Orientation Estimation of Rotated Sonar Image Targets via the Wavelet Subimage Energy Ratio

Peng Zhang, Jinsong Tang, Heping Zhong, Haoran Wu, Han Li, and Yue Fan

***Abstract*—Precise orientation information facilitates target recognition of sonar images. Traditional orientation estimation methods, including the Hough transform method, have poor anti-noise abilities, so they cannot achieve high estimation precision for sonar images, which usually have low signal-to-noise ratios. The convolutional neural network (CNN) regression method is sensitive to image affine transformations and requires extensive computation, so it cannot achieve strong robustness or high speed. To achieve high-precision, high-speed, and robust orientation estimation of sonar image targets, we present a novel orientation estimation method via the wavelet subimage energy ratio (WSER). The WSER varies with the rotation angles of images and has the highest value when the long axes of targets are vertical. It is translational and scale invariant and does not need supervised training. Therefore, we propose estimating orientations of sonar image targets by finding the max of the WSER to achieve high precision, high speed, and strong robustness to the translation and scale transformations of images. The results of experiments on a self-made sonar image dataset show that the mean absolute error (MAE) of the proposed method is 7.9°, while the MAEs of the CNN regression method and traditional methods are 10.1 and 22.9°, respectively. In addition, the proposed method is six times faster than the CNN regression method. The proposed orientation estimation method has also been applied to align the orientation of sonar images and CNN can achieve state-of-the-art performance for rotated target recognition using the aligned images. Data and codes are publicly available.**

***Index Terms*—Automatic target recognition, orientation estimation, rotated targets, sonar image, wavelet subimage energy.**

I. Introduction

SONAR images are a powerful tool for underwater target recognition [1], [2]. Rotated target recognition of sonar images has attracted increasing interest [1], [3], [4], [5], [6], [7], [8], [9], because existing target recognition algorithms of images have been found to have difficulty to recognizing arbitrarily rotated targets of sonar images. However, current studies on rotated target recognition are limited to target detection [7], [8], [9] and classification [2], [4], [5], [6], and far too little attention has been paid to orientation estimation. In fact, the precise orientation information of rotated targets can not only be used in pose estimation tasks but also in improving rotated target recognition [10].

The orientation estimation of rotated targets on sonar images is similar to the orientation estimation of general optical images, which is also referred to as tilt detection, or tilt estimation. Existing orientation estimation methods of general optical images can be divided into two categories, traditional methods that use manually designed features, and deep learning methods that use automatically learned features. In traditional methods, the Hough transform method [11] was the earliest approach for estimating orientation, it uses the edge features of images to detect lines, and uses the slopes of lines to estimate the orientation of targets. The Radon transform method [12], sometimes called the projection method, projects a gray image from different angles, and takes the angle with the maximum projection value as the orientation of targets. The Fourier transform method [13] uses frequency domain features, and takes the angle of the maximum spectral density as the orientation of targets.

The advantage of traditional methods is that the calculation speed is fast, and the image features do not need to be learned. However, one of the disadvantages is that the features extracted by traditional methods are sensitive to noise. Hence, they cannot achieve high orientation estimation precision for sonar images, which usually have lower signal-to-noise ratios (SNRs) than general optical images.

To achieve high orientation estimation precision, researchers have developed deep learning methods. The convolutional neural network (CNN) regression method [10], [14], [15], [16] is the most effective deep learning orientation estimation method. It models the orientation estimation task as an angle regression problem. Therefore, the orientation labels of rotated target images are learned in the training stage, and can be directly predicted in the test stage.

The advantage of CNN regression methods is that they can achieve high precision of orientation estimation, as long as there is a sufficient quantity of training data. However, their disadvantages are as follows: 1) Deep convolutional features need to be learned through end-to-end supervised learning. Supervised learning requires many sonar images to have orientation labels, while labeling orientations of images is time-consuming and easily introduces to introduce bias. In addition, owing to the lack of sonar images of underwater targets, there are usually not enough training data with labels to train a well-generalized

Manuscript received 3 July 2022; revised 20 September 2022; accepted 9 October 2022. Date of publication 17 October 2022; date of current version 26 October 2022. This work was supported by the National Natural Science Foundation of China under Grant 42176187. *(Corresponding author: Jinsong Tang.)*

Peng Zhang, Jinsong Tang, Heping Zhong, Haoran Wu, and Han Li are with the Institute of Electronic Engineering, Naval University of Engineering, Wuhan 430000, China (e-mail: pengzhang.ai@outlook.com; jinsongtangwh@163.com; zheping525@sohu.com; wuhaoran_wh@163.com; lh371082@163.com).

Yue Fan is with the Central China Normal University, Wuhan 430000, China (e-mail: 75112248@qq.com).

This article has supplementary downloadable material available at https://doi.org/10.1109/JSTARS.2022.3215068, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2022.3215068

CNN [1]. 2) Because the targets detected from sonar images usually have different sizes, and center locations, the orientation estimation task for real sonar images faces the challenge of scale and translation transformation. However, CNN has been proved to have poor robustness to scale and translation [17]. 3) The inference speed of CNN is slower than that of traditional methods. CNN has a large amount of calculation when calculating deep convolutional features. In addition, because the size of a sonar image is often larger than that of a general optical image, CNN models for sonar images requires more cost greater computation than CNN models for optical images [18].

From the abovementioned review, it can be concluded that the existing orientation methods cannot simultaneously achieve high precision, high speed, and strong robustness in the orientation estimation task of sonar images. Therefore, how to construct an orientation estimation method with high precision, speed, and robustness for sonar images remains unclear. To address this question, this article presents a novel orientation estimation method of rotated sonar image targets via the wavelet subimage energy ratio (WSER).

The WSER is defined as the ratio of the energy of wavelet subimages to the energy of the input images. We find that WSER varies with the rotation angle of input images and achieves the highest value when the long axis of targets is normal oriented. Thus, we propose using the max of WSER to estimate the orientation of the arbitrarily rotated target inside an image. Moreover, we prove that the WSER is invariant to both the translation and scale transformation, and has a strong antinoise ability. Therefore, the proposed method can precisely predict the orientation of rotated sonar image targets, and is robust to the translation and scale transformations of the targets. Furthermore, it is slightly slower than the traditional methods significantly faster than the CNN regression method.

Moreover, based on the proposed orientation estimation method, we investigate applying orientation information to facilitate rotated target classification. We align the arbitrarily rotated orientation to the same orientation based on the targets' orientations estimated by the proposed method, and achieve state-of-the-art performance in rotated target classification using the aligned images.

The contributions of this article include as follows.

1) This article presents a novel orientation estimation method via the WSER. The WSER is found to be rotation symmetrical, which means that it varies with the rotation angles of input images and achieves the max value when the long axis is vertical. Moreover, the WSER is proven to be invariant to both translation and scale transformation. The proposed method can precisely estimate orientations of arbitrarily rotated targets on sonar images.
2) This article advances the understanding of wavelet decomposition. This article uses wavelet subimage energy as an orientation descriptor for the first time, while the existing orientation estimation methods only use wavelet decomposition as an edge detector.
3) This article demonstrates that precise orientation information can be applied to improve the performance of rotated target recognition. The proposed orientation estimation method via the WSER is successfully applied to align the orientation of rotated sonar images. Using the aligned images, we can achieve state-of-the-art classification performance for rotated target recognition.

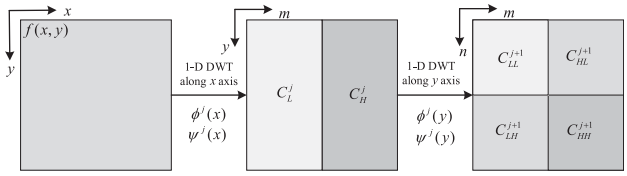

Fig. 1. Calculation process of 2-D DWT of an image.

II. Method

A. Image Wavelet Decomposition

Image wavelet decomposition refers to 2-D orthogonal discrete wavelet transform (DWT) [19], which entails performing 1-D transform along the rows and columns successively. The calculation process of 2-D DWT of an image $f(x, y)$ is shown in Fig. 1, where the size of $f(x, y)$ is $N \times N$. The calculation process is divided into two steps.

First, apply 1-D DWT to $f(x, y)$ along the x-axis. Given the scale function $\phi^j(x)$ and the wavelet function $\psi^j(x)$ at resolution$2^j (j \in Z)$, the 1-D DWT result C_L^{j+1}, C_H^{j+1} can be written as $C_L^{j+1} = \langle f(x, y), \phi^j(x)\rangle$and $C_H^{j+1} = \langle f(x, y), \psi^j(x)\rangle$, respectively.$C_L^{j+1}$, C_H^{j+1} are called subimages. They have a size of $N/2 \times N$. C_L^{j+1} contains the low frequencies of the rows of $f(x, y)$. C_H^{j+1} contains the high frequencies of the rows of $f(x, y)$.

Second, apply 1-D DWT to C_L^{j+1} and C_H^{j+1} along the y-axis. Given the scale function $\phi^j(y)$ and the wavelet function $\psi^j(y)$ at resolution $2^j (j \in Z)$, the 2^*D DWT result of $f(x, y)$ can be finally written as

$$C_{LL}^{j+1} = \left\langle f(x, y), \phi^j(x) \cdot \phi^j(y)\right\rangle \tag{1}$$

$$C_{HL}^{j+1} = \left\langle f(x, y), \psi^j(x) \cdot \phi^j(y)\right\rangle \tag{2}$$

$$C_{LH}^{j+1} = \left\langle f(x, y), \phi^j(x) \cdot \psi^j(y)\right\rangle \tag{3}$$

$$C_{HH}^{j+1} = \left\langle f(x, y), \psi^j(x) \cdot \psi^j(y)\right\rangle \tag{4}$$

C_{LL}^{j+1}, C_{LH}^{j+1}, C_{HL}^{j+1}, and C_{HH}^{j+1} are subimages with a size of $N/2 \times N/2$. C_{LL}^{j+1} contains the low frequencies of both the rows and columns of $f(x, y)$ and depicts the $f(x, y)$ at a lower resolution. C_{LH}^{j+1} contains the low frequencies of the rows of $f(x, y)$ and the high frequencies of the columns of $f(x, y)$ and depicts the vertical components of $f(x, y)$. C_{HL}^{j+1} depicts the horizontal components of $f(x, y)$ and C_{HH}^{j+1} depicts the diagonal components of $f(x, y)$.

Fig. 2 shows the single-level 2-D DWT of four images, which are composed of white geometric figures (rectangle, square, ellipse, and circle) against black backgrounds. The Haar orthonormal wavelet basis [21] is chosen to compute the DWT. Owing to its single-level DWT, j is taken as 0, and four subimages are denoted as C_{LL}^1, C_{LH}^1, C_{HL}^1, and C_{HH}^1. As it is shown in Fig. 2,

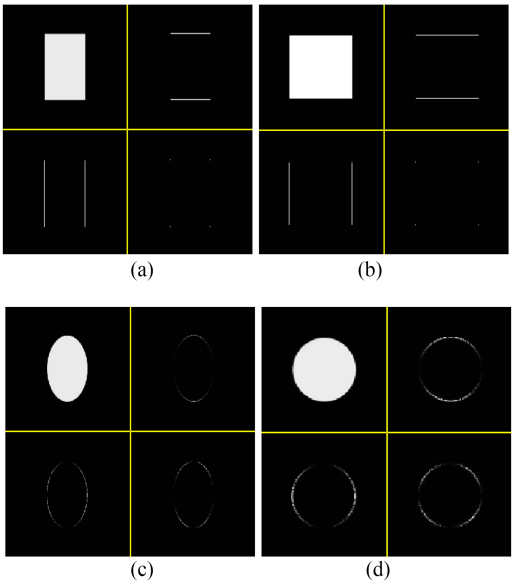

Fig. 2. DWT of images of geometric figures. (a) Rectangle. (b) Square. (c) Ellipse. (d) Circle.

C^1_{LH} and C^1_{HL} extract the vertical and horizontal components of the geometric figures. Applying DWT to C^1_{LL}, two-level DWT will be obtained. Analogously, multilevel DWT will be obtained. In this article, only the single-level DWT is used in the estimation process of image orientation.

B. Wavelet Subimage Energy Ratio

1) Definition: Let the energy of the input image $f(x, y)$ and its single-level DWT $C^{j+1}_{LL}, C^{j+1}_{LH}, C^{j+1}_{HL}, C^{j+1}_{HH}$ be E_{raw} and E^1_{LL}, E^1_{LH}, E^1_{HL}, E^1_{HH}. According to the energy conservation law of DWT [19], [20], E_{raw} can be written as

$$E_{\text{raw}} = E^1_{LL} + E^1_{LL} + E^1_{LH} + E^1_{HL} + E^1_{HH}. \tag{5}$$

Denote the value of each subimage at (m, n), $0 \le m, n \le N/2$, as $C^1(m, n)$. Then, according to the standard orthogonality of the wavelet basis, the energy value of each wavelet sub-band image in (5) can be calculated by

$$E^1 = \sum_m \sum_n \left| C^1(m, n) \right|^2. \tag{6}$$

Define the ratio of E^1_{LH} and E_{raw} to be the LH wavelet subimage energy ratio, denoted as R_{LH}

$$R_{LH} = \frac{E^1_{LH}}{E_{\text{raw}}} = \frac{E^1_{LH}}{E^1_{LL} + E^1_{LH} + E^1_{HL} + E^1_{HH}}. \tag{7}$$

Similarly, R_{HL} can be obtained by calculating the ratio of E^1_{HL} and E_{raw}.

2) Rotation Symmetry: This section summarizes the properties of the WSER varying with the rotation angle of image targets.

By rotating geometric figures in Fig. 2 by $[-180°, 180°]$, performing DWT on the rotated images, and calculating R_{LH}, the changes of R_{LH} over the rotation angles are obtained, as shown in Fig. 3. As it can be seen, R_{LH} is sensitive to the change in the rotation angle of the geometric figures except for circles. For the rectangle and ellipse, which have 180-degree rotation symmetry, the R_{LH} has the same two peak values, located at $0°$ and $180°$; For the square and circle, which have 90-degree rotation symmetry, the R_{LH} has the same four peak values, located at $-180°$, $-90°$, $0°$, and $180°$.

As can be seen from Fig. 3, the change of R_{LH} over the rotation angle of input images also has rotation symmetry, which corresponds to the rotation symmetry of geometric figures.

Similarly, R_{HL} has rotation symmetry. However, for geometric figures like rectangles and ellipses, which have different lengths of the long axis and short axis, the peak value of R_{HL} exhibits a $90°$ shift from that of R_{LH}.

3) Translation Invariance and Scale Invariance: This section theoretically derives the characteristics of the WSER R_{LH} changing with the translation and scale transformation of the image. To simplify the derivation, we assume that the pixel value of the target area in the input image $f(x, y)$ is constant, and the pixel value of the background area is 0. The reason for such an assumption is that the target area on the sonar image is a usually a strong echo, and the background area near the target area is usually a weak echo. The geometry figures shown in Fig. 2 meet such an assumption. Because the R_{LH} used in this article is the ratio of the energy of a single-level subimage C^1_{LH} to the total energy, only the single-level DWT is considered in the theoretical derivation process.

Performing single-level DWT on $f(x, y)$, the subimage C^1_{LH} can be obtained. Denote the wavelet basis corresponding to C^1_{LH} as $\psi^{0,m,n}_{LH}(x, y)$. According to (3), $\psi^{0,m,n}_{LH}(x, y) = \phi^{0,m,n}(x) \cdot \psi^{0,m,n}(y)$. Then, the value of C^1_{LH} at (m, n) as $C^{1,m,n}_{LH}$ can be represented as

$$C^1_{LH}(m, n) = \left\langle f(x, y), \psi^{0,m,n}_{LH}(x, y) \right\rangle. \tag{8}$$

On the basis of (8), E^1_{LH} and R_{LH} can be calculated through (6) and (7).

First, we derive the characteristics of the WSER R_{LH} changing with the translation transformation of images. If we translate the image by a and b along the x-axis and y-axis, respectively, the subimage of the translated image $f(x - a, y - b)$ can be denoted as $C'^1_{LH}(m, n)$, which can be written as

$$\begin{aligned} C'^1_{LH}(m, n) &= \left\langle f(x - a, y - b), \psi^{0,m,n}_{LH}(x, y) \right\rangle \\ &= \left\langle f(x', y'), \psi^{0,m,n}_{LH}(x' + a, y' + b) \right\rangle \\ &= \left\langle f(x', y'), \psi^{0,m-a,n-b}_{LH}(x', y') \right\rangle \\ &= C^1_{LH}(m - a, n - b). \end{aligned} \tag{9}$$

This result implies that when the image is translated, the corresponding subimage is also translated. If displacements a and b still enable the translated geometric figures to fall within the image boundary, then the target area on the subimage is also within the subimage boundary. We denote the energy of the subimage corresponding to the translated image as E'^1_{LH}, and

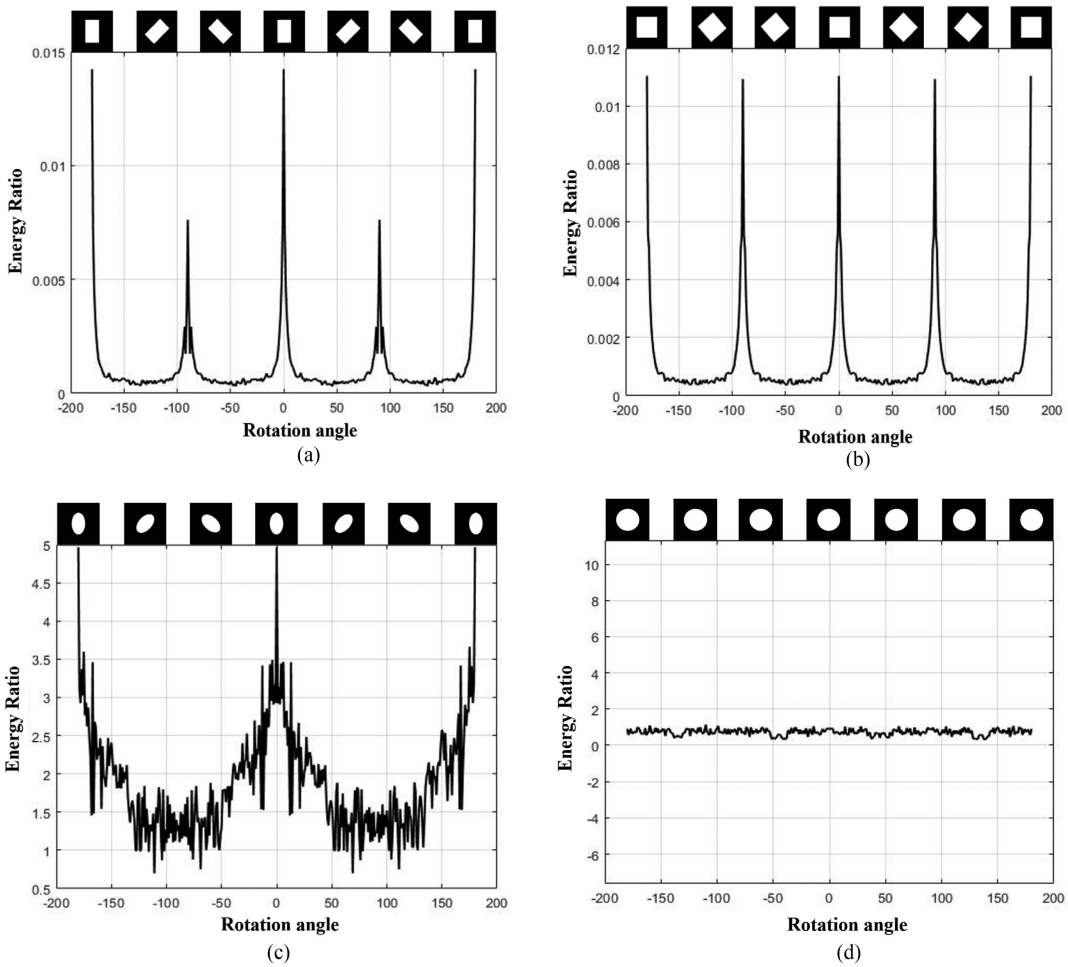

Fig. 3. Changes of R_{LH} over the rotation angle of input images. (a) Rectangle. (b) Square. (c) Ellipse. (d) Circle.

then

$$E'^1_{LH} = \sum_m \sum_n \left|C'^1_{LH}(m,n)\right|^2 = \sum_m \sum_n \left|C^1_{LH}(m,n)\right|^2 = E^1_{LH}. \tag{10}$$

This demonstrates that if the geometric figures are still within the image boundary after shifting a and b pixels, the energy $E'^1_{LH}$of the shifted subimage is the same as E^1_{LH}. Because the shapes of the geometric figures inside images do not change before and after translation, E_{raw}does not change before and after the image translation. Therefore, the WSER R_{LH} has a certain translation invariance.

Then, we derive the characteristics of the WSER R_{LH} changing with the scale transformation of images. Let the length of a target on the input image along the x-axis be Δx, and the length of the target on the subimage along the m-axis be Δm. If we enlarge the target area s_x times along the x-axis, the image of the enlarged target can be denoted as $f(x/s_x, y)$, and the length of the target on $f(x/s_x, y)$ is $s_x \cdot \Delta x$. We denote the subimage of $f(x/s_x, y)$ as C''^1_{LH}, and then

$$C''^1_{LH}(m,n) = \left\langle f(x/s_x, y), \psi^{0,m,n}_{LH}(x,y) \right\rangle. \tag{11}$$

Because only the length of the target on the input images increases by $s_x \cdot \Delta m$ times, the pixel values of $f(x/s_x, y)$ does not change. Therefore, only the length of the target area on the subimage changes to $s_x \cdot \Delta m$, the pixel values in the target area of the subimages C''^1_{LH} and C^1_{LH} do not change. Thus, the energy E''^1_{LH} of the subimage C''^1_{LH} can be derived as

$$\begin{aligned} E''^1_{LH} &= \sum_{s_x \cdot \Delta m} \sum_{\Delta n} \left|C''^1_{LH}(m,n)\right|^2 \\ &= s_x \cdot \sum_{\Delta m} \sum_{\Delta n} \left|C''^1_{LH}(m,n)\right|^2 \\ &= s_x \cdot \sum_{\Delta m} \sum_{\Delta n} \left|C^1_{LH}(m,n)\right|^2 \\ &= s_x \cdot E^1_{LH}. \end{aligned} \tag{12}$$

In other words, when the target of $f(x, y)$ enlarges s_x times along the x-axis, the energy E''^1_{LH} of the subimage C''^1_{LH} also increases s_x times. Because the image is only enlarged by s_x times in the x-axis direction and the pixel value of the target area remains unchanged, the total energy E_{raw} of the enlarged image also increases s_x times. Therefore, R_{LH} does not change.

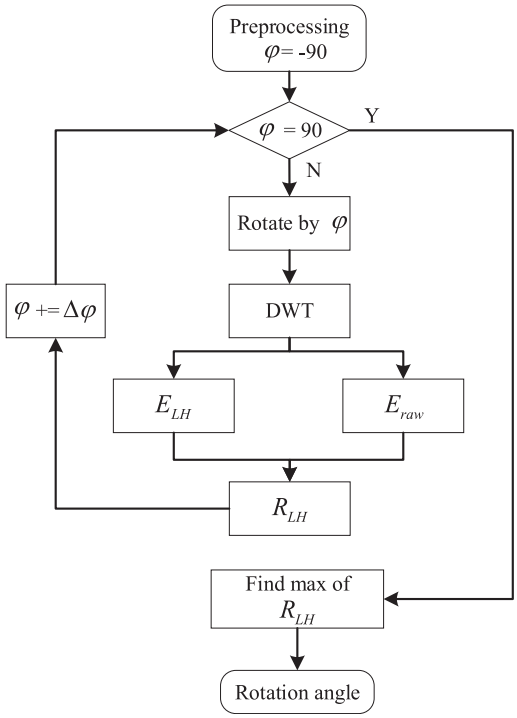

Fig. 4. Calculation flow of the proposed orientation estimation algorithm via the WSER.

Moreover, by a similar derivation process, we can draw the conclusion that R_{LH} also does not change if the target area is enlarged by s_y times along the y-axis.

In summary, R_{LH} and R_{HL} have certain translation and scale invariance. Such translation and scale invariance can ensure that the rotation symmetry of R_{LH} and R_{HL} does not change when the positions or sizes of the targets inside the images change within a certain range.

C. Orientation Estimation Algorithm of Rotated Targets via the WSER

Rotation symmetry is the basis of using R_{LH} or R_{HL} to estimate the orientation of image targets. If the image target rotates to a new orientation, the position of the maximum value of R_{LH} or R_{HL} also shifts accordingly. Based on this property, a novel orientation estimation algorithm for rotated image targets via the WSER is proposed, as shown in Fig. 4.

The calculation flow of the proposed algorithm is divided into three main steps.

Step 1: Image preprocessing. Preprocessing is done as follows:

1) convert image into grayscale image;
2) resize the image to $N \times N$;
3) denoise the image;
4) add a circle mask with the diameter N to the image, and set the values of the pixels outside the mask to 0.

Step 2: Calculation of $\{R_{LH}(\varphi)\}$. The calculation is done as follows:

1) rotate the input image by φ, bilinear interpolation is used in the rotation operation;
2) perform DWT to the rotated image, Harr wavelet is chosen as the wavelet basis;
3) calculate $R_{LH}(\varphi)$ through (6) and (7);
4) gradually increase φ at a sampling interval of $\Delta\varphi$, and repeat (a), (b), and (c) until the $R_{LH}(\varphi)$ corresponding to all the φ is calculated. All the $R_{LH}(\varphi)$ values form the $\{R_{LH}(\varphi)\}$.

Step 3: Orientation estimation. Find the maximum value of $\{R_{LH}(\varphi)\}$, and take the corresponding rotation angle as the estimation of orientation. That is

$$\hat{\varphi} = \arg\max_{\varphi} \{R_{LH}(\varphi)\}. \tag{13}$$

The $\hat{\varphi}$ estimated by R_{LH} is the orientation of the long axis of the target. In addition, the $\hat{\varphi}$ can also be estimated by R_{HL}, and the estimation is the orientation of the short axis of the target.

In terms of computational complexity, the two main factors that affect the computational load of the proposed algorithm are the input image size N and the rotation times M. When using the fast wavelet transform algorithm [20], [21] to calculate the wavelet decomposition coefficients of the input image, the computational load is proportional to $N \times N$. Because the calculation flow in Fig. 4 needs M times of DWT to the input image, the computational complexity can be expressed as $O(M \times N \times N)$. This shows that given a fixed input image size N, the computational complexity or computational time is linearly related to the rotation times M. And M is linearly related to $1/\Delta\varphi$ when the rotation angle interval of the input image is given. Therefore, computation time T is linearly related to $1/\Delta\varphi$. That is

$$T \propto \frac{1}{\Delta\varphi}. \tag{14}$$

In terms of robustness, it is known from the translation and scale invariance of R_{LH} that, the proposed algorithm has the robustness to scale and translation transformations of image targets. Such robustness implies that the proposed method can be used to estimate the target orientation after a practical target detection task, in which the targets in the detected image area are usually of different sizes and usually have different translations.

III. Results and Discussions

A. Dataset Preparation for Orientation Estimation of Rotated Sonar Image Targets

On the basis of target detection results from actual sonar images, this section describes the production of an orientation estimation dataset of sonar image targets. The dataset consists of detected rotated target images, and the rotation orientation labels of targets. The orientation is manually labeled as the rotation angle of the long axis of targets, and the function of the orientation labels is to evaluate the precision of orientation estimation algorithms.

The process for making the dataset is shown in Fig. 5, and is divided into three steps.

1) Predict a rectangular bounding box of rotated targets from the original large-size sonar image. Fig. 5(a) shows how

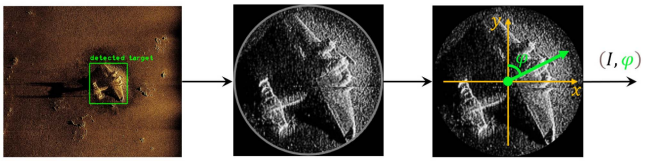

Fig. 5. Dataset creation process for the orientation estimation of sonar image targets. (a) Image in SCTD dataset. (b) Preprocessed image of target. (c) Orientation labeling.

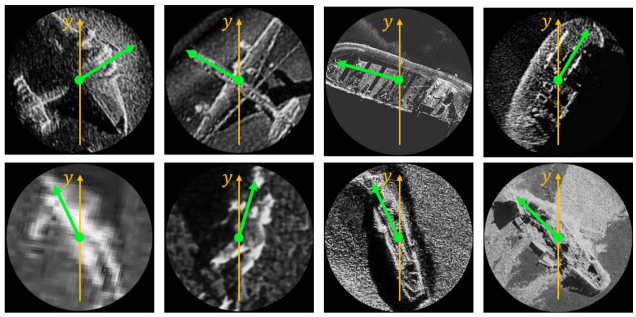

Fig. 6. Examples of labeled orientations of various sonar image targets (the labeled orientations are shown as green arrows).

faster R-CNN [22] was applied to detect targets in the sonar common target dataset [1]. The green rectangular in Fig. 5(a) shows the bounding box of rotated targets. The detection process simulates the process of predicting the target bounding box in the actual rotating target recognition task. In addition to using the faster R-CNN, other region recommendation or detection algorithms [23] can also be applied to predict the bounding boxes of targets.

2) Extract the rotated targets according to the predicted bounding box and perform preprocessing. The preprocessing follows the image preprocessing steps in Fig. 4. In this article, N was set as 224, and the average denoising algorithm with a kernel size of 11 was applied to denoise the image. The preprocessed image is shown in Fig. 5(b).
3) Label the orientation of rotated targets. As shown in Fig. 5(c), the positive direction of the y-axis is taken as, the angle between the 0° and the long axis of targets is labeled as φ, and the symbol of φ is set as negative if the long axis of targets is in a clockwise direction and positive if the long axis is in a counterclockwise direction. Finally, a self-made orientation estimation dataset composed of rotated target images I and the corresponding orientation labels φ was obtained. The self-made dataset includes 124 "ship," 19 "aircraft," and 27 "dummy" targets. Some examples of the labeled orientations of various sonar image targets from the self-made dataset are shown in Fig. 6.

B. WSER of Sonar Rotated Targets

The WSER of "aircraft," "dummy," and "ship" images shown in Fig. 6, was calculated according to the process shown in Fig. 4, φ is sampled within $[-180°, 180°]$, and the sampling interval $\Delta\varphi$ is set to $1°$. The calculated $\{R_{LH}(\varphi)\}$ is shown in Fig. 7.

By comparing Fig. 7 with Fig. 3, we can find that the overall shape of the R_{LH} curves of the sonar image targets is similar to that of the geometric figures in Fig. 3. The R_{LH} of these sonar image targets have approximately 180-degree rotation symmetry. When the value of φ is in $[-90°, 90°]$, R_{LH} has only one peak point when the long axes of targets are vertical.

However, unlike the R_{LH} curves of the geometric figures, the R_{LH} curves of the sonar image targets are likely to have more local maximum points. This is because the shape of the target on the actual sonar image is irregular. For example, although both Fig. 5.3(c) and (d) shows the R_{LH} curves of "dummy" targets, the poses of the targets are different, which causes a great difference in the shapes of these two R_{LH} curves. In addition, the R_{LH} curves of the sonar image targets are likely to have stronger noise, this is because the real sonar images have stronger noise than the geometric figures in Fig. 2.

In summary, this section has demonstrated that: 1) The R_{LH} of actual images have rotational symmetry to the rotation of sonar images. Based on the rotational symmetry, the proposed orientation estimation algorithm shown in Fig. 4 can be applied to estimate the rotation orientation of the sonar image targets. 2) The R_{LH} curves of the sonar image targets are likely to have more local maximum points and stronger noise, so$\Delta\varphi$ should be small enough to sample the true maximum of R_{LH} and accurately estimate the long axis of the target. However, the smaller $\Delta\varphi$ is, the slower the proposed algorithm is. To make a tradeoff between precision for orientation estimation and speed for orientation estimation, Sections III-C and III-D will evaluate the precision, and speed of the proposed orientation estimation algorithm under different $\Delta\varphi$.

C. Precision for Orientation Estimation

This section evaluates the orientation estimation precision for the algorithm proposed in Fig. 4. The experimental process to calculate orientation estimation precision is shown in Fig. 8, which is divided into three steps as follows.

1) Input an image I_i from the self-made dataset.
2) Set $\Delta\varphi = 1$, calculate $\{R_{LH}(\varphi)\}$ according to Fig. 4, and estimate the orientation of the target inside the input image.
3) Calculate the estimation error. Let the labeled value and estimated value of the orientation of the ith image be φ_i and $\hat{\varphi}_i$, respectively. The mean absolute error (MAE) of orientation estimation is calculated by

$$\text{MAE} = \frac{1}{N_I}\sum_{i=1}^{N_I}|\varphi_i - \hat{\varphi}_i| \tag{15}$$

where N_I is the number of images in the self-made dataset. The lower the value of MAE is, the higher the precision of orientation estimation is.

Following the calculating process in Fig. 8, we show the MAE of orientation estimation for "aircraft," "dummy," "ship," and the average value of MAE are shown in Table I. For comparison, Table I also gives the MAE of three traditional methods and the CNN regression method. The Hough transform method uses the Canny operator to extract edges. The CNN regression method

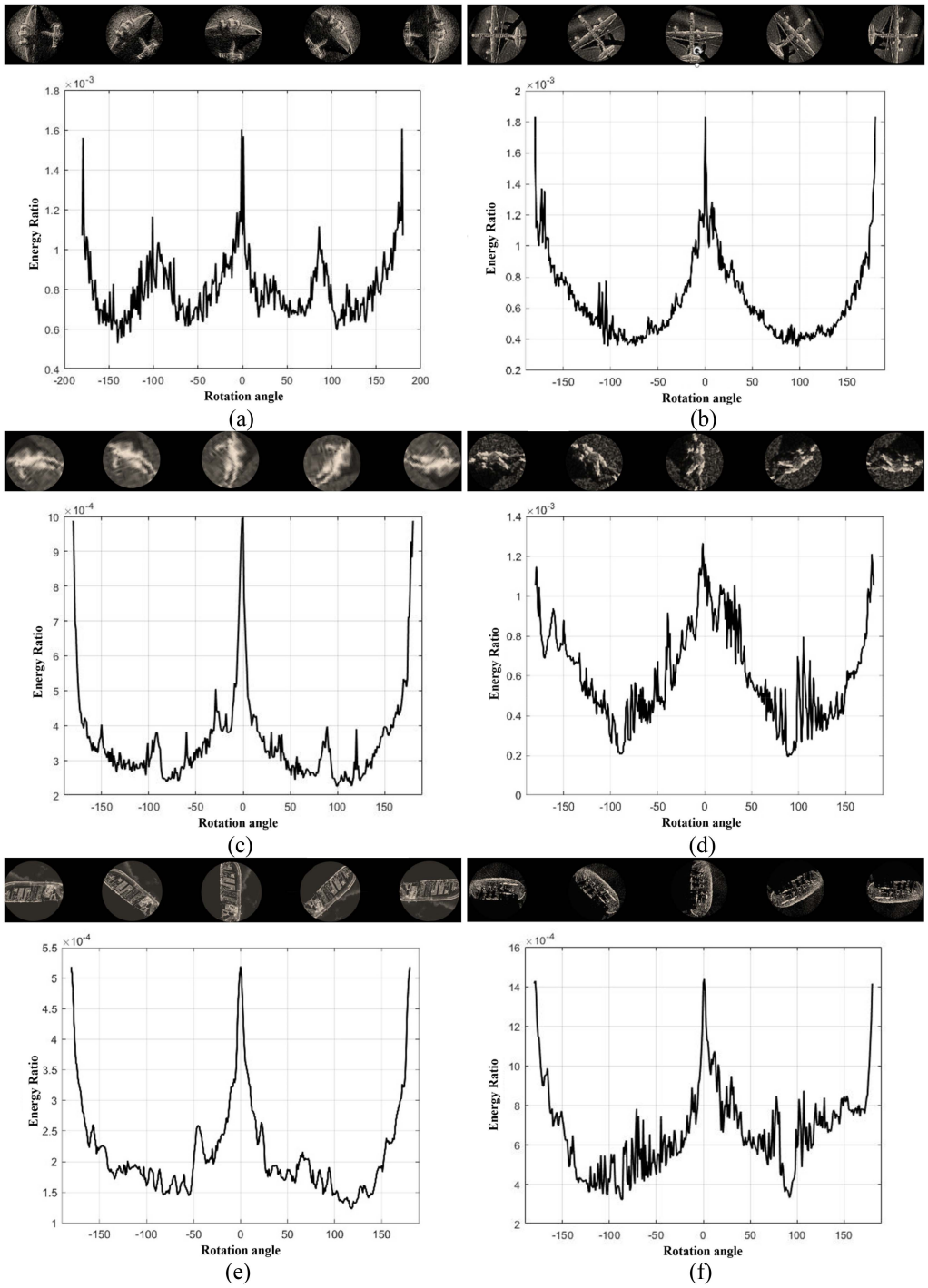

Fig. 7. Wavelet subimage energy ratio of sonar targets. (a) Aircraft 1. (b) Aircraft 2. (c) Dummy 1. (d) Dummy 2. (e) Ship1. (f) Ship2.

TABLE I
MAE FOR ORIENTATION ESTIMATION OF DIFFERENT METHODS

Methods	aircraft	dummy	ship	average
Hough Transform method	30.9815	27.3716	23.4083	27.2538
Radon Transform method	34.5250	27.1579	18.9259	26.8696
Fourier Transform method	25.8992	24.1713	18.8686	22.9797
CNN Regression method	11.0508	8.6256	10.6667	10.1144
The proposed method (without mask)	11.9083	10.3684	6.8148	9.6971
The proposed method (with mask)	11.2667	8.3158	4.3333	**7.9719**

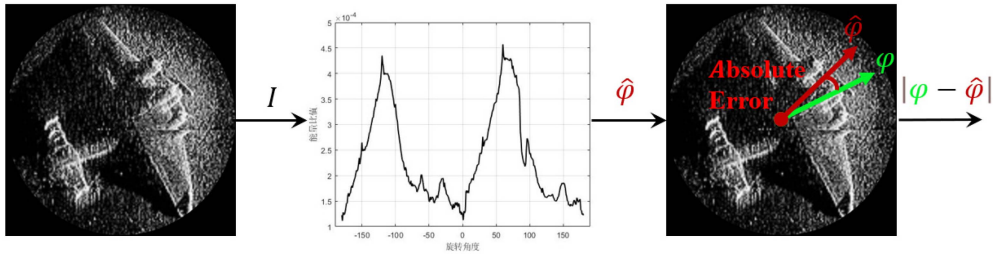

Fig. 8. Calculating process of MAE of orientation estimation.

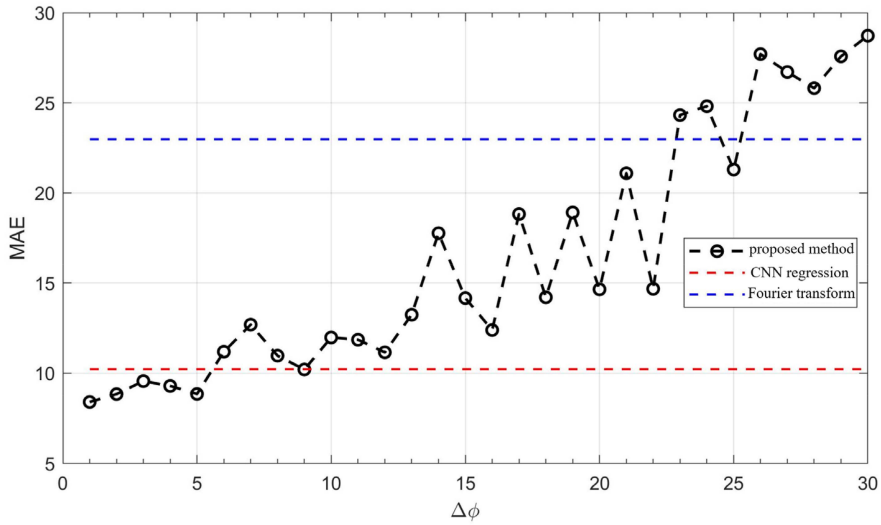

Fig. 9. Curve of average MAE under different $\Delta\varphi$.

uses a ResNet50 network, which has two output neurons, to predict the sine and cosine of the rotation angle. Because the CNN regression model needs to be trained, the self-made dataset made was divided into training and validation data subsets in a ratio of 1:1. Moreover, owning to the small sample amount inside the self-made sonar image dataset, ImageNet pretrained weights are used to initialize the training of CNN regressor. After 20 epochs of training, the trained CNN regressor is applied to calculate the average MAE on the validation dataset.

The MAE for orientation estimation of different methods is shown in Table I. The proposed orientation estimation method achieves a very low average MAE, which is only 7.9719°. As a comparison, the lowest MAE achieve by traditional methods is 22.9797, and the MAE of the CNN regression method is 10.1144. Therefore, the orientation estimation precision of the proposed method is not only significantly better than that of traditional methods that do not rely on supervised learning, but also better than that of the CNN regression method that uses deep supervised learning. This demonstrates that the proposed method can achieve higher orientation estimation precision for sonar images.

In addition, comparing the bottom two rows of Table I, it can also be found that if the circular mask is not used, the MAE of the proposed method slightly increases to 9.6971. This validates that adding a circular mask is helpful to improve the orientation estimation accuracy of the proposed method.

Table I gives the MAE of the proposed method under $\Delta\varphi = 1°$. The input images are rotated 360 times to calculate the $R_{LH}(\varphi)$, so this may require excessive computations. The most direct acceleration trick is to increase $\Delta\varphi$, but how to choose $\Delta\varphi$ or how the $\Delta\varphi$ affects the average MAE is still unknown. To answer this question, $\Delta\varphi$ was set to $\{1, 2, 3, \ldots, 30\}$ in turn, and the average MAE is calculated under each $\Delta\varphi$. The curve of the average MAE under different $\Delta\varphi$ is shown in Fig. 9. For clear comparison, the MAE of Fourier transform method and CNN regression method are plotted as horizontal dash lines in Fig. 9. As it can be seen, with the increase of $\Delta\varphi$, the average MAE shows an increasing trend, i.e., the average MAE of this method is positively related to the value of $\Delta\varphi$. When $\Delta\varphi \leq 5°$, the MAE of proposed method is lower than that of the CNN regression method; When $\Delta\varphi \leq 23°$, the MAE of the proposed method is lower than the Fourier transform method. This implies that it is possible to achieve orientation estimation precision higher than the existing methods using a higher $\Delta\varphi$. Increasing $\Delta\varphi$ is beneficial for reducing the computations of the proposed method.

It can also be found from Fig. 9 that the MAE curve exhibits instability if $\Delta\varphi > 13°$. The reason can be explained using the characteristics of the R_{LH} curve of real sonar images. Owing to the large difference of the pose of underwater targets and image noise, the R_{LH} curves of sonar image targets have multiple local maximum values and obvious noise. When the sampling

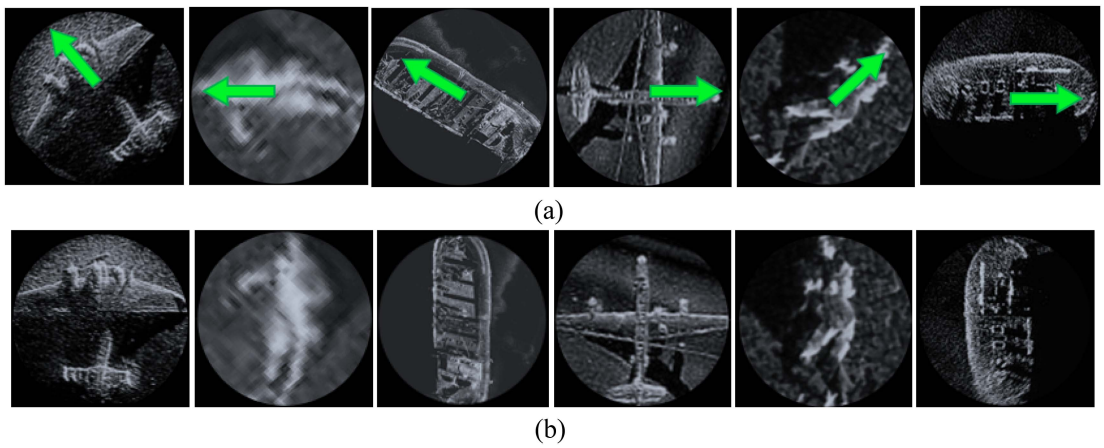

Fig. 10. Examples of estimated orientation and tilt correction using the proposed method (watch https://www.bilibili.com/video/BV1er4y1V7j3/ for more results). (a) Sonar images and estimated orientations. (b) Tilt correction of sonar images using the estimated orientation.

interval $\Delta\varphi$ is small, the points near the maximum of R_{LH} can be easily sampled, so the orientation estimation error is small. In addition, when $\Delta\varphi$ is small, changing $\Delta\varphi$ by a small angle does not greatly change the shape of sampled $R_{LH}(\varphi)$. So, a small change of $\Delta\varphi$ is not likely to significantly affect the orientation estimation error. This result implies that the MAEcurve does not fluctuate greatly. However, when the sampling interval $\Delta\varphi$ is large, the points near the maximum of R_{LH} are hardly sampled, so the orientation estimation error is large. In addition, when $\Delta\varphi$ is large, the change in $\Delta\varphi$ greatly changes the shape of sampled $R_{LH}(\varphi)$. This result implies that the MAEcurve may fluctuate greatly. The abovementioned analysis is consistent with the results of Fig. 9.

When $\Delta\varphi = 5$, the estimated orientations of the sonar image targets are shown in Fig. 10(a). The proposed orientation estimation method can precisely estimate the orientations of the long axes of the rotating targets. Moreover, using the estimated target orientations, the arbitrarily rotated targets can be aligned to a unified orientation. As it was shown in Fig. 10(b), the target orientations of the aligned images are close, and the similarities between two images of the same category are stronger. One of the applications of the aligned images is to improve the performance for rotated target classification. To demonstrate this, a CNN classification model is trained and tested on the aligned sonar images. The results are shown in in the appendix of this article, which shows that the rotated target classification performance can be improved through the aligned images. This demonstrates that precisely orientation information can facilitate the recognition of arbitrarily rotated targets.

D. Speed for Orientation Estimation

Section III-C has analyzed the orientation estimation precision of the proposed method under different $\Delta\varphi$, this section analyzes the calculation speeds of the proposed method under different $\Delta\varphi$. For fair comparison, all methods are implemented on an Intel(R) Core (TM) i7-8550U@1.80 GHz CPU. The CNN Regression method is implemented by the Pytorch torchvision library, and the other methods were implemented by MATLAB.

TABLE II
AVERAGE CALCULATION TIME OF DIFFERENT METHODS

Methods	Average calculation time (s)
Hough transform method	0.0409
Radon transform method	0.0506
Fourier transform method	0.0523
CNN Regression method	1.2596
The proposed method (with mask)	0.2272

$\Delta\varphi$ is set to $\{1, 2, 3, \ldots, 30\}$ in turn, and the total calculation time is counted for each $\Delta\varphi$. Then, the average calculation time can be obtained by dividing the total calculation time by the number of the images in the self-made dataset, i.e., N_I. The curve of average calculation time under different $\Delta\varphi$ is shown in Fig. 11. As it is shown, the curve has the shape of an inverse proportional function, which is consistent with the theoretical analysis of (9).

Fig. 11 also shows the calculation time for the orientation estimation of existing methods. For a clear comparison, the average calculation time of the Fourier transform method and CNN regression method are plotted as horizontal dash lines in Fig. 11. As it can be seen, when increasing $\Delta\varphi$, the average calculation time of the proposed method is always lower than that of the CNN regression method, and gradually approaches the average calculation time of the Fourier transform method. From Fig. 9, we know that the orientation estimation precision of the proposed method is better than the existing method when $\Delta\varphi \leq 5$. Therefore, the average calculation time of the proposed method at $\Delta\varphi = 5$ is quantitatively compared with existing methods, and the results are shown in Table II. At this time, the average calculation time of the proposed method is about 4.34 times that of the Fourier transform method and one-sixth of the CNN regression method.

According to the results in Fig. 11 and Table II, we know that the orientation estimation time of the proposed method has an inverse proportional relationship with $\Delta\varphi$. To achieve the orientation estimation precision that is higher than existing

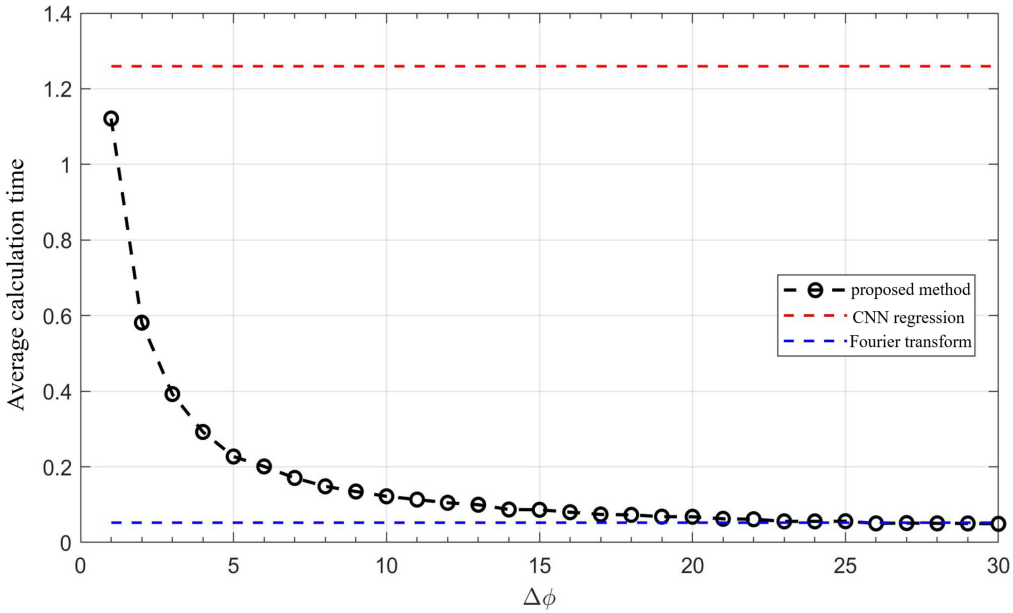

Fig. 11. Curve of average calculation time under different $\Delta\varphi$.

methods, $\Delta\varphi$ should be smaller than or equal to 5°. In this case, the proposed method is slightly slower than traditional methods but six times faster than the CNN regression method.

E. Robustness to Translation and Scale Transformation

From the conclusions of (10) and (12), we know that WSER R_{LH} has some translation and scale invariance. Translation and scale invariance can ensure that the rotational symmetry of R_{LH} and R_{HL} does not change when the position and size of the image targets change within a certain range. Therefore, the proposed method is robust to translation and scale transformation of image targets.

In order to examine the translation invariance of WSER, a "ship" target is randomly selected from the self-made sonar image dataset and is translated along the horizontal direction by 5, 20, 15, 20, and 25 pixels. The R_{LH} curve under different translation pixels is shown in Fig. 12(a). As it is shown, the R_{LH} curves under different translations are highly overlapped. This demonstrates that a certain degree of translation transformation does not affect the values of WSER. Such a result is consistent with the conclusion of (10) that R_{LH} has translation invariance. Translation invariance implies that small translations to the input image do not affect the orientation estimation performance of the proposed methods.

To evaluate the orientation estimation performance under different translations of input images, all the sonar images are translated along the horizontal direction by Δx pixels, and then the MAE is calculated using the process in Fig. 8. Changing Δx in $\{-30, -25, 20, 15, 10, 5, 0, 5, 10, 15, 20, 25, 30\}$, the MAE curve under different translations can be plotted as the black dashed line in Fig. 12(b). For comparison, the MAE curve under different translations of the CNN regression method is plotted as the red dashed line in Fig. 12(b). It can be seen from the figure that the MAE of the proposed method is relatively stable when image targets are translated, while the MAE of the CNN regression method is sensitive to minor translations.

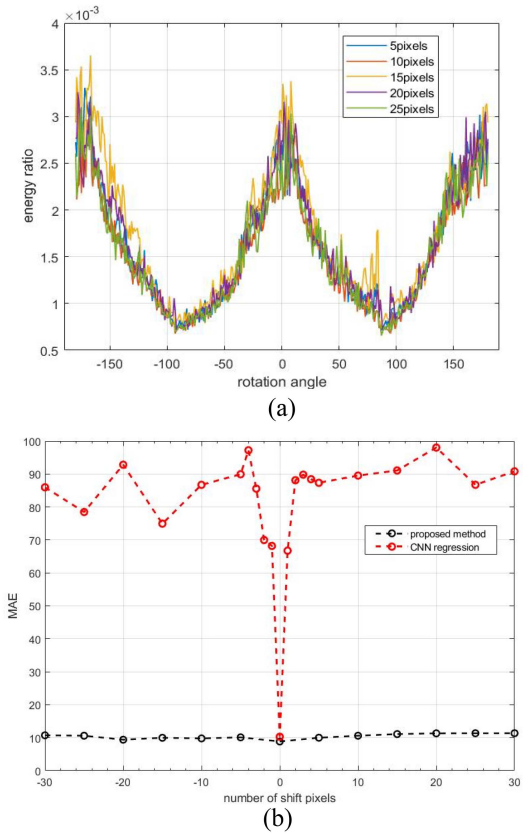

Fig. 12. R_{LH} and MAE under different translations. (a) R_{LH} curves under different translations. (b) MAE under different translations.

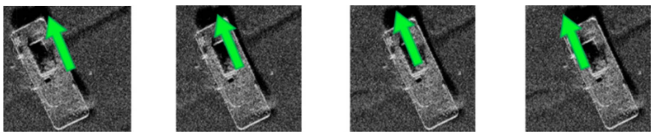

Fig. 13. Orientation estimation for an image target under different translations.

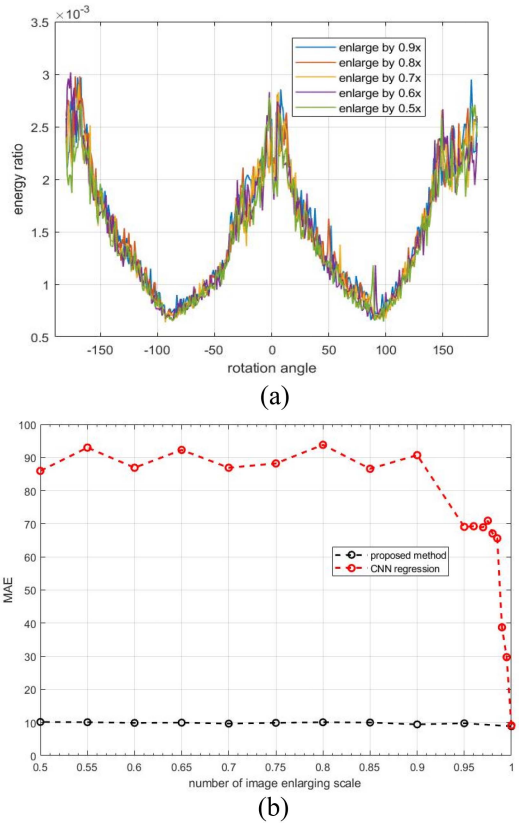

Fig. 14. R_{LH} and MAE under different target scales. (a) R_{LH} curves under different target scales. (b) MAE under different target scales.

Other existing methods are not chosen because their MAE is significantly higher than that of the proposed method, so it does not make much sense to examine their robustness.

In brief, the orientation estimation performance of the proposed method is robust to sonar image translations, and the robustness is stronger than that of the CNN regression method. Translation invariance means that when the detected target is not located at the center of the image, i.e., targets are under different translations, the proposed method can still precisely estimate the orientation of sonar image targets, as shown in Fig. 13.

In order to examine the scale invariance, a "ship" target was randomly selected from the self-made dataset and was enlarged by 0.9, 0.8, 0.7, 0.6, and 0.5 times. The enlarged image was padded with 0, and the padded image has a size of 224×224. After such preprocessing, the size of the targets can be changed while the input image size of orientation estimation algorithms is unchanged. The R_{LH} curve under different target scales is shown in Fig. 14(a). As it is shown, the R_{LH} curves under different scales are highly overlapped, which is consistent with the conclusion of (12) and demonstrates that R_{LH} has scale

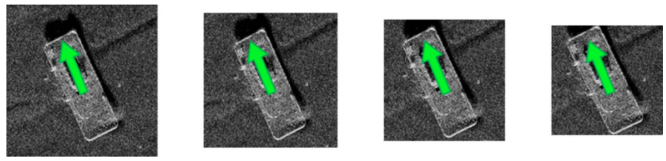

Fig. 15. Orientation estimation for an image target under different scales.

invariance. Such scale invariance implies that small-scale transformation to the targets of the input image does not affect the orientation estimation performance of the proposed methods.

To evaluate the orientation estimation performance under different scale transformations to image targets, sonar image targets are enlarged by s times, and then the MAE is calculated using the process in Fig. 8. s is named the target scale in this article. By adjusting s in $\{1, 0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55, 0.5\}$, the MAE curve under different target scales can be plotted as the black dashed line in Fig. 14(b). For comparison, the MAE curve under different target scales of the CNN regression method is plotted as the red dashed line in Fig. 14(b). As it is shown, the MAE of the proposed method is relatively stable when image targets are on different scales, while the MAE of the CNN regression method is sensitive to merely small-scale transformations to targets.

The orientation estimation performance of the proposed method is robust to sonar image scale transformations of sonar image targets, and the robustness is stronger than the CNN regression method. Scale invariance means that no matter whether the predicted bounding box is too large or small, the proposed method can still precisely estimate the orientation of sonar image targets, i.e., targets are under different scales, which is shown in Fig. 15.

This section has used real sonar images to demonstrate that the WSER is translational and scale invariant to some extent. Such invariance makes the proposed orientation method via the WSER robust to both the translation and scale transformations of sonar images, which implies that it can achieve robust orientation estimation for targets detected from sonar images.

IV. Conclusion

Rotated target recognition is a crucial issue of sonar image recognition, but little attention has been paid to orientation estimation. The orientation information of sonar image targets is beneficial to the automatic target recognition of sonar images, but is also a challenging area.

To achieve high-performance orientation estimation for sonar image targets, this article has proposed a novel orientation estimation method via the wavelet subimage energy ratio. By analyzing the properties of the WSER varying with the rotation angle of image targets, the energy ratio of the wavelet subimage is derived and demonstrated to have rotation symmetry, translation invariance, and scale invariance. Rotation symmetry is the basis of the proposed method to estimate target orientations, and translation and scale invariance can enhance the robustness of the proposed method.

Because of these excellent properties, the proposed method has outperformed existing methods in estimation precision. It has achieved a strong orientation estimation MAE of 7.9°. To achieve strong estimation MAE, the proposed method needs to use a relatively small rotation interval $\Delta\varphi$. As a result, the proposed method has been found to be six times faster than the CNN regression method while slightly slower than traditional methods. In addition, the proposed method has been found to be robust to both the translation and scale transformations of sonar images, which implies that it can achieve robust orientation estimation for targets detected from sonar images. In addition, it is important to emphasize that this method does not require the use of supervised training.

The proposed orientation method via WSER is the first to use wavelet subimage energy as an orientation descriptor, while the existing orientation estimation methods only use wavelet decomposition as an edge detector tool. Therefore, the proposed method not only enriches the orientation description method, but offers a new view to advance the understanding of wavelet decomposition and expands the application scenario of the wavelet decomposition.

The proposed orientation estimation method can be applied to orientation estimation tasks of other types of images, such as remote sensing images or medical images, which have to tackle the problem of arbitrary target orientations. Besides, the proposed orientation estimation method has been demonstrated to be beneficial for rotated target recognition (see Appendix).

Future studies should integrate the proposed orientation estimation method into a unified rotated target recognition framework, which can simultaneously output the positions, categories, and orientations of image targets.

APPENDIX

Existing rotated target classification algorithms without using orientation information are shown in Fig. 16. At the training stage, existing algorithms use rotated data augmentation (RDA) to train CNN. RDA actively rotates the training images into various orientations. At the test stage, CNN trained with RDA can directly classify arbitrarily rotated targets.

To demonstrate that this article can facilitate the recognition of arbitrarily rotated targets, we present a rotated target classification algorithm that uses the orientation information of rotated targets, and compare it with existing rotated target classification algorithms without using orientation information. The proposed rotated target classification algorithm is shown in Fig. 17, and has the same workflow of the training and testing. The workflow has three steps.

1) First, precisely estimate targets' orientations by the proposed orientation estimation method via the WSER.
2) Second, align the arbitrarily rotated image targets to the same orientation based on the estimated orientation.
3) Finally, train CNN classification models on the aligned training sonar images, or use the trained CNN to classify rotated targets via aligned images.

We apply these two algorithms to the rotated sonar images produced in Section III-A. Because the number of samples is small, the ImageNet pre-trained model is applied to initialize the training of CNNs. The classification accuracies for test images of these two methods are shown in Table III. As it is shown, the average classification accuracy of the proposed algorithm is 13.28% higher than that of the existing RDA method. Moreover, the existing RDA method requires ResNet152 CNN, which has a larger number of parameters, to achieve high classification accuracy. By contrast, the proposed rotated target classification algorithm via aligned images can achieve higher classification accuracy through the ResNet18 CNN, which has fewer parameters.

In terms of calculation time, the total calculation time of the proposed rotated target classification algorithm consists of the orientation alignment time and CNN inference time. The total calculation time of the data augmentation method only includes the CNN inference time. However, because the proposed algorithm requires CNNs with fewer numbers of parameters, the inference time of the CNNs is also shorter than that of CNNs used by the existing RDA method. As a result, the total calculation time cost by the proposed algorithm is lower than that of the existing RDA method (see Table IV).

In summary, compared with existing RDA methods, the proposed rotated target classification algorithm via aligned images can achieve higher classification accuracy than existing methods, and cost less inference time. This demonstrates that precise orientation information can facilitate the recognition of arbitrarily rotated targets.

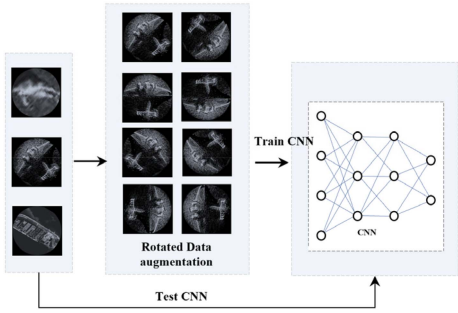

Fig. 16. Existing rotated target classification algorithm without using orientation information.

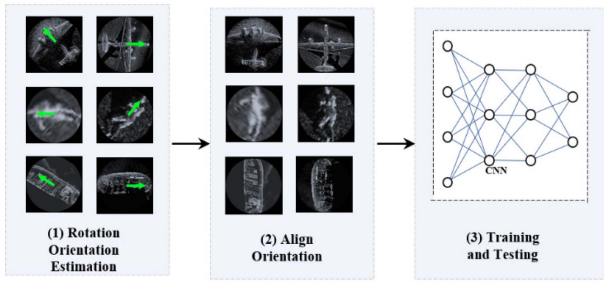

Fig. 17. Proposed rotated target classification algorithm via aligned images.

TABLE III
CLASSIFICATION ACCURACIES

Method		airplane (%)	'dummy (%)	ship(%)	Average (%)
Existing RDA method	ResNet18	66.62	59.11	63.3	63.01
	ResNet152	84.80	89.50	81.51	85.27
The proposed method (ResNet18)		100.00	100.00	95.65	**98.55**

TABLE IV
AVERAGE CALCULATION TIME

Method		Average calculation time (s)
Existing RDA method (ResNet152)		3.5299
The proposed method (ResNet18)	orientation alignment	0.2272
	CNN inference	1.1607
	total	**1.3879**

REFERENCES

[1] P. Zhang, J. Tang, H. Zhong, M. Ning, D. Liu, and K. Wu, "Self-Trained target detection of radar and sonar images using automatic deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, Jul. 2021, doi: 10.1109/TGRS.2021.3096011.

[2] D. P. Williams, "On the use of tiny convolutional neural networks for human-expert-level classification performance in sonar imagery," *IEEE J. Ocean. Eng.*, vol. 46, no. 1, pp. 236–260, Jan. 2021, doi: 10.1109/JOE.2019.2963041.

[3] Z. Peng, J. Tang, H. Zhong, M. Ning, and Y. Fan, "Rotated target recognition of sonar images via convolutional neural networks with rotated inputs," in *Proc. 14th Int. Conf. Digit. Image Process.*, Wuhan, China, 2022, pp. 1–7.

[4] I. D. Gerg and V. Monga, "Structural prior driven regularized deep learning for sonar image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, Jan. 2021, doi: 10.1109/TGRS.2020.3045649.

[5] M. Valdenegro-Toro, "Object recognition in forward-looking sonar images with convolutional neural networks," in *Proc. OCEANS MTS/IEEE Monterey*, 2016, pp. 1–6.

[6] X. Wang, J. Jiao, J. Yin, W. Zhao, X. Han, and B. Sun, "Underwater sonar image classification using adaptive weights convolutional neural network," *Appl. Acoust.*, vol. 146, pp. 145–154, 2019.

[7] D. Karimanzira, H. Renkewitz, D. Shea, and J. Albiez, "Object detection in sonar images," *Electronics*, vol. 9, no. 7, 2020, Art. no. 1180.

[8] Y. Yu, J. Zhao, Q. Gong, C. Huang, G. Zheng, and J. Ma, "Real-time underwater maritime object detection in side-scan sonar images based on transformer-YOLOv5," *Remote Sens.*, vol. 13, no. 18, pp. 1–28, 2021.

[9] G. Neves, M. Ruiz, J. Fontinele, and L. Oliveira, "Rotated object detection with forward-looking sonar in underwater applications," *Expert Syst. Appl.*, vol. 140, 2020, Art. no. 112870.

[10] H. Penedones, R. Collobert, F. Fleuret, and D. Grangier, "A document skew detection method using run-length encoding and the hough transform," in *Proc. 10th Int. Conf. Pattern Recognit.*, 1990, pp. 464–468.

[11] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Radon transform orientation estimation for rotation invariant texture analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 1004–1008, Jun. 2005, doi: 10.1109/TPAMI.2005.126.

[12] G. S. Peake and T. N. Tan, "A general algorithm for document skew angle estimation," in *Proc. Int. Conf. Image Process.*, 1997, pp. 230–233.

[13] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," in *Proc. Adv. Annu. Conf. Neural Inf. Process. Syst. 28*, Montreal, Quebec, Canada, 2015, pp. 2017–2025.

[14] F. Massa, R. Marlet, and M. Aubry, "Crafting a multi-task CNN for viewpoint estimation," in *Proc. Brit. Mach. Vis. Conf.*, York, U.K., Sep. 19–22, 2016, pp. 1–12.

[15] Z. Peng, J. Tang, H. Zhong, M. Ning, and Y. Fan, "Pre-rotation only at Inference-time: A way to rotation invariance," in *Proc. 14th Int. Conf. Digit. Image Process.*, Wuhan, China, 2022, pp. 1–10.

[16] A. Azulay and Y. Weiss, "Why do deep convolutional networks generalize so poorly to small image transformations," *J. Mach. Learn. Res.*, vol. 20, no. 184, pp. 1–25, 2019.

[17] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient CNN architecture design," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 116–131.

[18] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[19] T. Yao and H. Sun, *Advanced Digital Signal Processing*, 2nd ed. Wuhan, China: Huazhong Univ. Sci. Technol. Press, 1999.

[20] I. Daubechies, "Ten lectures on wavelets," in *Proc. SIAM*, 1992, pp. 56–58.

[21] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[22] L. Liu et al., "Deep learning for generic object detection: A survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, 2020, doi: 10.1007/s11263-019-01247-4.

Peng Zhang received the B.S. degree in radar engineering and the M.S. degree in signal and information processing in 2016, and 2018, respectively, from the Naval University of Engineering, Wuhan, China, where he is currently working toward the Ph.D. degree in underwater acoustical engineering.

His research interests include deep learning and its application in pattern recognition and signal processing.

Jinsong Tang received the Ph.D. degree in electronic engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1996.

He is currently a Professor with the Naval University of Engineering, Wuhan, China. His current research interests include synthetic aperture sonar (SAS), Interferometric SAS, and underwater acoustic communication.

Dr. Tang was honored in the list of the "100 Talents Programme" issued by the Chinese Academy of Science in 1999.

Heping Zhong received the B.S. and M.S. degrees in computer science and technology and the Ph.D. degree in underwater acoustical engineering from the Naval University of Engineering (NUE), Wuhan, China, in 2005, 2007, and 2011, respectively.

He is currently an Associate Professor with NUE. His current research interests include synthetic aperture radar (sonar) interferometry software development, interferometric signal processing, and parallel computing.

Haoran Wu received the B.S. degree in radar engineering, and the M.S. and Ph.D. degrees in underwater acoustical engineering from the Naval University of Engineering (NUE), Wuhan, China, in 2012, 2014, and 2018, respectively.

He is currently an Associate Professor with NUE. His current research interests include synthetic aperture radar (sonar), and radar (sonar) signal processing, and automatic target recognition.

Han Li received the B.S. degree in radar engineering in 2018 from the Naval University of Engineering (NUE), Wuhan, China, where he is currently working toward the M.S. degree in information and communication engineering.

He was doing research on SAR image processing in Tsinghua University from 2016 to 2018. His current research interests include computer vision and its application in radar (sonar) interferometry.

Yue Fan received the M.S. degree in electronic engineering in 2009 from Central China Normal University, Wuhan, China, where he is currently working toward the Ph.D. degree in radio physics.

She is currently an Associate Professor with the Naval University of Engineering, Wuhan, China. Her current research interests include synthetic aperture sonar image processing and automatic target recognition.