the Movement of Whale

IV. FUTURE WORK AND RECOMMENDATION

There are some works we need to do further. Detection and tracking of two or more moving objects is the immediately next level of enhancement. Segmentation of multiple smaller objects and combining two or more approaches for detection and finally considering the better one, will give more efficient results. There is opportunity for further work.

V. CONCLUSION

This paper has provided an efficient method to detect moving objects using background subtraction framework.

The strength of this technique is that it is robust and efficient in segmenting the moving object. In order to prove our claim, we have demonstrated the proposed method for different video sequence. The main contribution towards the success of the detection in our method is that we have applied Adaptive Gaussian Mixture Model for background and foreground frame generation, and have applied Kalman filtering for tracking the object in consecutive frames.

REFERENCES

1. SamanPoursoltan, Russell Brinkworth, Matthew Sorell "Biologically-inspired Video Enhancement Method For Robust Shape Recognition," University of Adelaide, Australia, IEEE, 2013.
2. Prabhakar C J & Praveen Kumar P U. "Feature Tracking of Objects in Underwater VideoSequences.", Kuvempu University, India, ACEEE 2012.
3. SaeedVahabiMashak, BehnamHosseini, S.A.R. Abu-Bakar."Background Subtraction for Object Detection under Varying Environments."UniversitiTeknologi Malaysia, IJCISIMA, 2012
4. M. Weber, M. Welling, and P. Perona. "Detecting ,Tracking And Counting Fish In Low Quality Unconstrained Underwater Videos" University of Edinburgh, Edinburgh, UK.
5. "Underwater ColorConstancy :Enhancement of Automatic Live Fish Recognition" M. Chambah, D. Semani, A. Renouf, P. Courtellemont, A. Rizzi, Université de La Rochelle, France, Dept. of Information Technology - University of Milano/Italy.
6. CodrutaOrnianaAncuti, CosminAncuti, Tom Haber and Philippe Bekaert "FUSION-BASED RESTORATION OF THE UNDERWATER IMAGES," Belgium, IEEE, 2011.
7. MohamedAmer, Emil Bilgazyev "Fine-grained Categorization of Fish Motion Patterns in Underwater Videos", Oregon State University, ICCV, 2011
8. PiotrJasiobedzki, Stephen Se, Michel Bondy, and Roy Jakola,"Underwater 3D mapping and pose estimation for ROVoperations", OCEANS 2008, pp. 1-6, September 2008.
9. Li Ma, KaihuaWu,L. Zhu, "Fire smoke detection in videoimages Using Kalman filter and Gaussian Mixture Colormodel", International Conference on Artificial Intelligence andComputational Intelligence, 2010
10. Li Xu, Feihu Qi, Renjie Jiang, YunfengHao, Guorong Wu,"Shadow Detection and Removal in Real Images: A Survey"Shanghai JiaoTong University, P.R. China, June 2006.
11. J Shin, S Kim, et all, "Optical flow-based real-time object tracking using non-prior training active feature model"Computer Vision and Image Understanding, June 2005,Volume 98, Issue 3, Pages 462-490.
12. B.Hosseini, SaeedVahabiMashak and A.S.R Abu Bakar,"Human Movement Based on Rule based classifier" Asia Modelling Symposium, May 2010.
13. S.S. Beauchemin and J.L. Barron, "The Computation of Optical Flow", ACM Computing Surveys, 1995 Vol. 27,No.3, PP 43- 467.
14. S.Y. Chien, S.Y. Ma and L.G. Chen ," Efficient moving objectsegmentation algorithm using background registrationtechnique" IEEE Trans Circuits Syst. Video Technol, JULY2002
15. Hu Fuyuan, Zhang Yanning, et all, "A New Method of MovingObject Detection and Shadow Removal" JOURNAL OF ELECTRONICS (CHINA), July 2007
16. Dongxiang Zhou, Hong Zhang and NilanjanRay,"TextureBased Background Subtraction", Proceedings of theInternational Conference on Information and Automation,Zhangjiajie China, IEEE Conference, June 20 -23, 2008.
17. YunlongGuo, Bo Yang, Yangyang Ming, Aidong Men, "AnEffective Background Subtraction Under the Mixture of MultipleVarying Illuminations",Second InternationalConference on Computer Modeling and Simulation, IEEEConference, 2010.
18. V. Brandou, A. G. Allais, M. Perrier, E. Malis, P. Rives, J.Sarrazin, and P. M. Sarradin, "3D Reconstruction of NaturalUnderwater Scenes Using the Stereovision System IRIS", OCEANS 2007 - Europe , pp. 1-6, 2007.
19. David G Lowe, "Distinctive image features from scaleinvariantkeypoints", *International Journal of Computer Vision*, vol. 60(2),pp. 91-110, 2004.
20. David G Lowe, "Object recognition from local scaleinvariantfeatures", Proceedings of the International Conference on ComputerVision, vol. 2, pp. 1150-1157, 1997.