

Article

# Underwater Object Segmentation Based on Optical Features

Zhe Chen <sup>1</sup> , Zhen Zhang <sup>1</sup>, Yang Bu <sup>2</sup>, Fengzhao Dai <sup>2</sup>, Tanghuai Fan <sup>3</sup> and Huibin Wang <sup>1,\*</sup>

<sup>1</sup> College of Computer and Information, Hohai University, Nanjing 211100, China; chenzhe@hhu.edu.cn (Z.C.); zz\_hhuc@hhu.edu.cn (Z.Z.)

<sup>2</sup> Laboratory of Information Optics and Opto-Electronic Technology, Shanghai Institute of Optics and Fine Mechanics, Shanghai 201800, China; buyang@siom.ac.cn (Y.B.); fzdai@siom.ac.cn (F.D.)

<sup>3</sup> School of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, China; fantanghuai@nit.edu.cn

\* Correspondence: hbwang@hhu.edu.cn; Tel.: +86-(25)8378-7750

Received: 1 December 2017; Accepted: 9 January 2018; Published: 12 January 2018

**Abstract:** Underwater optical environments are seriously affected by various optical inputs, such as artificial light, sky light, and ambient scattered light. The latter two can block underwater object segmentation tasks, since they inhibit the emergence of objects of interest and distort image information, while artificial light can contribute to segmentation. Artificial light often focuses on the object of interest, and, therefore, we can initially identify the region of target objects if the collimation of artificial light is recognized. Based on this concept, we propose an optical feature extraction, calculation, and decision method to identify the collimated region of artificial light as a candidate object region. Then, the second phase employs a level set method to segment the objects of interest within the candidate region. This two-phase structure largely removes background noise and highlights the outline of underwater objects. We test the performance of the method with diverse underwater datasets, demonstrating that it outperforms previous methods.

**Keywords:** underwater object segmentation; optical features; level-set-based object segmentation; artificial light guidance

## 1. Introduction

In recent years, interest in underwater computer vision has increased, taking advantage of the development of artificial illumination technologies and high-quality sensors [1–4]. To an extent, this progress has addressed some challenges for close-range imaging. However, for common underwater vision tasks where sight distance is large, extremely strong attenuation can seriously reduce the performance of artificial illumination. This, combined with strong hazing and interference from sky light, make underwater objects indistinguishable [5–7]. In such cases, commonly used image features, such as color, intensity and contours, are not well characterized [8].

Two key aspects to consider for underwater artificial light are optical characteristics such as its strength and spectrum, and its collimation, which is the region covered by the light. The former is accounted for by many underwater imaging and image processing methods, such as those extending underwater visibility [9,10]. However, no existing methods determine the importance of light collimation for underwater computer vision tasks.

The collimation of artificial light can provide novel cues to aid underwater object segmentation, since artificial light often collimates directly on objects of interest. Hence, we can initially identify candidate object regions following the guidance provided by artificial light. However, besides artificial light, sky light entering the camera can make local optical environments significantly inhomogeneous. Moreover, artificial light commonly mixes with other light, such as ambient light and object radiation. These factors tend to obstruct artificial light recognition.

In order to address these challenges, this paper investigates the potential of an optical feature-based artificial light recognition method in underwater environments. The recognition results are regarded as candidates for further object segmentation. The primary novelty of our method is that it is the first investigation that uses artificial light collimation as a cue for underwater object segmentation. The approach includes optical feature extraction and a calculation technique to recognize the region of artificial light, and artificial light recognition to guide underwater object segmentation.

The remainder of this paper is organized as follows. Section 2 reviews related work on underwater object segmentation and detection. Section 3 describes the optical feature extraction and calculation method for artificial light component recognition, which is used as a guidance in segmenting the candidate region. In Section 4, the level-set-based object segmentation method is performed within candidate regions, generating final underwater object segmentation results. Section 5 presents the results obtained by the proposed artificial light recognition and underwater object segmentation methods, and compares the performance of the proposed underwater object segmentation method with state-of-the-art methods using diverse underwater images. Finally, the conclusions are presented in Section 6.

## 2. Related Works

In contrast to the significant progress in ground image object segmentation, very few methods have been proposed to successfully segment objects in underwater data. In general, underwater object segmentation can be classified into two categories: methods based on prior knowledge of any special objects, and methods that are independent of any prior experience and are used for uncertain objects.

### 2.1. Prior-Knowledge-Based Methods

The advantage of prior-knowledge-based underwater object segmentation methods lies in their robust ability to identify special objects. For example, Yu et al. used a pre-defined color feature to detect and segment the artificial underwater landmarks (AULs) in an underwater vision system. This method has proven to be reliable for navigation when the camera is located near underwater objects [11]. Lee et al. designed a docking system using a LED ring with five large lights. Aiming to accurately track the position of an autonomous underwater vehicle, the structure of the docking-marks was carefully pre-designed and their brightness carefully adjusted. This information was input to a camera system as prior knowledge for performing detection of the docking marks and subsequent segmentation [12]. Besides the features given by these pre-designed man-made objects, many types of prior knowledge can be introduced via machine learning methods. A typical sample is the use of Markov random fields in vision control systems [13,14]. In this method, the MRF model was used as a preprocessor to calculate the relationship between raw image data and corresponding corrected images, and the output of model was applied for detecting underwater coral objects. Additionally, many methods demonstrate the advantage of prior structural features for underwater object segmentation. Negre et al. presented a scale and rotationally invariant feature for self-similar landmark (SSL) detection. From the results, it was evident that structural features were more robust for underwater object detection than color features [15]. Maire et al. extracted rectangular Haar wavelet features and the rotation of integral images to detect uniquely identifiable objects for automatic docking control [16].

### 2.2. Methods without Prior Information

In contrast to prior-knowledge-based approaches, the second type of underwater object segmentation methods is independent of prior experience. This type of technique generally focuses on uncertain aquatic objects using a two-phase structure to overcome the challenge of variable underwater optical environments. For these systems, a preprocessor is commonly introduced to perform image enhancement and as a second stage for object segmentation. In principle, this strategy can better adapt to underwater environments. Lee et al. proposed an updated underwater image restoration method to correct the object colors. The contribution of image preprocessing to the underwater object

detection was demonstrated by comparing the results before and after image preprocessing [17]. Kim et al. fused color correction and template matching to segment well-designed landmarks, and its performance was demonstrated by both indoor test water and outdoor natural pools [18]. Moreover, the two-phase structure in some studies includes a coarse-to-fine strategy. For example, Edgington et al. employed the Itti saliency detection method to first identify candidates. However, the usage of an Itti model in underwater environments is somewhat questionable. Features such as intensity, color, and orientation used by the Itti model are nominally suited for the typical ground scenes, while these features are unreliable in underwater conditions [19]. Further, the saliency detection method employed by Rizzini as a preprocessor can initially identify the candidate region of the objects. Object segmentation in the successive phase was achieved by using a low-pass filter [20]. Chuang et al. used the phase Fourier transform (PFT)-based saliency detection method to detect patches of underwater objects. The texture features were then extracted to classify fish images [21]. Zhu et al. integrated mean-shift oversegmentation and saliency detection methods to completely segment the underwater object. Image restoration was also used in this method for correcting underwater samples [22]. Recently, Chen et al. proposed a novel depth-feature-based region of interest (ROI) detection method. The ROI detection result is further corrected using Otsu's method [23] for underwater object detection. From the presented results, the depth information was key for ROI detection; however, the method was vulnerable in environments where artificial and sky light seriously affect depth estimation [24]. Differences between the previously proposed work [24] and the method presented in this paper can be concluded in three aspects. Firstly, the previous underwater object detection method was achieved by a multi-feature fusion strategy and independent of any prior about underwater environments. In contrast, the method in this paper is totally based on a novel optical discipline given by the knowledge of underwater imaging: artificial light commonly collimates the underwater ROI (region of interest), thus can guide underwater object segmentation process. Secondly, the depth information in previous work was used to identify the ROI, which can only be achieved in relatively homogeneous environments. However, the underwater artificial light makes underwater optical environments inhomogeneous, which will block the usage of the previous method. In contrast, the method proposed in this paper is adaptive to inhomogeneous environments and has the ability to segment underwater objects using the cue originated from the artificial light. At last, the level set model is introduced in this paper to generate accurate underwater object segmentation results. However, the level set model is not taken advantage in the previous work [24], since the previous work aims to detect the existence of underwater objects. Hence, a simple OTSU model is enough and employed in the previous work.

### *2.3. Comparison to Previous Works*

The work described above shows the success of coarse-to-fine strategies for underwater object segmentation and is an important motivation for research presented in this paper. However, in contrast to previously proposed methods, our approach is not specialized for any special objects and is independent of prior knowledge. Moreover, the novel cue provided by artificial light guidance is considered in our first step rather than the commonly used image enhancement or saliency detection methods.

### *2.4. Proposed Method*

The framework of the proposed method is shown in Figure 1. In the first phase, various optical features including the global intensity contrast, channel variation, intensity-position, and red channel contrast are extracted from the underwater images. These optical features are then calculated in a discrimination model to recognize the distribution of artificial light, which then guides the candidate region segmentation in the fourth stage, which is achieved by the simple OTSU method. Finally, a level-set method is applied to the candidate region, generating the final results for underwater object segmentation.

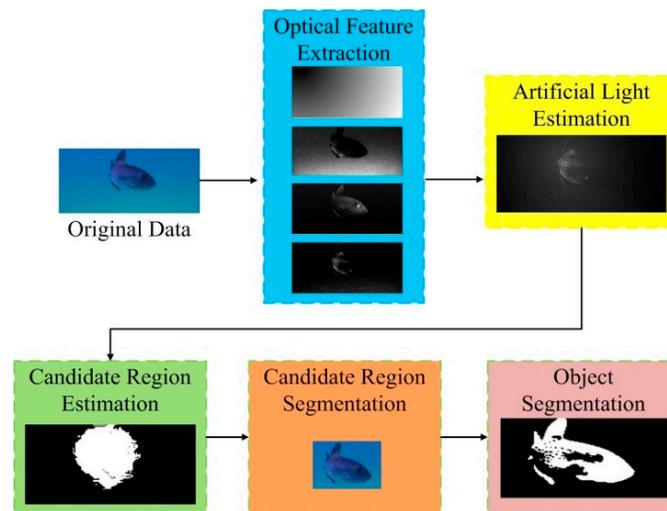


Figure 1. Framework of the proposed underwater object segmentation method.

### 3. Optical Feature Extraction and Artificial Light Recognition

Artificial light can provide a novel cue for underwater object segmentation; however, its recognition using underwater optical features is a core issue. Optically, artificial light propagating through water presents many characteristics different from other light. The optical attenuation and light scattering as we know are two main factors which affect underwater light [25–30], as follows:

- (i) *Optical attenuation.* The optical attenuation will cause a significant degeneration of underwater light power [25]. Besides, this attenuation effect exponentially changes with the sight distance. Since the distance between artificial lumination and reflectors, and the distance between reflectors and the camera are both limited, the degeneration of underwater artificial light is relatively small comparing to underwater natural light [26]. This will generate a higher intensity in local regions which at the same time presents a significant contrast to the surrounding regions. Moreover, the underwater light attenuation factor is also wavelength-selective [26]. For natural light in water, the red channel is extremely weak, which can be partly compensated by artificial light [27,28]. Hence, the red channel intensity of underwater artificial light is relatively high and has a large contrast to the natural light component.
- (ii) *Light scattering.* The underwater light scattering factor will cause a halo effect in the region collimated by the artificial light [29,30]. In this region, the intensity of any points is inversely proportional to their distances from the brightest point. However, this halo effect is not presented in other regions due to the homogeneousness of underwater natural light.

Based on these characteristics above, the discriminative features of the artificial light can be extracted into four aspects. First, artificial light is more intense than other optical inputs, and consequently, its average intensity in the artificial light region is higher in the entire image. Second, the propagation distance of artificial light is shorter than other optical inputs, making artificial light channel variation smaller and the channels of the reflected artificial light more balanced. Third, a halo effect appears surrounding the collimation point. The intensity distribution within the artificial light region is exponentially modified with a radius, holding an intensity-position mode. Finally, as compared to the natural optical components, the red channel of artificial light is significantly more intense, which can cause its emergence in the artificial light region in the red channel.

#### 3.1. Underwater Optical Feature Extraction

In general, four types of optical features are extracted in this paper: the global intensity contrast, channel variation, intensity position, and red channel contrast. The absolute difference in the intensity

values is calculated to scale the intensity contrast at points  $x$  and  $y$ . The global intensity contrast at point  $x$  is the summation of these point-to-point differences as follows:

$$C_x^i = \sum_{\forall y \in N} C(I_x^i, I_y^i) = \sum_{\forall y \in N} \|I_x^i - I_y^i\| \quad (1)$$

where  $C(I_x^i, I_y^i)$  is the absolute difference between points  $x$  and  $y$  in the intensity value,  $N$  denote the points within underwater images, and the super scrip  $i$  is the label of the intensity.

The channel variation feature is computed as the summation of the differences between the channel intensity and the average intensity, as follows:

$$V_x^c = \sqrt{\sum V(I_x^c, I_x^i)} = \sqrt{(I_x^r - I_x^i)^2 + (I_x^g - I_x^i)^2 + (I_x^b - I_x^i)^2} \quad (2)$$

where  $V(I_x^c, I_x^i)$  is the variance of the point  $x$  in the  $r * g * b$  color space.

The intensity-position relation is scaled by the distance between point  $x$  and the point which has the largest intensity in the whole image, as follows:

$$D_x^d = \exp(D(\mathbf{x}, \mathbf{m})) = \exp\left(\sqrt{(\xi_1 - \xi_2)^2 + (\gamma_1 - \gamma_2)^2}\right) \quad (3)$$

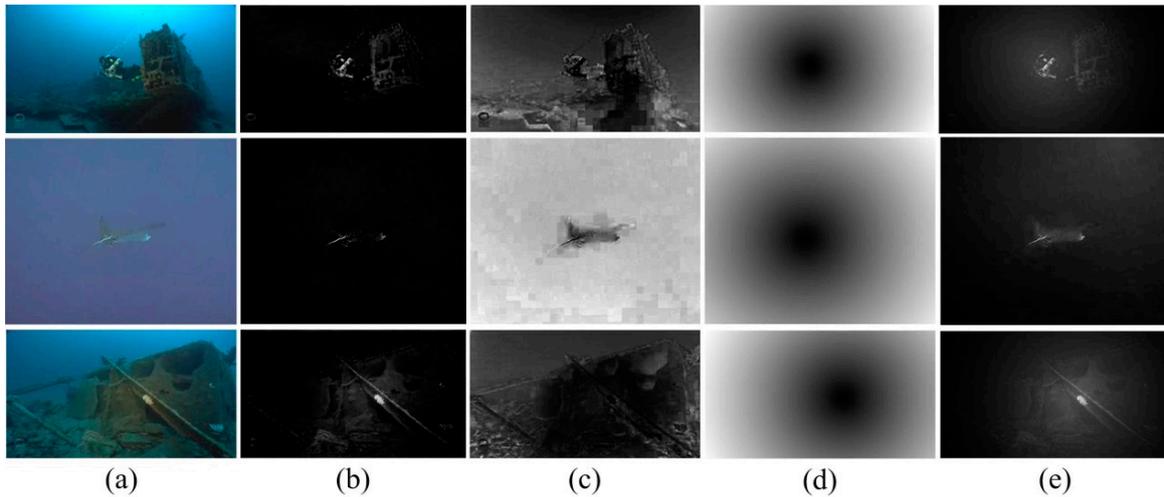
where  $D(\mathbf{x}, \mathbf{m})$  is the Euclidean distance between point  $x$  and point  $m$  which has the largest intensity in the whole image,  $\mathbf{x} = [\xi_1, \gamma_1]$ ,  $\mathbf{m} = [\xi_2, \gamma_2]$  are the spatial coordinates of the points  $x$  and  $m$ .

The red channel contrast at point  $x$  is the summation of point-to-point differences in the red channel, as follows:

$$C_x^r = \sum_{\forall y \in N} C(I_x^r, I_y^r) = \sum_{\forall y \in N} \|I_x^r - I_y^r\| \quad (4)$$

where  $C(I_x^r, I_y^r)$  is the absolute difference between points  $x$  and  $y$  in the red channel, and the super scrip  $r$  is the label of the red channel.

The optical feature extraction results are shown in Figure 2.



**Figure 2.** Underwater optical features. (a) original underwater image; (b) global intensity contrast; (c) channel variation; (d) intensity-position; (e) red channel contrast.

### 3.2. Artificial Light Recognition

Artificial light recognition follows four optical principles which are given or concluded by prior knowledge about underwater optical characteristics [25–30], as follows:

1. The inverse relationship between the global intensity contrast and the intensity position relation features: in the artificial light region, higher-contrast points are closer to the largest intensity point and vice versa.
2. The correspondence between the global intensity contrast and the red channel contrast features: within the artificial light region, an increase in the global intensity contrast corresponds to an increasing red channel contrast.
3. The inverse relationship between the global intensity contrast and the channel variation features: points with larger global intensity contrast have lower channel variation in the artificial light region, and vice versa.
4. The correspondence between the intensity-position relation and the channel variation features: within the artificial light region, the lower channel variation must be located in points that are closer to the maximum intensity point.

In this section, the discrimination function is proposed following the simulation of principles 1–4. Here, we generally use two-dimensional correlation [31] to scale the inverse relationship and correspondence between different optical features. According to the principle 1, the inverse relationship between the global intensity contrast and the intensity position relation can be established as  $\text{corr2}(C^i, (1 - D^d))$ , while according to the principle 2, the correspondence between the global intensity contrast and the red channel contrast can be calculated as  $\text{corr2}(C^i, C^r)$ . Similarly, the inverse relationship between the global intensity contrast and the channel variation (principle 3), and the correspondence between the intensity-position relation and the channel variation features (principle 4) can be calculated as  $\text{corr2}(C^i, (1 - V^c))$  and  $\text{corr2}(D^d, V^c)$ , respectively. Large values of two-dimensional correlations will be obtained with significantly strong inverse relationship and correspondence between optical features. If all the two-dimensional correlations are of large values, the artificial light collimation is recognized in corresponding regions. Hence, the comprehensive discrimination function is formulated as the multiplicative of the correlation calculations as

$$S = \text{corr2}(C^i, (1 - D^d)) \times \text{corr2}(C^i, C^r) \times \text{corr2}(C^i, (1 - V^c)) \times \text{corr2}(D^d, V^c) \quad (5)$$

where  $S$  is the discrimination function for the artificial light region;  $C^i$ ,  $D^d$ ,  $C^r$ , and  $V^c$  are the matrices for the  $C_x^i$ ,  $D_x^d$ ,  $C_x^r$ , and  $V_x^c$ , respectively; and  $\text{corr2}()$  is used to calculate the two-dimensional correlation between matrices. According to Equation (5), if and only if all correlations (including inverse relationship and correspondence) reach significant high values,  $S$  will be of a high value which identifies the region covered by the artificial light. However, the divergence between correlations will result in a relative low value of  $S$ . In this case, the high intensity in local regions or the red channel may be caused by the sky light or bioluminescence which cannot correctly guide the underwater object segmentation. If all correlations are of extreme low values, the underwater environment may be seriously affected by any sources with specially designed spectrum.

We apply a threshold function to recognize the region of the artificial light as follows:

$$L_{\text{Artificial Light}} = \begin{cases} 1 & \text{if } S > T \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $T$  is a threshold.

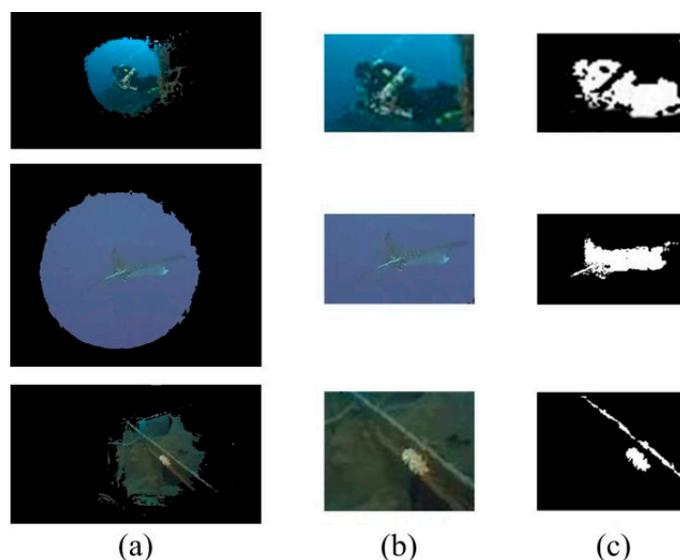
According to Equation (6), if  $L_{\text{Artificial Light}} = 1$ , the artificial light exists, and the corresponding strength of the artificial light is calculated as

$$W = (C^i + C^r - D^d - V^c) \times L_{\text{Artificial Light}} \quad (7)$$

where  $W$  is the strength of the artificial light.

### 3.3. Candidate Region Identification

Here, we use Otsu's method to segment the map given by the artificial light recognition result  $W$ , extracting the candidate object regions [23]. The rationale behind the application of this method is twofold. Otsu's method adapts well to processing the  $W$  maps, since the candidate region is distinguishable from the background in the gray histogram. However, Otsu's method is efficient and linear with the size of the maps. Samples of segmentation results are shown in Figure 3a. Within the region given by Otsu's method, the largest interior rectangle window is given to further identify the candidate regions, as shown in Figure 3b. In addition, Figure 3c shows the object segmentation result in the candidate regions. The contribution of the artificial light guidance can be demonstrated by Figure 3, that the artificial light region (Figure 3a) can correctly identify the candidate regions of underwater objects (Figure 2). Within the largest interior rectangle window of the artificial light region (Figure 3b), the object segmentation results (Figure 3c) can correctly occupy the ground truth and largely remove background noises.



**Figure 3.** Underwater object candidate regions. (a) artificial light recognition; (b) candidate object region segmentation; (c) our object segmentation. The original images are shown in the first column of Figure 2.

## 4. Level Set Based Underwater Object Segmentation

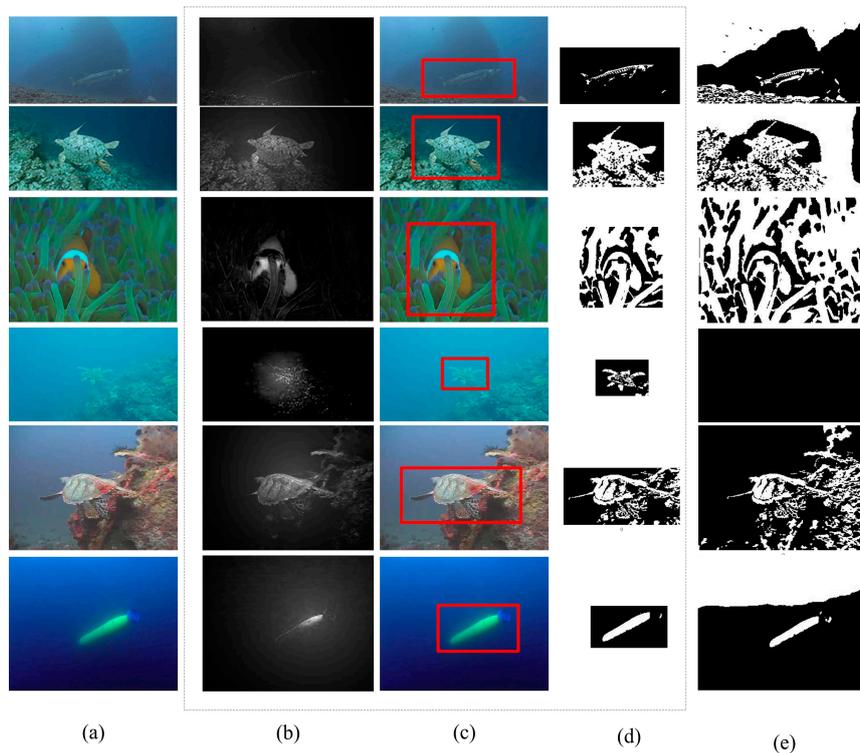
In the post-processing phase, our underwater object segmentation results are generated by performing the level-set-based method in candidate regions. Specifically, we employ here a method using parametric kernel graph cuts [32]. Here, graph cuts that minimize the loss function in a kernel-introduced space are established as

$$\psi_{\kappa}(\{u_l\}, v) = \sum_{l \in L} \sum_{p \in R_l} (\phi(u_l) - \phi(I_p))^2 + \alpha \sum_{\{p,q\} \in N} r(v(p), v(q)) \quad (8)$$

where  $\psi_{\kappa}(\{u_l\}, v)$  is the measurement by the kernel-induced, non-Euclidean distances between the observations and the regions' parameters,  $u_l$  is the piecewise-constant model parameter of region,  $I_p$  are the original image parameters,  $r(v(p), v(q))$  is a smoothness regularization function established by the truncated squared absolute difference, and  $\alpha$  is a positive factor.

This loss function is then minimized by an iterative two-stage method. In the first stage, the labeling is fixed and the loss function is optimized with respect to the parameter  $\{u_l\}$ . Then, an optimal labeling search is performed using the optimal parameter provided by the first

stage. The level-set-based underwater object segmentation results with and without artificial light guidance are shown in Figure 4. The first column (Figure 4a) shows the original underwater images. The second column (Figure 4b) presents the artificial light recognition results, and the detected candidate regions are presented with a red rectangle in the third column (Figure 4c). The fourth column (Figure 4d) shows the results of the level-set-based method under the guidance of the artificial light, and in the last column (Figure 4e) the object segmentation results without artificial light guidance are presented. We find that the artificial light guidance significantly contributes to background removal, while without this prior guidance the segmented objects are indistinguishable and seriously overlap with background noise.



**Figure 4.** Phases of underwater object segmentation and comparison of results. (a) original underwater image; (b) artificial light recognition; (c) candidate object region segmentation; (d) our object segmentation; (e) object segmentation without artificial guidance.

## 5. Experimental Results

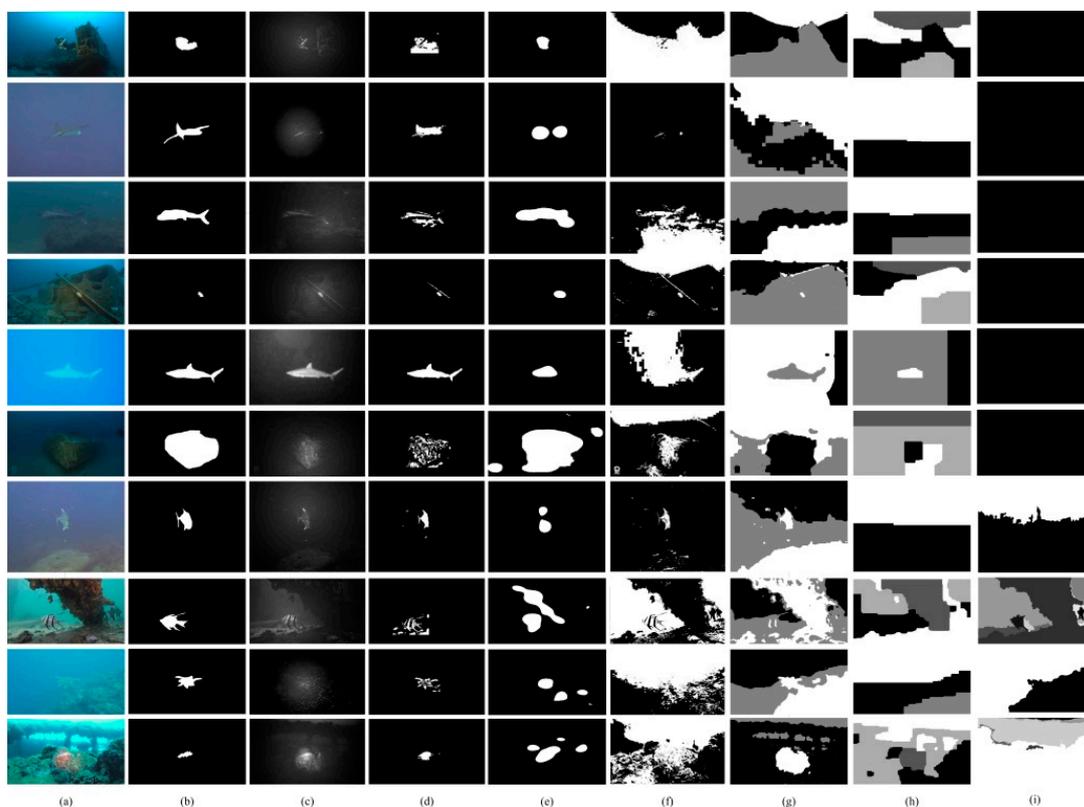
### 5.1. Experiment Setup

The performance of the proposed method is demonstrated through comparison with five state-of-the-art saliency models. These models have shown excellent performance in various datasets and have been frequently cited in the literature. The compared methods include the updated saliency detection (HFT [33]) and statistical models (BGGMM [34], FRGMM [35]), and optimization (Kernel\_GraphCuts [32], ROISEG [36]). Some of these methods have been successfully tested with underwater samples. Since our method is independent of training, new machine learning-based methods are not included in our experiment. The code for the baseline methods was downloaded from the websites provided by their authors.

### 5.2. Dataset

Underwater images available on YouTube were collected to establish the benchmark for the experimental evaluations [37–45]. These data were collected by marine biologists, tourists and

autonomous underwater vehicles. More than 200 images were included in our underwater benchmarks. Some of these samples include participants located in coast areas where the underwater visibility is high while object appearances are seriously disturbed by large amount of optical noises [37,40–42,44] (i.e., the first, third, fourth eighth, ninth and the last rows in Figure 5). There are also samples presenting scenes under open seas where underwater objects are not clear due to the large sight distance [38,43,45] (i.e., the second, fifth and the seventh rows in Figure 5). Moreover, our dataset includes underwater images collected in environments with very weak natural light [39] (i.e., the sixth row in Figure 5). All test images were acquired by cameras with fixed focal length lens appended with underwater LED artificial illumination equipment. The highest brightness of these underwater lumination devices range from 5200 lm to 30,000 lm. The ground truth of each image was labeled by 10 experienced volunteers who major in researches of computer vision. Experimental samples used in this paper can be downloaded from our website [46]. The largest human visual contrast is used as the principle to segment underwater objects from background and the averaged results were deemed to be the ground truth. The images in this dataset differed in many aspects, including participants, context, and the optical inputs present. The diversity of these images ensures a comprehensive and fair evaluation of the methods. The threshold  $T$  for the decision function in Equation (6) is defined as 0.6 for all experiments.



**Figure 5.** Result comparison. (a) underwater image; (b) ground-truth; (c) artificial light recognition; (d) our method; (e) HFT; (f) BGGMM; (g) FRGMM; (h) Kernel\_GraphCuts; (i) ROISEG.

### 5.3. Evaluation Methods

The PASCAL criterion [47] was used to evaluate the overlap of the detection results and ground truth

$$c = \frac{\Omega' \cap \Omega}{\Omega' \cup \Omega} \quad (9)$$

where  $\Omega'$  is the detection;  $\Omega$  is the labeled ground-truth; and  $c$  is degree of the overlap. Obviously, the larger the value of  $c$ , the more robust the algorithm. Moreover, the performance of our method was

evaluated by six criteria [17], the precision (Pr), similarity (Sim), true positive rate (TPR), F-score (FS), false positive rate (FPR), and percentage of wrong classifications (PWC).

$$\begin{aligned} \text{Pr} &= \frac{tp}{tp+fp}, \text{TPR} = \frac{tp}{tp+fn}, \text{Fs} = 2 \times \frac{\text{Pr} \times \text{TPR}}{\text{Pr} + \text{TPR}} \\ \text{Sim} &= \frac{tp}{tp+fp+fn}, \text{FPR} = \frac{fp}{fp+tn}, \text{PWC} = 100 \times \frac{fn+fp}{tp+tn+fp+fn} \end{aligned} \quad (10)$$

where  $t_p$ ,  $t_n$ ,  $f_p$  and  $f_n$  denote the numbers of the true positives, true negatives, false positives, and false negatives, respectively.

#### 5.4. Results

Figure 5 shows the experimental results of our artificial light recognition, our object segmentation methods and five compared methods with ten typical test images. The optical environments of the test underwater image (Figure 5a) are diverse. The participants in the underwater images are different. Moreover, random image noise is obvious in these images. Together, these factors impose a serious challenge for object segmentation.

The first column in Figure 5 shows the original images; the second column is the ground-truth; the third and fourth columns respectively present the results of our artificial light recognition and our object segmentation methods. The last five columns respectively present the results given by the five compared methods. From the results, the most competitive performance is demonstrated by our method which can correctly segment underwater objects and remove the complicated background. The most comparable results are presented in the fifth column, provided by the HFT method. From the results, the HFT method works well in relatively pure scenes, such as the samples in the second and third rows. However, its performance is significantly poorer in scenes including textural backgrounds, such as the samples in the last three rows. From the results shown in the sixth and seventh columns, the statistical model can detect the objects of interest, but do not have the capability to recognize objects from the background. Moreover, the statistical models are also vulnerable to optical noise. As the result, the detected objects of interest in the fourth and fifth columns seriously overlapped with the background and the optical noise. Generally, without the guidance of the artificial light, the level-set-based method could not adapt to the underwater scenes, providing insignificant and vague results. Recall that the second phase of our proposed method is established by the level-set-based method (Kernel\_GraphCuts), so that the importance of the artificial light guidance can be highlighted by comparing the results in the fourth and the eighth columns. The objective of these level-set-based methods is to segment the image into discrete regions that are homogeneous with respect to some image feature. However, for underwater images, hazing and ambient scattered light significantly inhibit the emergence of the objects of interest. Considering the entire image, the margin between the objects and background is relatively indistinguishable, while other transitions, such as those between sky and ambient parallel light, may be mistaken as regional boundaries. Alternatively, our method uses the artificial light collimation to guide the object segmentation process. The level-set-based image segmentation processing in our method focuses on the candidate regions that present a relatively higher contrast between objects and the surrounding background. In this case, the object contours are the most optimal boundaries with which to minimize the loss function (Equation (8)). This is the factor underlying the success for our underwater object segmentation method.

To further examine the quantitative performance of our method, Table 1 shows the differences in the average performance of the different methods on the 200 underwater images. The PASCAL criterion ( $\bar{C}$ ), the precision (Pr), true positive rate (TPR), F-score (FS), similarity (Sim), false positive rate (FPR), and percentage of wrong classifications (PWC) are used here as the quantitative criteria. From the results, our method performed best, scoring the first in four criteria and the second in three criteria. Specifically, the highest scores of the criteria  $\bar{C}$ , Pr, TPR and Sim indicate that our segmentation results can exactly occupy bodies of underwater objects. The robustness against the background noises is demonstrated by the criteria FS, FPR and PWC. According to these three criteria,

the best performance is achieved by the HFT method, which, however, mistakes parts of the objects as the background. As a result, there are many holes existing in the results given by the HFT method (Figure 5). Although the second rather than the first best performance is given by our method referred to the criteria FS, FPR and PWC, our method can remove most of background noises, demonstrating the contribution of our artificial light guidance for underwater object segmentation. Generally, the best comprehensive performance is achieved by our method. However, the other four compared methods (BGGMM, FRGMM, Kernel\_GraphCuts and ROISEG), as shown in Table 1, scored worse than the proposed method. This indicates that these four methods cannot well adapt to the underwater environments and the results given by these methods are insignificant, which is in accordance with the qualitative results shown in Figure 5.

**Table 1.** Average performance comparison of HFT, BGGMM, FRGMM, Kernel\_GraphCuts, ROISEG and our method using diverse underwater image data.

Method	$\bar{C}$	$P_r$	TPR	FS	Sim	FPR	PWC
HFT + OTSU	0.5628	0.5426	0.7912	0.5858	0.4352	0.0331	3.5898
BGGMM	0.3075	0.3076	0.7480	0.2987	0.2071	0.2329	23.3065
FRGMM	0.3708	0.3387	0.8475	0.4109	0.3150	0.1253	12.5406
Kernel_GraphCuts	0.1178	0.2345	0.7198	0.1782	0.1156	0.2259	23.9979
ROISEG	0.1042	0.3509	0.2804	0.0906	0.0656	0.1219	11.8211
Our method	0.7164	0.6327	0.7968	0.5162	0.4479	0.0355	7.1233

## 6. Conclusions

The work presented in this paper is a novel investigation of the usage of optical features for underwater computer vision. Various features, such as the global intensity contrast, channel variation, intensity position, and the red channel contrast, were extracted and jointly used to establish a decision function for artificial light recognition. The recognition results provide strong guidance for underwater object segmentation. This new method overcomes challenges caused by difficult underwater optical environments, such as light attenuation and hazing. The outstanding performance of our method was demonstrated by experiments on diverse underwater images.

Although the evaluation results presented in our paper preliminarily demonstrate the performance of the artificial light guidance for underwater object segmentation, the robustness of our method is required to be tested on more underwater images. Moreover, this work is a starting point that sheds light on the contribution of optical models for underwater computer vision tasks. In further work, we aim to introduce additional prior optical knowledge into state-of-the-art computer vision methods by model reconstruction, establishing more elaborate models for underwater optical environments.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China (No. 61501173, 61563036, 61671201), the Natural Science Foundation of Jiangsu Province (No. BK20150824), the Fundamental Research Funds for the Central Universities (No. 2017B01914), and the Jiangsu Overseas Scholar Program for University Prominent Young & Middle-aged Teachers and Presidents.

**Author Contributions:** Zhe Chen designed the optical feature extraction and object segmentation methods, and drafted the manuscript. Huibin Wang provided a theoretical analysis of the proposed method, Zhen Zhang organized the experiments and performed data analysis, and Yang Bu and Fengzhao Dai collected the underwater data and evaluated the performance of our proposed method. Tanghuai Fan provided funding support and the equipment for experiments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kocak, D.M.; Dalgleish, F.R.; Caimi, F.M.; Schechner, Y.Y. A focus on recent developments and trends in underwater imaging. *Mar. Technol. Soc. J.* **2008**, *42*, 52–67. [[CrossRef](#)]
2. Fan, S.; Parrott, E.P.; Ung, B.S.; Pickwell-MacPherson, E. Calibration method to improve the accuracy of THz imaging and spectroscopy in reflection geometry. *Photonics Res.* **2016**, *4*, 29–35. [[CrossRef](#)]

3. Zhao, Y.; Chen, Q.; Gu, G.; Sui, X. Simple and effective method to improve the signal-to-noise ratio of compressive imaging. *Chin. Opt. Lett.* **2017**, *15*, 46–50. [[CrossRef](#)]
4. Le, M.; Wang, G.; Zheng, H.; Liu, J.; Zhou, Y.; Xu, Z. Underwater computational ghost imaging. *Opt. Express* **2017**, *25*, 22859–22868. [[CrossRef](#)] [[PubMed](#)]
5. Ma, L.; Wang, F.; Wang, C.; Wang, C.; Tan, J. Monte Carlo simulation of spectral reflectance and BRDF of the bubble layer in the upper ocean. *Opt. Express* **2015**, *23*, 74–89. [[CrossRef](#)] [[PubMed](#)]
6. Satat, G.; Tancik, M.; Gupta, O.; Heshmat, B.; Raskar, R. Object classification through scattering media with deep learning on time resolved measurement. *Opt. Express* **2017**, *25*, 66–79. [[CrossRef](#)] [[PubMed](#)]
7. Akkaynak, D.; Treibitz, T.; Shlesinger, T.; Loya, Y.; Tamir, R.; Iluz, D. What is the space of attenuation coefficients in underwater computer vision? In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
8. Fang, L.; Zhang, C.; Zuo, H.; Zhu, J.; Pang, L. Four-element division algorithm to focus coherent light through a turbid medium. *Chin. Opt. Lett.* **2017**, *5*, 1–5.
9. Mullen, L.; Lee, R.; Nash, J. Digital passband processing of wideband-modulated optical signals for enhanced underwater imaging. *Appl. Opt.* **2016**, *55*, 18–24. [[CrossRef](#)] [[PubMed](#)]
10. Zhang, R.; Zhang, W.; He, C.; Zhang, Y.; Song, J.; Xue, C. Underwater Imaging Using a  $1 \times 16$  CMUT Linear Array. *Sensors* **2016**, *16*, 312. [[CrossRef](#)] [[PubMed](#)]
11. Yu, S.C.; Ura, T.; Fujii, T.; Kondo, H. Navigation of autonomous underwater vehicles based on artificial underwater landmarks. In Proceedings of the OCEANS, Honolulu, HI, USA, 5–8 November 2001.
12. Lee, P.M.; Jeon, B.H.; Kim, S.M. Visual servoing for underwater docking of an autonomous underwater vehicle with one camera. In Proceedings of the OCEANS, San Diego, CA, USA, 22–26 September 2003.
13. Dudek, G.; Jenkin, M.; Prahacs, C.; Hogue, A.; Sattar, J.; Giguere, P.; Simhon, S. A visually guided swimming robot. In Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2–6 August 2005.
14. Sattar, J.; Dudek, G. Robust servo-control for underwater robots using banks of visual filters. In Proceedings of the IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009.
15. Negre, A.; Pradalier, C.; Dunbabin, M. Robust vision-based underwater target identification and homing using self-similar landmarks. In *Field and Service Robotics*; Springer: Berlin/Heidelberg, Germany, 2008.
16. Maire, F.D.; Prasser, D.; Dunbabin, M.; Dawson, M. A vision based target detection system for docking of an autonomous underwater vehicle. In Proceedings of the 2009 Australasian Conference on Robotics and Automation, University of Sydney, Sydney, Australia, 2–4 December 2009.
17. Lee, D.; Kim, G.; Kim, D.; Myung, H.; Choi, H. Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean Eng.* **2012**, *48*, 59–68. [[CrossRef](#)]
18. Kim, D.; Lee, D.; Myung, H.; Choi, H.T. Artificial landmark-based underwater localization for AUVs using weighted template matching. *Intell. Serv. Robot.* **2014**, *7*, 175–184. [[CrossRef](#)]
19. Edgington, D.R.; Salamy, K.A.; Risi, M.; Sherlock, R.E.; Walther, D.; Koch, C. Automated event detection in underwater video. In Proceedings of the OCEANS, San Diego, CA, USA, 22–26 September 2003.
20. Rizzini, D.L.; Kallasi, F.; Oleari, F.; Caselli, S. Investigation of vision-based underwater object detection with multiple datasets. *Int. J. Adv. Robot. Syst.* **2015**, *12*, 1–13. [[CrossRef](#)]
21. Chuang, M.C.; Hwang, J.N.; Williams, K. A Feature Learning and Object Recognition Framework for Underwater Fish Images. *IEEE Trans. Image Proc.* **2016**, *25*, 1862–1872. [[CrossRef](#)] [[PubMed](#)]
22. Zhu, Y.; Chang, L.; Dai, J.; Zheng, H.; Zheng, B. Automatic object detection and segmentation from underwater images via saliency-based region merging. In Proceedings of the OCEANS, Shanghai, China, 10–13 April 2016.
23. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
24. Chen, Z.; Zhang, Z.; Dai, F.; Bu, Y.; Wang, H. Monocular Vision-Based Underwater Object Detection. *Sensors* **2017**, *17*, 1784. [[CrossRef](#)] [[PubMed](#)]
25. Chiang, J.Y.; Chen, Y.C. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Proc.* **2012**, *21*, 1756–1769. [[CrossRef](#)] [[PubMed](#)]
26. Duntley, S.Q. Light in the sea. *JOSA* **1963**, *53*, 214–233. [[CrossRef](#)]
27. Galdran, A.; Pardo, D.; Picón, A.; Alvarez-Gila, A. Automatic red-channel underwater image restoration. *J. Vis. Commun. Image Represent.* **2015**, *26*, 132–145. [[CrossRef](#)]

28. Jaffe, J.S.; Moore, K.D.; McLean, J.; Strand, M.P. Underwater optical imaging: Status and prospects. *Oceanography* **2001**, *14*, 66–76. [CrossRef]
29. Hou, W.W. A simple underwater imaging model. *Opt. Lett.* **2009**, *34*, 2688–2690. [CrossRef] [PubMed]
30. Ma, Z.; Wen, J.; Zhang, C.; Liu, Q.; Yan, D. An effective fusion defogging approach for single sea fog image. *Neurocomputing* **2016**, *173*, 1257–1267. [CrossRef]
31. Two-Dimensional Correlation Model. Available online: <https://cn.mathworks.com/help/signal/ref/-xcorr2.html> (accessed on 11 May 2017).
32. Salah, M.B.; Mitiche, A.; Ayed, I.B. Multiregion image segmentation by parametric kernel graph cuts. *IEEE Trans. Image Proc.* **2011**, *20*, 545–557. [CrossRef] [PubMed]
33. Li, J.; Levine, M.D.; An, X.; Xu, X.; He, H. Visual saliency based on scale-space analysis in the frequency domain. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 996–1010. [CrossRef] [PubMed]
34. Nguyen, T.M.; Wu, Q.J.; Zhang, H. Bounded generalized Gaussian mixture model. *Pattern Recognit.* **2014**, *47*, 32–42. [CrossRef]
35. Nguyen, T.M.; Wu, Q.J. Fast and robust spatially constrained gaussian mixture model for image segmentation. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, *23*, 21–35. [CrossRef]
36. Donoser, M.; Bischof, H. ROI-SEG: Unsupervised color segmentation by combining differently focused sub results. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 17–22 June 2007.
37. Bubble Vision. Underwater Vision. Available online: <http://www.youtube.com/user/bubblevision> and <http://www.youtube.com/watch?v=NKmc5dIVSRk&hd=1> (accessed on 5 November 2016).
38. Valeo Films Inc. Underwater Vision. Available online: <https://www.youtube.com/watch?v=P7257ozFHkI> (accessed on 19 August 2008).
39. Monterey Bay Aquarium Research Institute. Underwater Vision. Available online: <https://www.youtube.com/watch?v=1T70Ev2AYs> (accessed on 13 November 2009).
40. Divertanaboo. Underwater Vision. Available online: <https://www.youtube.com/watch?v=Pt4ib8VIFVA> (accessed on 6 May 2010).
41. Virkof23. Underwater Vision. Available online: <https://www.youtube.com/watch?v=2AzEeh87Z38> (accessed on 16 July 2008).
42. SASSub Aviator Systems. Underwater Vision. Available online: [https://www.youtube.com/watch?v=a9\\_iVF4EA-o](https://www.youtube.com/watch?v=a9_iVF4EA-o) (accessed on 8 June 2012).
43. VideoRay Remotely Operated Vehicles. Underwater Vision. Available online: <https://www.youtube.com/watch?v=BNq1v6KCANo> (accessed on 27 October 2010).
44. Bubble Vision. Underwater Vision. Available online: [https://www.youtube.com/watch?v=kK\\_hJZo-7-k](https://www.youtube.com/watch?v=kK_hJZo-7-k) (accessed on 24 December 2012).
45. Tigertake0736. Underwater Vision. Available online: <https://www.youtube.com/watch?v=NKmc5dIVSRk> (accessed on 1 February 2010).
46. Chen, Z. Underwater Object Detection. Available online: <https://github.com/9434011/underwater-object-detection> (accessed on 9 January 2018).
47. Everingham, M.; Van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [CrossRef]

