

## Research Article

# Underwater Object Recognition Using Transformable Template Matching Based on Prior Knowledge

Jianjiang Zhu,<sup>1</sup> Siqian Yu,<sup>2,3</sup> Zhi Han ,<sup>2</sup> Yandong Tang,<sup>2</sup> and Chengdong Wu<sup>3</sup>

<sup>1</sup>School of Electrical & Automatic Engineering, Changshu Institute of Technology, Changshu 215500, China

<sup>2</sup>State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

<sup>3</sup>Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Zhi Han; hanzhi@sia.cn

Received 17 September 2018; Revised 28 December 2018; Accepted 10 January 2019; Published 3 February 2019

Academic Editor: Oscar Reinoso

Copyright © 2019 Jianjiang Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Underwater object recognition in sonar images, such as mine detection and wreckage detection of a submerged airplane, is a very challenging task. The main difficulties include but are not limited to object rotation, confusion from false targets and complex backgrounds, and extensibility of recognition ability on diverse types of objects. In this paper, we propose an underwater object detection and recognition method using a transformable template matching approach based on prior knowledge. Specifically, we first extract features and construct a template from sonar video sequences based on the analysis of acoustic shadows and highlight regions. Then, we identify the target region in the objective image by fast saliency detection techniques based on FFT, which can significantly improve efficiency by avoiding an exhaustive global search. After affine transformation of the template according to the orientation of the target, we extract normalized gradient features and calculate the similarity between the template and the target region, which can solve various difficulties mentioned above using only one template. Experimental results demonstrate that the proposed method can well recognize different underwater objects, such as mine-like objects and triangle-like objects and can satisfy the demands of real-time application.

## 1. Introduction

Underwater object detection and recognition for sonar images is one of the most important challenges in ocean exploration [1–3]. Because of the complexity of the underwater acoustic environment, the multipath effect of underwater acoustic propagation, and the sonar scan angle, sonar images are prone to be affected by complicated background noise, and the target may be misrecognized causing false target confusion, object rotation, and scaling.

Threshold-based and model-based methods have commonly been used for underwater object detection in the past. Model-based approaches, such as Markov random field (MRF) [4, 5] and active contour [6–8] are computationally intensive, and the active contour model is greatly affected by the initial contour. Threshold-based methods, such as the Tsallis entropy-based method [9, 10], often suffer from low-quality sonar images. Meanwhile, confusion from fake

targets and complex background noise also lead to the low recognition rate of common threshold segmentation methods. Recently, saliency detection techniques have developed rapidly and have been applied in many fields, such as image segmentation, image localization, and object detection [11, 12]. In general, the localization accuracy of saliency detection methods under complex background can be further improved, but segmenting the region of the salient object in an image is very robust.

Template matching is another effective method for target detection and recognition, which finds objects in an image by calculating the similarity between the template and the subwindow of the object image. The template matching scheme based on feature correlation [13–18] has achieved good results in cases where the target and template have the same orientation and scale. As a result, the template matching method has a good reference to object recognition in sonar images. However, in the sonar image acquisition

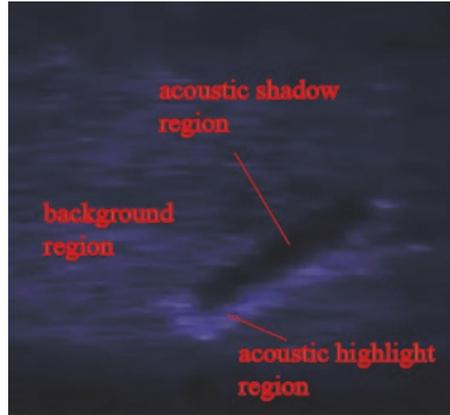


FIGURE 1: Sonar image of a mine-like object.

process, due to the relative motion between the sonar device and the target, the orientation of the object in different sonar images often varies, i.e., object rotation. This object rotation leads to a low recognition rate in traditional template matching methods, such as shape template matching (STM) and normalized cross-correlation (NCC). To find the rotated object, a common method is to create templates in multiple orientations and then match them one by one, which has very low efficiency and cannot achieve the demands of real-time application.

The abovementioned methods generally disregard the characteristics of sonar images, while, in our applications, we realize that some prior knowledge is very useful for better analyzing sonar images. As shown in Figure 1, a sonar image of mine-like objects generally consists of an acoustic shadow region, an acoustic highlight region, and a background region. It is notable that the shape of the acoustic shadow region is similar to that of the target. In particular, the shape of the acoustic shadow is relatively stable and more salient than the background region in the sonar image [19]. This prior knowledge is well considered in our approach and plays an important role in object detection and recognition.

In this paper, inspired by the advantages of saliency detection and template matching methods, we propose an underwater object detection and recognition method using transformable template matching based on the prior knowledge of sonar imaging. Experimental results verify that the proposed method can well locate and recognize different underwater objects, such as mine-like objects and triangle-like complex objects. The main contributions of this paper are as follows: (1) after studying the characteristics of targets in sonar images, design a template for underwater target recognition and localization based on prior knowledge; (2) we adopt the saliency detection technique based on FFT to segment the target region and narrow the scope of template matching; (3) by integrating saliency detection, affine transformation, and template matching, we can efficiently locate and recognize targets, which can satisfy real-time application demand.

## 2. Proposed Method

As shown in Figure 2, we first extract the features and construct a template from sonar video sequences based on the analysis of the acoustic shadow and highlight region. Then, we achieve coarse localization of the target in the objective image via fast saliency detection based on the spectral residual, which can significantly improve the running efficiency by avoiding exhaustive global searching for the target. We expand the salient region to the same size as the template and identify it as the target region. After affine transformation [20] of the template according to the orientation of the target region, we extract normalized gradient features and calculate the similarity between the template and the target region, which can be robust against various effects, such as fake targets, complicated backgrounds, rotation, and scaling of the object. Finally, according to the similarity score, we can identify the target.

*2.1. Construction of Template Based on Prior Knowledge.* As mentioned above, the shape of the acoustic shadow region is similar to that of the object, relatively stable, and more salient than the background region in the sonar image. Therefore, the structural features of the object can be used for similarity measures, such as the gradient orientation feature of the acoustic shadow and highlight regions. The template should contain abundant and sufficiently clear structural information to ensure the matching accuracy. In general, the template should be constructed by a deep learning or regression model. However, due to the high cost of collecting sonar images, positive training samples are very rare. Therefore, we use image difference and morphological operator to extract the structural information of objects to construct the template.

First, we extract the candidate object by the image difference and threshold segmentation. The corresponding equations are as follows:

$$D(\mathbf{x}, \mathbf{y}) = \mathbf{F}(\mathbf{x}, \mathbf{y}) - \mathbf{I}(\mathbf{x}, \mathbf{y}), \quad (1)$$

$$D_1 = \{(x, y) \in D \mid t_{min} \leq D(\mathbf{x}, \mathbf{y}) \leq t_{max}\}, \quad (2)$$

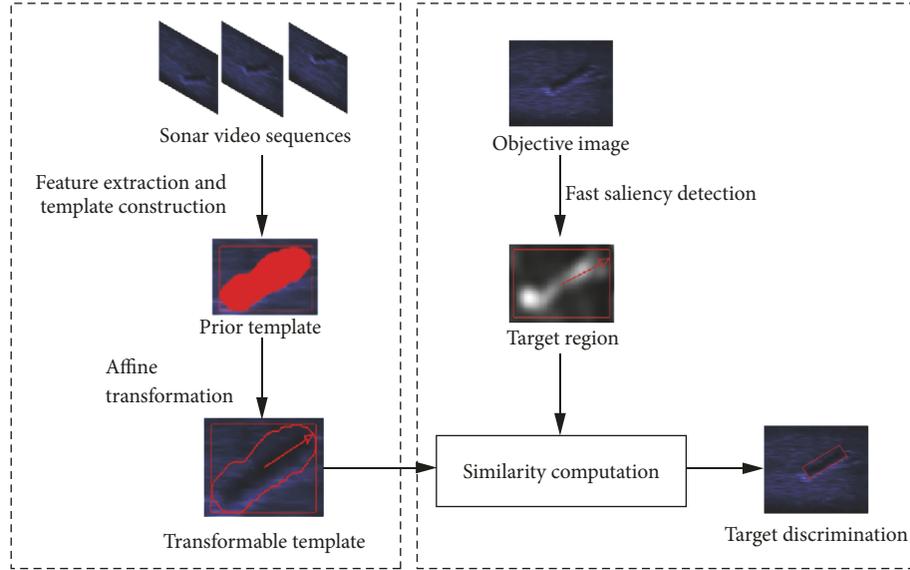


FIGURE 2: Overview of the proposed method.

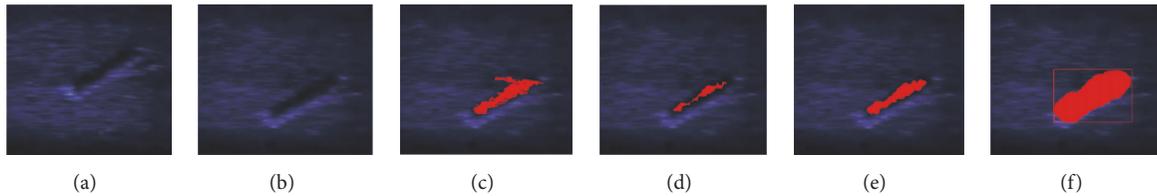


FIGURE 3: Construction process of the template based on prior knowledge: (a) is the previous frame, (b) is the next frame, (c) is the result of image difference and threshold segmentation, (d) is the erosion result of the candidate object, (e) is the dilation result of the candidate object, and (f) is the template containing the acoustic shadow and highlight region.

where  $I$  is the previous frame image,  $F$  is the next frame image,  $D$  is the result of the image difference,  $D_1$  is the result of performing the threshold operation for  $D$ , and  $t_{min}$  and  $t_{max}$  are the minimum and maximum thresholds, respectively. The candidate object is shown in Figure 3(c).

Then, to smooth the boundary of the candidate object, we conduct erosion and dilation operations for the candidate object by

$$D_2 = (D_1 \ominus B) \oplus B, \quad (3)$$

where  $D_2$  is the result of performing erosion and dilation operations for  $D_1$  using structure element  $B$ . The results are shown in Figures 3(d) and 3(e).

Finally, to contain both the acoustic shadow and highlight regions in the template, we conduct a dilation operation by

$$T = D_2 \oplus C, \quad (4)$$

where  $T$  is the result of performing a dilation operation for  $D_2$  using structure element  $C$ . In addition, we construct the template through computing the minimum enclosing rectangle of the dilated object and the background region. The template is shown in Figure 3(f).

## 2.2. Fast Saliency Detection and Similarity Measure

*Step 1* (search scope identification via fast saliency detection). After the template is created, another key factor in the proposed method is how to determine the search scope quickly. Inspired by literature [21], we achieve the coarse localization of the object in the objective image via fast saliency detection based on the spectral residual, which can significantly improve the time efficiency. After fast Fourier transform of the image and computing the logarithm of the amplitude, the log spectrum is obtained. The spectral residual can be obtained by subtracting the average curve from the amplitude frequency response curve of the image. A saliency map can be obtained by using inverse Fourier transform to transform spectral residuals into the spatial domain, which can obtain the coarse position of the object. The corresponding equations are as follows:

$$\begin{aligned} R(f) &= L(f) - h_n(f) * L(f) \\ &= \log A(f) - h_n(f) * \log A(f), \end{aligned} \quad (5)$$

where  $A(f) = R(F[I(x)])$  is the amplitude spectrum of the Fourier frequency domain,  $\log A(f)$  is the log spectrum of

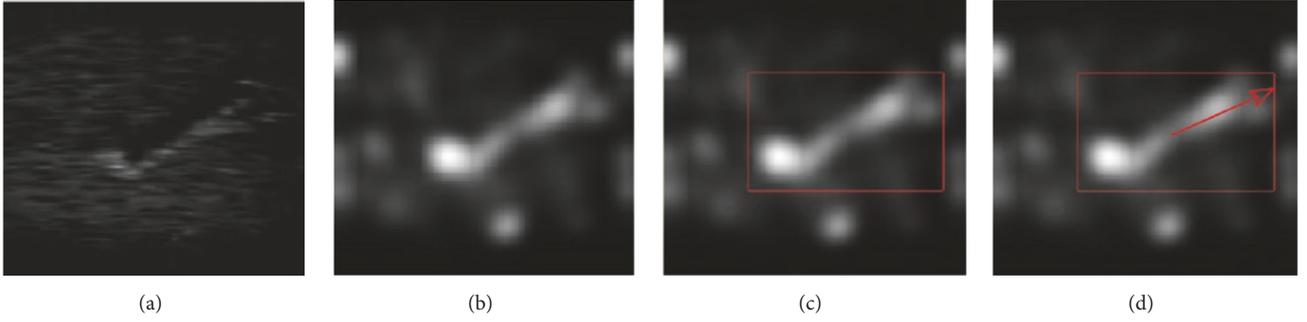


FIGURE 4: Search scope identification via fast saliency detection: (a) is the gray image, (b) is the saliency map, (c) is the target region after performing the threshold and morphology operator, and (d) is the orientation of the target region.

the amplitude,  $h_n(f)$  is a local mean filter, to simulate the average amplitude frequency response via convolution with the log  $A(f)$ , and  $R(f)$  denotes the statistical singularities, called spectral residuals.

$$P(f) = S(F[I(x)]), \quad (6)$$

where  $P(f)$  is the phase spectrum of the image  $I(x)$  and  $F$  is the Fourier Transform.

$$S(x) = g(x) * F^{-1}[\exp(R(f) + P(f))]^2, \quad (7)$$

where  $S(x)$  is the saliency map,  $F^{-1}$  is the inverse Fourier transform, and  $g(x)$  is a Gaussian filter to smooth the saliency map for better visual effects.

For the saliency map, the coarse position of the target can be segmented by a threshold. Considering the scale of the object, we perform the morphology operator to extend the target region to the size of the prior template. The target region contains some backgrounds, which can help to improve the accuracy of template matching. Through fast locating the target region, it can avoid searching the whole image for the target. Furthermore, we compute the orientation and central coordinates of the target and perform the affine transformation of the template for solving the rotation problem. Figure 4 illustrates the results of the search scope identification via fast saliency detection.

*Step 2* (template matching via the similarity measure). As shown in Figure 1, the acoustic shadow usually follows the highlight region due to the wide sonar scan angle, and the boundary of acoustic highlight and shadow region has salient edge features. Therefore, we use the edge gradient features [22] to calculate the similarity between the template and the target region. We calculate the gradient vector of the points in the prior template and the target region by (8), and represent them by  $P_i$  and  $Q_i$  (see (9)), respectively.

$$\begin{aligned} G_x(x, y) &= [-1, 0, 1] * I(x, y), \\ G_y(x, y) &= I(x, y) * [-1, 0, 1]^T, \end{aligned} \quad (8)$$

where  $I(x, y)$  is the input image.

$$\begin{aligned} P_i &= (G_{xi}, G_{yi}), \\ Q_i &= (G_{vi}, G_{wi})^T \end{aligned} \quad (9)$$

As mentioned above, the target region extracted by fast saliency detection is the same size as the prior template, and affine transformation of the prior template is also the same orientation as the target region. Therefore, we can calculate the similarity by the normalized dot product of the gradient vectors of the prior template and the target region. The corresponding equation is as follows:

$$Sc = \frac{1}{n} \sum_{i=1}^n \frac{|G_{xi} \cdot G_{vi} + G_{yi} \cdot G_{wi}|}{\sqrt{G_{xi}^2 + G_{yi}^2} \cdot \sqrt{G_{vi}^2 + G_{wi}^2}}, \quad (10)$$

where  $Sc$  is the score of the similarity measure.

To further improve the efficiency of the algorithm, the minimum score threshold  $Sc_j$  is set, which represents the partial sum of the dot product up to the  $j$ th element of the vector  $P$ . When the score of the similarity measure achieves the user-defined threshold, the evolution of the sum can be discontinued, which can further accelerate the recognition process.

$$Sc_j = \frac{1}{j} \sum_{i=1}^j \frac{|G_{xi} \cdot G_{vi} + G_{yi} \cdot G_{wi}|}{\sqrt{G_{xi}^2 + G_{yi}^2} \cdot \sqrt{G_{vi}^2 + G_{wi}^2}} \quad (11)$$

*Step 3* (object discrimination according to the similarity score). Comparing the similarity score with the user-defined minimum score threshold, we can discriminate whether there is an object. Then, we extract the center coordinates of the identified object and draw the minimum outer rectangle according to the orientation of the object.

### 3. Experimental Results and Discussion

*3.1. Experimental Results Comparing the Localization and Recognition Accuracy.* To evaluate the localization and recognition accuracy of the proposed method, we captured mine-like object images in a real underwater environment by



FIGURE 5: Comparison results of different object recognition methods.

an automatic underwater vehicle (AUV) embedded dual-frequency identification sonar (DIDSON). The original images can be divided into three categories: the first type contains real objects to be recognized, the second type contains false targets but no real object, and the third type contains only the background but no object. The corresponding acoustic shadow region and highlight region of the object in different original images are different. The acoustic shadow area shape of the false target is similar to that of the real object, but the surrounding highlight area is not salient, and there is no obvious edge. In addition, there are different colors and textures in the background of the sonar images. The proposed method should recognize the real object rather than the false

target and not be disturbed by the background in the sonar images.

We compare our approach with some state-of-the-art methods, such as STM, NCC, SURF, and random forest (RF). Representative results are shown in Figure 5. The first 7 rows are the target images, the 8th row is the false target image, and the 9th row is the background image. As shown, the STM method can identify most of the real objects located in a bright background and achieve high localization accuracy, but it cannot recognize the target that is partly submerged in the dark background. The NCC method can correctly identify the real object in sonar images and can achieve high localization accuracy. However, similarly, it may miss

TABLE 1: Recognition rate and false alarm rate of the proposed method.

	A	B	C
Original images	23	6	23
Correctly recognized images	$A_1=23$	$B_1=5$	$C_1=23$
Incorrectly recognized images	$A_2=0$	$B_2=1$	$C_2=0$
Recognition rate (%)	98.1		
False alarm rate (%)	1.9		

TABLE 2: Recognition rate and false alarm rate using different methods.

	STM	NCC	SURF	RF	Proposed
Recognition rate (%)	84.6	90.4	88.5	92.3	98.1
False alarm rate (%)	15.4	9.6	11.5	7.7	1.9

the target that is partly submerged in the dark background, and, also, it is prone to be disturbed by false targets with similar color features. As SURF is a feature point matching method, we regard the targets detected by SURF when the number of matching feature points is more than 8 and the number of points in the target region is more than 70%. The advantage of SURF is that it does not misidentify false targets and background. However, it cannot precisely locate the target, because the matching points are generally distributed in the highlight region and the background, but very few are distributed in the acoustic shadow region. As RF is a training method, we randomly select 54 target images and 60 background images as the training set and use the remaining 23 target images and 29 background images as the testing set. The number of classification trees is set to 100. The RF method reaches the second highest classification accuracy. However, it misidentifies the false target due to adopting the color histogram as a feature and cannot locate the target in sonar images. Finally, the proposed method can not only correctly identify real objects, false targets, and backgrounds in sonar images but also precisely locate the target. The proposed method achieves the highest performance and visual effect.

To quantify the recognition effect, we define the recognition rate  $R$  and false alarm rate  $F$  as follows:

$$\begin{aligned} R &= \frac{A_1 + B_1 + C_1}{A + B + C}, \\ F &= \frac{A_2 + B_2 + C_2}{A + B + C}, \end{aligned} \quad (12)$$

where  $A$  is the number of images with a real object,  $B$  is the number of images with a false target, and  $C$  is the number of images with only the background. We selected 52 original images as the experimental sample data. The recognition rate and false alarm rate of the proposed method are shown in Table 1.

The experimental data of different methods are shown in Table 2.

**3.2. Experimental Results for Object Rotation Recognition.** To validate the robustness of the proposed method in the case

of object rotation, we conduct the recognition experiment for object rotation. As shown in Figure 6, the orientation of the object in the original image is different. If only a fixed template is used, the traditional template matching method cannot recognize the object. It is awkward to build a template library according to multiple orientations. In the proposed method, only one prior template is needed, an affine transformation of the prior template is performed according to the orientation difference between the prior template and the object identified by fast saliency detection, and the object in the objective image is matched by the rotated template. The experimental results in the last column in Figure 6 illustrate that the proposed method is robust for object rotation. The recognition rate reaches 94.2%.

For object scaling, the ratio between the area of the object identified by fast saliency detection and the area of the object in the prior template is calculated and used as a scaling factor. After an affine transformation of the prior template is performed according to the scaling factor, the object in the search image is matched by the scaling template.

**3.3. Experimental Results for Object Shape Recognition.** To further verify the extensibility of the proposed method, we conduct another complicated object recognition experiment. As shown in Figure 7, the shape of the object is similar to a triangle, and several float balls are bound to the object. Because the triangle-like object is composed of multiple floating balls and a triangular frame, the structure, shape, and edge features of the triangle-like object are much more complex than that of the mine-like object. Furthermore, due to the float balls floating in the water, acoustic shadows of the float balls are varied in position and submerged in the background. All of these make it more difficult to recognize. However, the highlight regions of the triangular frame and the float balls are still salient to the background region in the sonar image.

For the triangle-like object recognition using the proposed method, first, we extract the structural information of the objects and construct the prior template using the image difference and the morphological operator. Figure 8

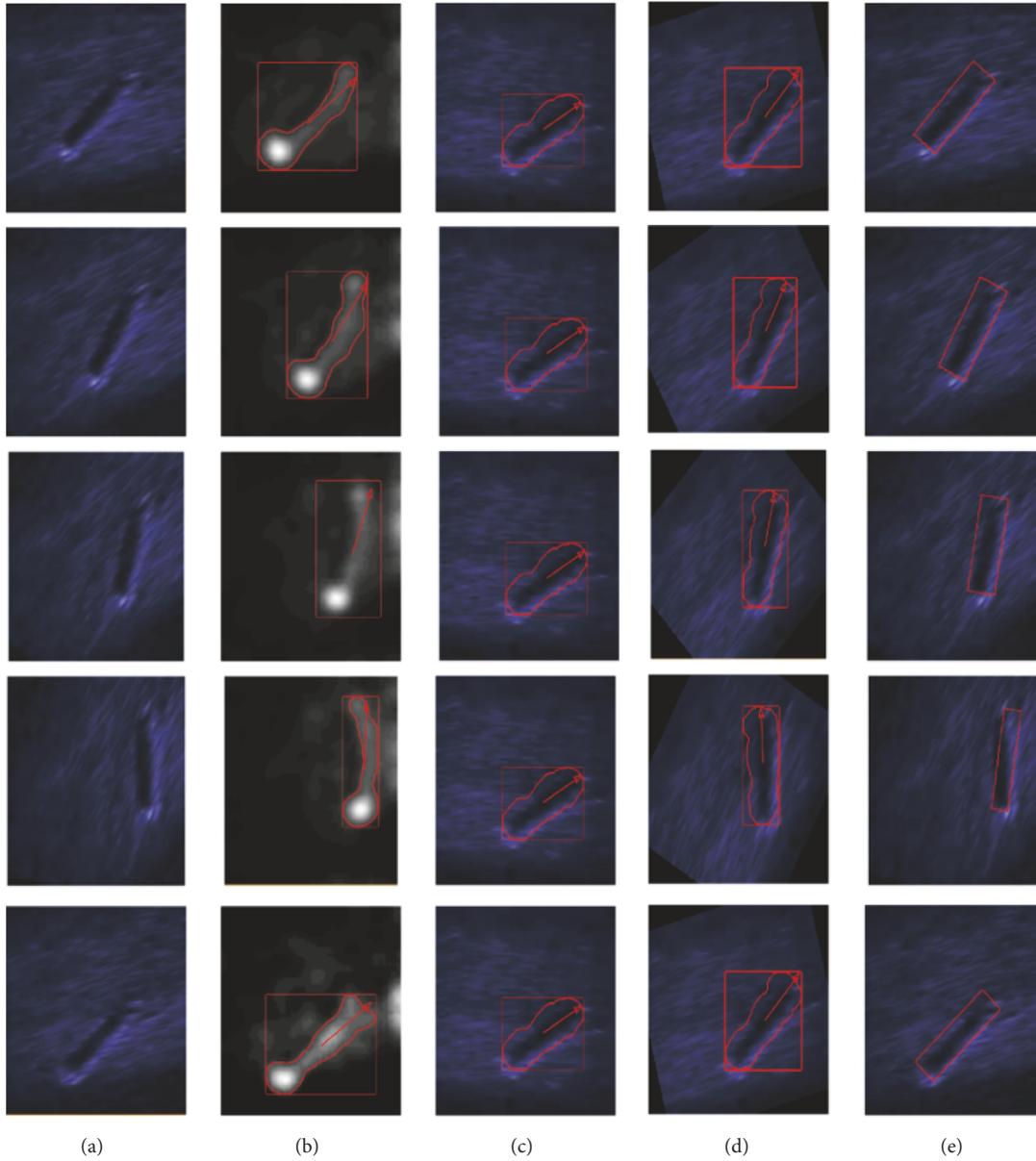


FIGURE 6: Recognition experiment for object rotation: (a) shows the original images containing object with different orientations, (b) shows the result of fast saliency detection and the orientation of the salient object is extracted, (c) shows the prior template, (d) shows the rotated template according to the orientation of the salient object, and (e) shows the result of the object recognition according to the rotated template.

illustrates the construction process of the prior template for the triangle-like object recognition.

Then, we select the sample images for the experiment. Since the previous experiment has already verified that the proposed method is not interfered with by a false target and complex background, the experiment on those images will not be repeated. The original images as the experimental samples all contain the object to be recognized. Each object to be recognized has a different position in the sonar image, some of which are close to the image boundary.

Finally, we use the proposed method to test the sample images and obtain the experimental results. We use

a rectangle to mark the identified object area. The more complete the object area covered by the rectangle, the higher the recognition accuracy. As shown in Figure 9, for all the sample images, the proposed method correctly identifies the object. Through carefully observing the identified object's minimum outer rectangle, we find that most of the minimum outer rectangle accurately covers the object; only the object in the last row is not covered completely by the minimum outer rectangle. The experimental results validate the ability of the proposed method to recognize different objects.

Furthermore, we demonstrate the comparison results of triangle-object recognition via different methods in Figure 10.

TABLE 3: Recognition rate of triangle-like objects by different methods.

	STM	NCC	SURF	RF	Proposed
Recognition rate (%)	7.7	76.9	84.6	-	92.3
False alarm rate (%)	92.3	23.1	15.4	-	7.7

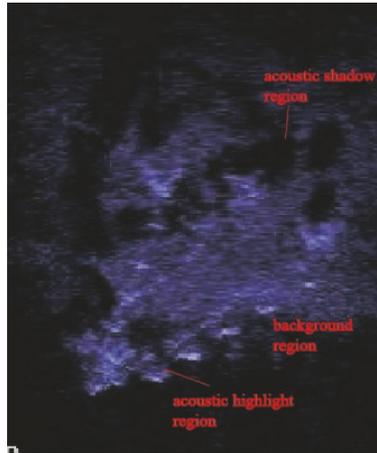


FIGURE 7: The sonar image containing a triangle-like object.

The experimental results of STM are not ideal. STM basically fails in this experiment due to the complex texture of the triangle-object. NCC and SURF are effective for most of the cases, but the performances are not as good as ours. Please note that the matching points of SURF on the triangle-like object are more than those of the mine-like object, which shows that SURF is more suitable for large target matching in an image. In this experiment, we do not compare with RF because the dataset is too small to adequately train RF. The quantitative results are shown in Table 3.

#### 4. Conclusions

In this paper, aiming at the key problems for underwater object recognition, such as object rotation, false target interference and complex backgrounds, and recognition ability of different objects, we propose an underwater object recognition method using transformable template matching based on prior knowledge. We locate the target region in the objective image via fast saliency detection techniques based on FFT, which can avoid searching the whole image for the object. We extract the features of the target region, such as central coordinates, orientation, and area; then, the corresponding affine transformation of the template is performed according to these features, which can well adapt to the rotation of objects in different sonar images using only one template. The normalized gradient

orientation feature is used for calculating the similarity between the target region and the template, which can adapt to various interferences, such as false targets and complex backgrounds. Experimental results demonstrate that the proposed method can well recognize and locate different underwater objects, such as mine-like objects and triangle-like objects, and can achieve the demands of real-time application.

#### Data Availability

The image type data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

#### Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants no. 61773367 and no. 61303168 and the Youth Innovation Promotion Association CAS no. 2016183.

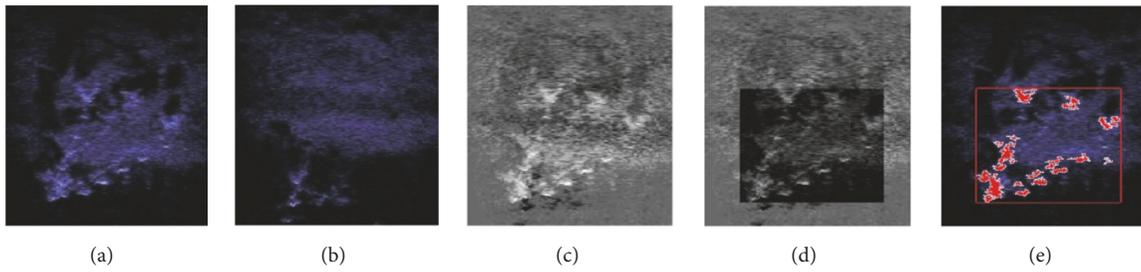


FIGURE 8: Template construction process of the triangle-like object: (a) is the image containing the triangle-like object, (b) is the image containing mainly background, (c) is the result of the image difference, and (d) and (e) are the template containing the acoustic highlight region and background.

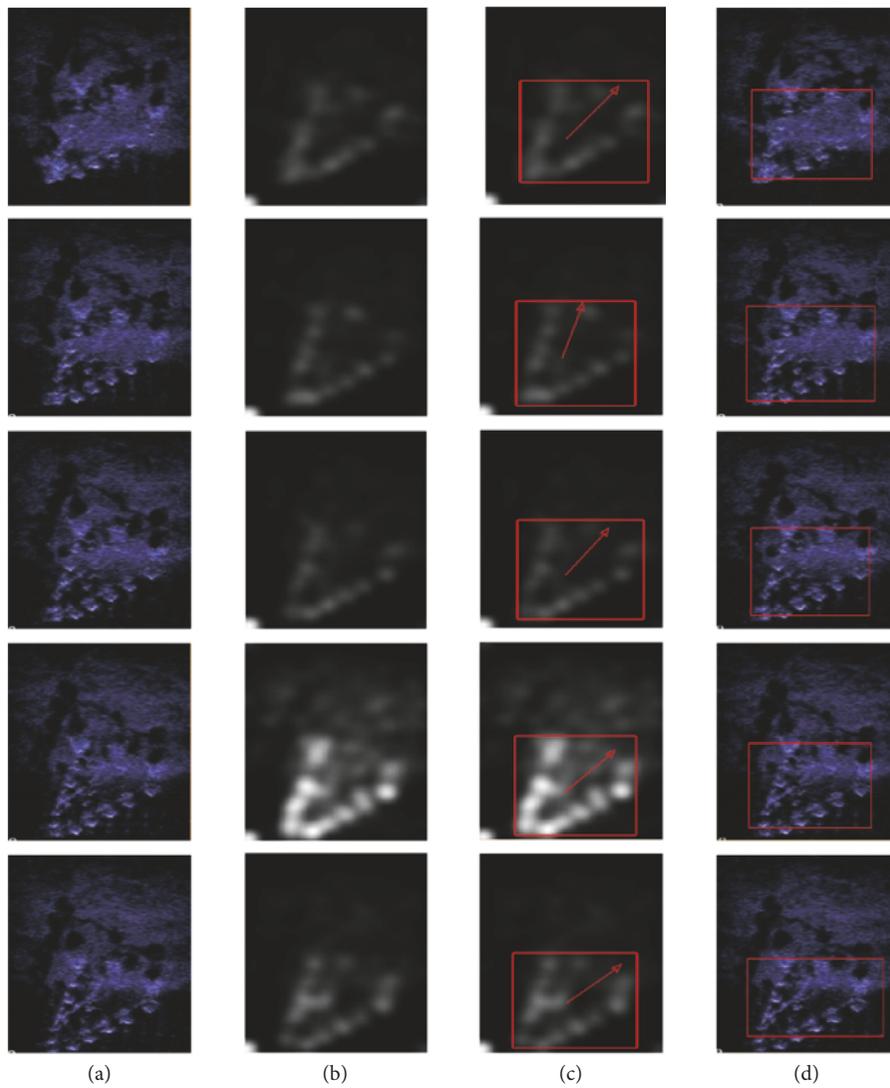


FIGURE 9: Results of triangle-like object recognition via the proposed method: (a) is the original image containing the triangle-like object, (b) shows the results after performing fast saliency detection, (c) shows the results after identifying object region and orientation, and (d) shows the results of the triangle-like object recognition.

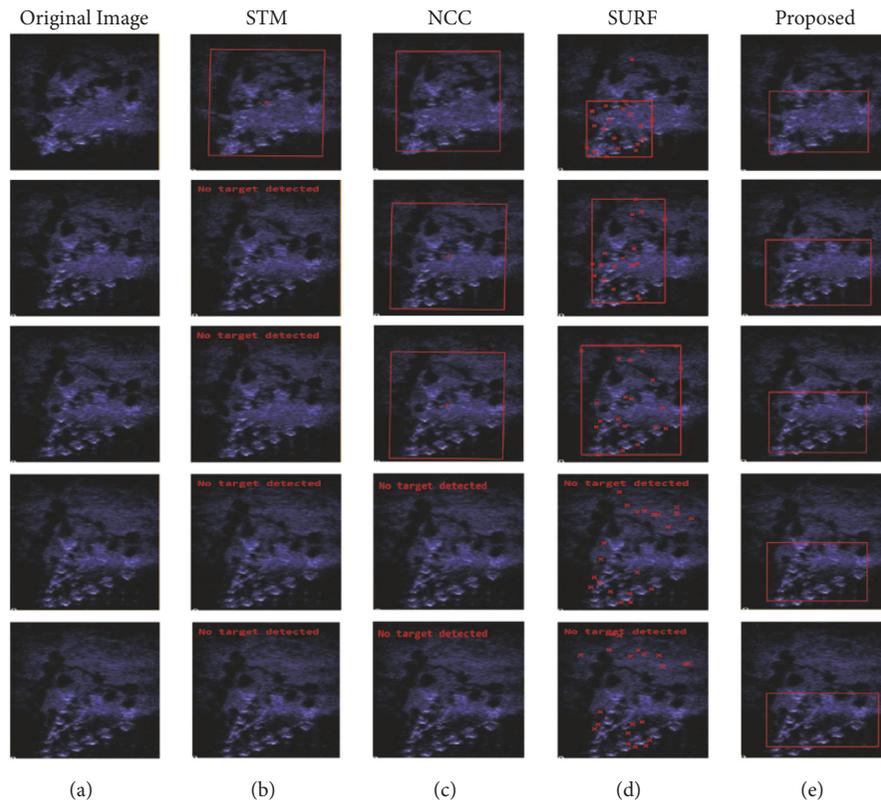


FIGURE 10: Comparison results of triangle-like object recognition via different methods.

## References

- [1] L. Zheng and K. Tian, "Detection of small objects in sidescan sonar images based on POHMT and Tsallis entropy," *Signal Processing*, vol. 142, pp. 168–177, 2018.
- [2] T. Celik and T. Tjahjadi, "A novel method for sidescan sonar image segmentation," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 2, pp. 186–194, 2011.
- [3] G. G. Acosta and S. A. Villar, "Accumulated CA-CFAR Process in 2-D for Online Object Detection from Sidescan Sonar Data," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 558–569, 2015.
- [4] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Markov random field and fuzzy logic modeling in sonar imagery: application to the classification of underwater floor," *Computer Vision and Image Understanding*, vol. 79, no. 1, pp. 4–24, 2000.
- [5] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical MRF model," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216–1231, 2000.
- [6] X. Wang, L. Guo, J. Yin, Z. Liu, and X. Han, "Narrowband Chan-Vese model of sonar image segmentation: A adaptive ladder initialization approach," *Applied Acoustics*, vol. 113, pp. 238–254, 2016.
- [7] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
- [8] B. Lehmann, S. K. Ramanandan, K. Siantidis, and D. Kraus, "Extended active contours approach for mine detection in synthetic aperture sonar images," in *Proceedings of the International Symposium on Ocean Electronics*, vol. 40, no. 5, pp. 73–78, 2011.
- [9] Q. Lin and C. Ou, "Tsallis entropy and the long-range correlation in image thresholding," *Signal Processing*, vol. 92, no. 12, pp. 2931–2939, 2012.
- [10] P. M. Rajeshwari, D. Rajapan, G. Kavitha, and C. M. Sujatha, "Multilevel Tsallis entropy based segmentation for detection of object and shadow in SONAR images," in *Proceedings of the 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems, SPICES 2015*, vol. 71, no. 2, pp. 1–5, India, February 2015.
- [11] T. Xi, W. Zhao, H. Wang, and W. Lin, "Salient object detection with spatiotemporal background priors for video," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3425–3436, 2017.
- [12] F. Guo, W. Wang, J. Shen et al., "Video Saliency Detection Using Object Proposals," *IEEE Transactions on Cybernetics*, vol. 48, no. 11, pp. 3159–3170, 2018.
- [13] J.-H. Chen, C.-S. Chen, and Y.-S. Chen, "Fast algorithm for robust template matching with m-estimators," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 230–243, 2003.
- [14] A. Sibiriyakov, "Fast and high-performance template matching method," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011*, pp. 1417–1424, USA, June 2011.
- [15] T. Dekel, S. Oron, M. Rubinstein, S. Avidan, and W. T. Freeman, "Best-Buddies Similarity for robust template matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*, pp. 2021–2029, USA, June 2015.

- [16] E. Elboher and M. Werman, "Asymmetric correlation: a noise robust similarity measure for template matching," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3062–3073, 2013.
- [17] Y. Hel-Or, H. Hel-Or, and E. David, "Matching by tone mapping: Photometric invariant template matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 317–330, 2014.
- [18] I. Talmi, R. Mechrez, and L. Zelnik-Manor, "Template matching with deformable diversity similarity," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 1311–1319, USA, July 2017.
- [19] N. Kumar, U. Mitra, and S. S. Narayanan, "Robust Object Classification in Underwater Sidescan Sonar Images by Using Reliability-Aware Fusion of Shadow Features," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 592–606, 2015.
- [20] S. Korman, D. Reichman, G. Tsur, and S. Avidan, "Fast-match: fast affine template matching," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2331–2338, June 2013.
- [21] X. D. Hou and L. Q. Zhang, "Saliency detection: a spectral residual approach," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, June 2007.
- [22] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1063–6919, June 2005.

