# Automatic object detection and segmentation from underwater images via saliency-based region merging

**Yafei Zhu[1], Xinxin Qiu[1], Xiaoqing Sun[1], Haiyong Zheng[1,*], Bing Zheng[1]**

**1 College of Information Science and Engineering, Ocean University of China, Qingdao, China**

∗ **E-mail: zhenghaiyong@ouc.edu.cn**

Object detection and segmentation from underwater images play an essential role in automatic tracking and recognition of underwater targets such as fishes. Generally, the objects in underwater images are not clearly visible due to low contrast and high noise in the environment, which increases the difficulty of detection and segmentation. Interactive schemes with a few simple user inputs are good solutions for robust detection and accurate segmentation. However, they are usually tedious and not effective in some situations. This paper proposes a novel automatic object detection and segmentation framework to deal with the above problems. Firstly, the dark channel prior [1] is applied to enhance the image by removing the "haze" that makes the underwater image unclear. Then, for detecting salient objects in the enhanced image with complicated background, we choose DRFI [2] model to generate the saliency map containing "saliency values" per pixel. And the obtained saliency map is further processed to determine the information of object and background. Finally, the region merging based interactive image segmentation method (called MSRM) [3] is improved by adding the determined object and background information as the user inputs so that the algorithm becomes automatic. The framework of our proposed method is shown in Fig. 1.
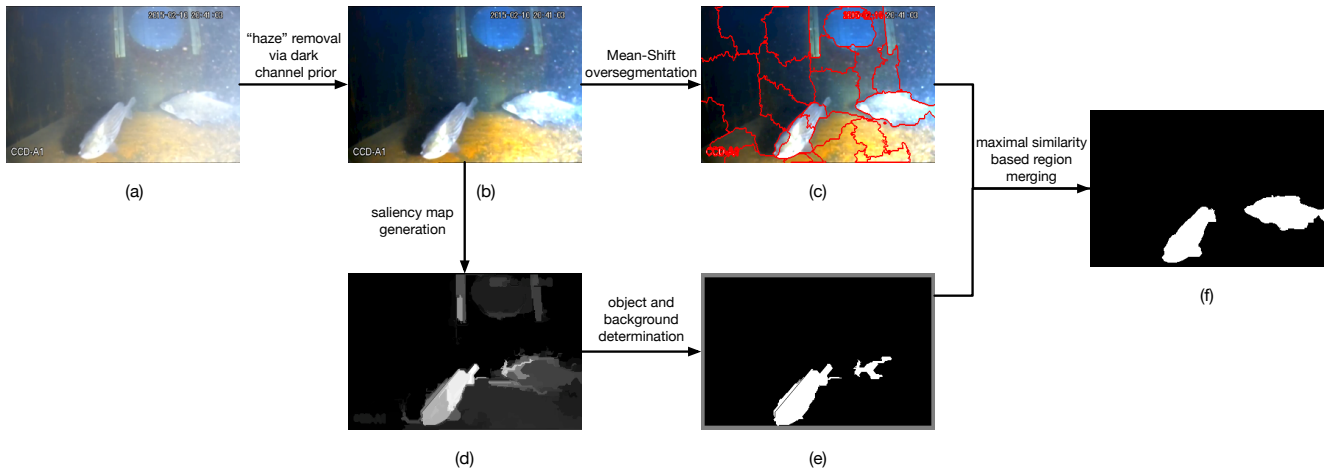


Fig. 1: The framework of our proposed method for object detection and segmentation from underwater images.

The dark channel prior was proposed to remove haze from a single input image [1]. It is based on the observation that most local patches in haze-free images contain some pixels which have very low intensities in at least one color channel. So dark pixels in the haze image can be used to estimate the haze's transmission and recover the hi-quality haze-free image. The underwater images are similar with the haze images as they are all degraded by medium. The experimental results verify that the dark channel prior can help to enhance the underwater image (Fig. 1(b)).

Objects in underwater images are generally more salient than the background, hence saliency detection can be used to "find" the objects from underwater images. We choose DRFI [2] to generate the saliency map of enhanced underwater image, since it integrates the regional contrast, regional property and regional backgroundness descriptors together to form the master saliency map, thus can deal with the challenging cases such as detecting salient objects from low-quality underwater images. For each pixel in the saliency map, the higher of the value, the more salient of it in the original image. In other words, normally, for an image, pixels which have higher values in

the corresponding saliency map belong to objects. Therefore, pixels whose values in the saliency map are greater than a certain threshold can be taken as object pixels. The Otsu method is adopted to compute the threshold from the saliency map for the determination of object information. We also observed that the image boundary is mostly background, so we take pixels in the 6-pixel wide narrow border region of the image as background information. In Fig. 1(e), the white parts are the object markers while the gray parts in the image boundary are the background markers, the remaining black parts are non-marked. These markers are used to determine the object and background information automatically instead of user-inputs.

Finally, MSRM [3] is improved for automatic segmentation by using the extracted object and background information:

1. Use Mean-Shift method to generate the oversegmented result (Fig. 1(c)) of the enhanced underwater image (Fig. 1(b)).

2. Indicate the initial marker type of each region. For each region in the oversegmented result, we regard it as background, object, or non-marker region using the extracted object and background information respectively.

3. Adopt the maximal similarity based region merging mechanism to label all the non-marker regions as either object region or background region recursively so that the objects can be extracted. The similarity between two regions is defined as the Bhattacharyya coefficient of the two corresponding region histogram vectors. If a region $A$ has the maximal similarity (within its neighborhood) with a region $B$, $A$ will be merged to $B$: $B = A \cup B$.
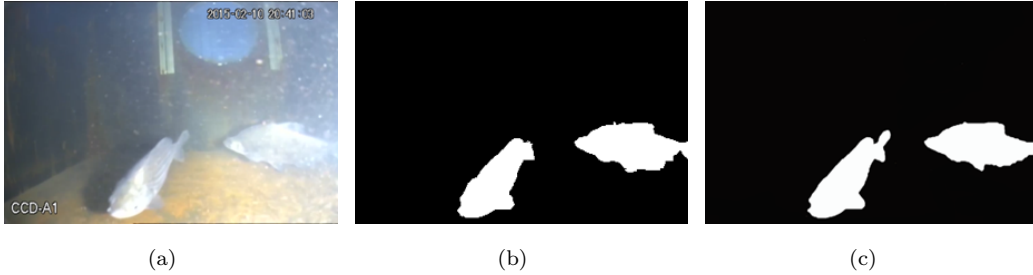


(a)　　　　　　　　　　(b)　　　　　　　　　　(c)

Fig. 2: Qualitative comparison: (a) Original image. (b) Our segmentation result. (c) Human labeled map.

It can be seen from Fig. 2(b) that it's efficient to segment the salient objects from the underwater image by the proposed method compared with the human labeled ground-truth Fig. 2(c).

# References

[1] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(12):2341–2353, 2011.

[2] Huaizu Jiang, Jingdong Wang, Zejian Yuan, Yang Wu, Nanning Zheng, and Shipeng Li. Salient object detection: A discriminative regional feature integration approach. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2083–2090. IEEE, 2013.

[3] Jifeng Ning, Lei Zhang, David Zhang, and Chengke Wu. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 43(2):445–456, 2010.