Small Maritime Target Detection Using Gradient Vector Field Characterization of Infrared Image

Ping Yang, Lili Dong, *Member, IEEE*, and Wenhai Xu

***Abstract*—Infrared small maritime target detection under strong ocean waves, a challenging task, plays a key role in maritime distress target search and rescue applications. Many methods based on directionality or gradient properties have proven to perform well for infrared images with heterogeneous scenarios. However, they tend to perform poorly when facing strong ocean wave background, mainly due to the following: 1) infrared images have low signal-to-clutter ratio with low intensity for small targets; 2) some waves have high local contrast that may be similar to or higher than targets. To solve these issues, a new method based on gradient vector field characterization (GVFC) of infrared images is proposed. First, we construct the gradient vector field and coarsely extract suspected targets. Then, gradient vector distribution measure (GVDM) is presented, which comprehensively integrates a synergistic homogeneity test based on Kolmogorov–Smirnov test with absolute difference standard deviation for gradient direction angle and regression analysis for gradient modulus. The proposed GVDM takes advantage of pixel-level gradient distribution property to further filtrate refined suspected targets. Moreover, gradient modulus horizontal local dissimilarity is proposed to measure the diversity of gradient modulus in horizontal direction between targets and waves, so as to enhance target saliency and suppress residual clutter simultaneously, which achieves preferable performance. Finally, a simple adaptive threshold is applied to confirm targets. Extensive experiments implemented on infrared maritime images with strong ocean waves demonstrate that the proposed method is superior to the state-of-the-art methods with respect to robustness and detection accuracy.**

***Index Terms*—Gradient vector, infrared maritime image, small target detection, vertical local dissimilarity.**

NOMENCLATURE

GVFC	Gradient vector field characterization.
GVDM	Gradient vector distribution measure.
GMHLD	Gradient modulus horizontal local dissimilarity.
GVF	Gradient vector field.
$\overrightarrow{\text{GV}}$	Gradient vector.
GM	Gradient modulus map.
BGM	Binary GM.
STs	Suspected targets.
NB	Neighboring background.
TNR	Target neighboring region.
SHT	Synergistic homogeneity test.
K-S	Kolmogorov–Smirnov.
ADSD	Absolute difference standard deviation.
CDF	Cumulative distribution function.
LRC	Linear regression classification.
RST	Refined ST.
SCR	Signal-to-clutter ratio.
SCRG	Signal-to-clutter ratio gain.
LCG	Local contrast gain.
BSF	Background suppression factor.
ROC	Receiver operating characteristic.
AUC	Area under the curve.

Manuscript received 10 June 2022; revised 18 August 2022; accepted 14 September 2022. Date of publication 10 February 2023; date of current version 15 February 2023. This work was supported in part by the Fundamental Research Funds for the Central Universities of China under Grant 3132019340 and Grant 3132019200, and in part by the High-tech Ship Research Project from Ministry of Industry and Information Technology of the Peoples Republic of China under Grant MC-201902-C01. *(Corresponding authors: Ping Yang; Lili Dong.)*

Ping Yang is with PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China.

Lili Dong and Wenhai Xu are with the Information and Communication Engineering Department, School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China (e-mail: donglili@dlmu.edu.cn; xuwenhai@dlmu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3210112

I. INTRODUCTION

DETECTING infrared maritime targets robustly is a key technology in sea surface distress target searching and rescue systems and will directly affect the effectiveness of subsequent target location and tracking [1], [2], [3]. Infrared images tend to have low SCR because of long imaging distance, and infrared small maritime targets' contours and shapes are not concrete and texture features are scarce [4]; moreover, due to the complex and variable marine weather conditions, humid and hot environment, sunlight, sea fog, and ocean waves have significant effects on imaging quality [5], which limits the infrared small maritime target detection performance; strong ocean waves typically present high radiation intensity in infrared image, which may be similar to real targets and easily submerge targets. Therefore, it remains a challenging task to develop effective infrared small maritime target detection algorithms with high performance and satisfactory accuracy.

Generally, existing infrared target detection methods can be roughly classified into two broad categories based on the number of image frames used: 1) sequential frames detection and 2) single-frame detection. Sequential frames detection methods, also known as track before detection methods, usually employ spatiotemporal information of multiframe images to identify targets, such as spatio-temporal tensor model [6], spatial–temporal local difference measure [7], multiple subspace learning and spatial–temporal infrared patch-tensor (MSLSTIPT) [8], novel spatiotemporal saliency method [9]. However, the processing procedures of sequential frames-based methods are complicated

and need more prior information; besides, their performance also relies on the information of single-frame processing. Therefore, developing a high-accuracy infrared target detection method based on a single frame image will be crucial. Single-frame detection methods, also termed as detection before track methods, have attracted much attention because of their easy implementation and the robust and high performance for detecting small targets in infrared images with complex and fickle background.

A. Related Work

In recent years, many researchers have dedicated to developing single-frame-based infrared target detection methods that generally contain the following four categories:

1) background-estimation-based method;
2) low-rank and sparsity-analysis-based method;
3) deep-learning-based method;
4) human visual system (HVS)-based method.

The conventional background-estimation-based methods design filters to predict background information followed by difference operation between raw image and estimated background. Top-hat transformation, a typical filter, utilizes mathematical morphology operations to capture difference information between the target and surrounding background and achieve target enhancement [10], [11]. Max-mean and max-median filters are also useful for acquiring background information, which conduces to infrared small target detection [12]. Soni et al. [13] employed two-dimensional least-mean-square adaptive filter to predict correlated background clutter, which was subtracted from original signal, then small targets were preserved in the residual component. However, these methods cannot eliminate the interference of complex background, such as random noise, heavy ocean waves, and strong cloudy clutter.

Recently, the low-rank and sparsity-analysis-based methods have come into much notice, and they mainly include two types: 1) low-rank matrix recovery theory and 2) tensor model. The former usually decomposes infrared image into sparse and low-rank components using the nonlocal self-correlation property of background and sparsity of targets for infrared target detection. Gao et al. [14] first introduced infrared patch-image (IPI) model that transformed target detection into the recovery of low-rank and sparse matrices. Then, low-rank matrix recovery theory has been gradually used in the infrared target detection field. He et al. [15] proposed low-rank and sparse representation model that combined sparse representation and basis of low-rank decomposition to segment target from noise and background. Wang et al. [16] added total variation regularization into the low-rank background recovery that was handled by principal component pursuit, which performs better in unsmooth and uneven background. However, they are time consuming and not suitable for practical applications. The latter, tensor model, extends patch-image into high-dimension matrixes, which is more effective for highly different scenes. Dai et al. [17] proposed reweighted infrared patch-tensor model and creatively handled with target-background separation problem as the optimization of sparse and low-rank tensors recovery. Subsequently, varied infrared patch-tensor (IPT)-based methods have been proposed, such as the partial sum of the tensor nuclear norm (PSTNN) [18] and improved IPT model based on nonconvex tensor fibered rank approximation [19]. To dig out more spatial and structural information, Zhao et al. [20] presented three-order tensor creation and Tucker decomposition. However, these methods usually have high computational complexity.

With the rapid and wide development of deep learning, it has been applied to the field of infrared target detection. Deep-learning-based methods excavate image properties by training the model on large samples, then detect infrared targets using the trained deep learning model. Generally, the deep convolutional neural network (CNN) is the commonly used network and beneficial to learning the hierarchical features of infrared images, such as lightweight CNN (L-CNN) and multisource feature cascade decision method [23], target-oriented shallow-deep features and effective small-anchor-based CNN detection method [25], and the salient target detection for infrared and visible image fusion (STDFusionNet) [24]. Generative adversarial network (GAN), another predominant deep learning model, builds an adversarial game between source images and generative images to learn the desired probability distribution. Zhao et al. [21] proposed infrared small target detection with GAN (IRSTD-GAN), which considers infrared target detection as the problem of image-to-image translation. Besides, Dai et al. [22] proposed a novel model-driven deep network named attention local contrast network (ALCNet), which implants traditional local contrast into the deep network. However, the performance of deep learning models rely on the training samples, which limits their application for complicated and volatile scenes.

HVS-based methods have caught attention and made great progresses for infrared small target detection in recent years. The predominant HVS-based methods can be subsumed under two broad categories, i.e., local contrast-based methods and directionality/gradient-based methods. Local contrast-based methods assume that there is great difference or discontinuity in radiation intensity between the target and neighborhood background on the infrared image. Inspired by biological vision mechanism, Chen et al. [26] first proposed local contrast measure (LCM) to achieve target enhancement and background suppression simultaneously. Then many researchers have proposed its variation, such as improved LCM (ILCM) [27], novel LCM (NLCM) [28], relative LCM (RLCM) [29], homogeneity-weighted LCM (HWLCM) [30], etc. However, these methods are sensitive to strong clutter and edges with high brightness. Directionality/gradient-based methods construct the final saliency map by measuring the gradient, directional derivative, or intensity difference between target and neighboring background (NB) in different directions. Liu et al. [31] transformed infrared image to infrared gradient vector field (GVF), and proposed multiscale flux density and gradient direction diversity to separate small targets from complex background clutter. Lu et al. [32] analyzed gradient properities using multidirectional derivative subbands. Besides, multiscale-patch-based contrast measure (MPCM) [33], absolute directional mean difference (ADMD) [34], local hypergraph dissimilarity measure [35], and multiscale trilayer LCM (TLLCM) [36] were proposed to estimate directional dissimilarity information of small target

neighborhood. These methods can preferably suppress strong edges. Moreover, some researchers integrated both local contrast and directionality/gradient properties, such as local intensity and gradient (LIG) [37] and fast adaptive masking and scaling with iterative segmentation (FAMSIS) [38], which takes more comprehensive information into account and deserves further research.

B. Motivation

Under marine environment, infrared images have low SCR and targets are small with low intensity because of the long imaging distance. Furthermore, for infrared maritime images with strong ocean waves, there are some waves with high local contrast; particularly, intensity of some strong ocean waves may be similar to or higher than that of small targets. Therefore, local contrast-based methods suffer from strong wave residuals, which causes high false alarm. From our observation, the target represents isotropic Gaussian characteristic in the neighborhood [37], [39], that is, the dissimilarity between the target and NB is similar in all directions. In contrast, the dissimilarity between ocean wave and corresponding neighborhood is different in horizontal and vertical directions [40].

In infrared maritime images with strong ocean waves, waves are serious interference that may be likely to cause severe false alarm detection. To date, researchers have observed the practicability of directionality or gradient property for infrared small target detection [41], [42], [43], [44], [45]. However, many infrared small target detection methods did not perform effectively on strong ocean wave suppression. Our crucial observation is that multidirectional dissimilarity is insufficient to separate infrared small targets from complicated background with strong ocean waves. In fact, infrared small targets and waves also differ in their gradient vector distribution. Based on this observation, utilizing both gradient vector distribution and directional dissimilarity simultaneously and appropriately will be helpful for improving the infrared small target detection performance under strong ocean wave background.

Therefore, we propose a novel framework, gradient vector field characterization (GVFC), for single-frame small target detection under strong ocean waves. The contributions of this article are mainly composed of the following three folds.

1) A GVFC framework is developed for infrared small maritime target detection under strong ocean waves by analyzing gradient vector feature, which employs multiple morphological attributes, i.e., distribution of pixel-level gradient direction angle and modulus, and gradient modulus local dissimilarity.
2) To improve the efficiency and effectiveness, we first introduce a synergistic homogeneity test (SHT) based on K-S test with ADSD to describe gradient direction angle distribution; besides, regression analysis of gradient modulus is considered to handle ocean wave clutter comprehensively. The proposed gradient vector distribution measure (GVDM) strictly achieves filtration of refined suspected targets (STs).
3) The gradient modulus horizontal local dissimilarity (GMHLD) based on horizontal direction information, which utilizes local directional dissimilarity property of gradient modulus, is promoted to drastically enhance target saliency and suppress residual wave clutter.

Our proposed GVFC method is implemented on various infrared maritime image datasets under adverse strong ocean waves. Experiments demonstrate that the proposed method achieves preferable and satisfactory detection performance with high accuracy, which outperforms the state-of-the-art methods.

The remainder of this article is structured as follows. In Section II, the theoretical basis of the proposed method is introduced in detail. Section III describes the experimental setup; besides, we analyze the key parameters of the proposed method and discuss result comparisons with baseline methods. Finally, we conclude this article in Section IV.

II. Proposed Method

The framework diagram of the proposed GVFC method is shown in Fig. 1. First, we construct GVF that contains gradient modulus and gradient direction. Then, we present GVDM that contains homogeneity analysis of gradient direction angle distribution and regression analysis of gradient modulus to achieve STs extraction. Due to the gradient modulus diversity of targets and waves in the horizontal direction, GMHLD is introduced to measure the saliency of STs, so as to enhance target and suppress residual ocean wave clutter. Finally, real targets can be segmented from background clutter using the adaptive threshold.

A. GVF Construction

For infrared maritime images with small targets captured under strong edge sea clutter conditions, the intensity of some clutter is similar to that of targets, which causes the difficulty for small infrared maritime target detection. However, the gradient of targets, background, and waves are distinct, which is beneficial to identifying real targets [40]. Considering that the raw infrared image can be regarded as a two-dimensional function $I(x, y)$, its GVF is a bivector that consists of partial derivatives in x and y directions and can be expressed as follows:

$$\mathrm{GVF} = \nabla I(x, y) = \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \tag{1}$$

where x and y directions correspond to horizontal and vertical directions, respectively. Thus, GVF can be calculated as follows:

$$\mathrm{GVF} = \begin{bmatrix} \dfrac{I(x+1, y) - I(x-1, y)}{2} \\ \dfrac{I(x, y+1) - I(x, y-1)}{2} \end{bmatrix}. \tag{2}$$

Take the real infrared image shown in Fig. 2 as an example, we utilize (2) to calculate the corresponding GVF for the target and wave labeled in raw image. In Fig. 2(b1) and (b2), the length and direction of ever arrow indicate the modulus and direction of the gradient vector ($\overrightarrow{\mathrm{GV}}$), respectively. Thus, to further intuitively illustrate the difference in GVF between target and wave, we decompose the gradient vector of every pixel into two components, namely direction angle and gradient modulus. And the decomposition process is shown in Fig. 3, which can

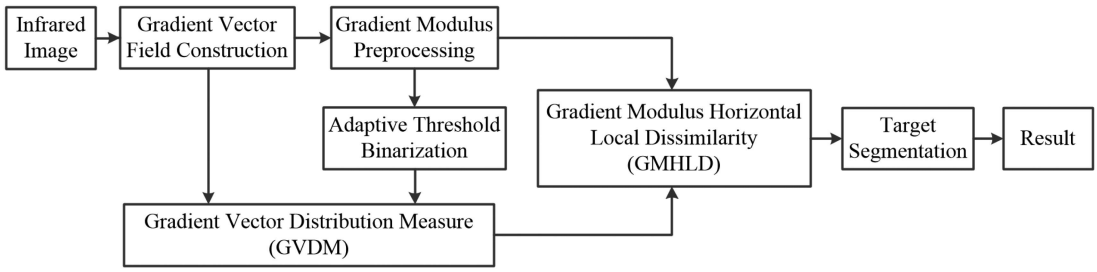

Fig. 1. Framework of the proposed method.

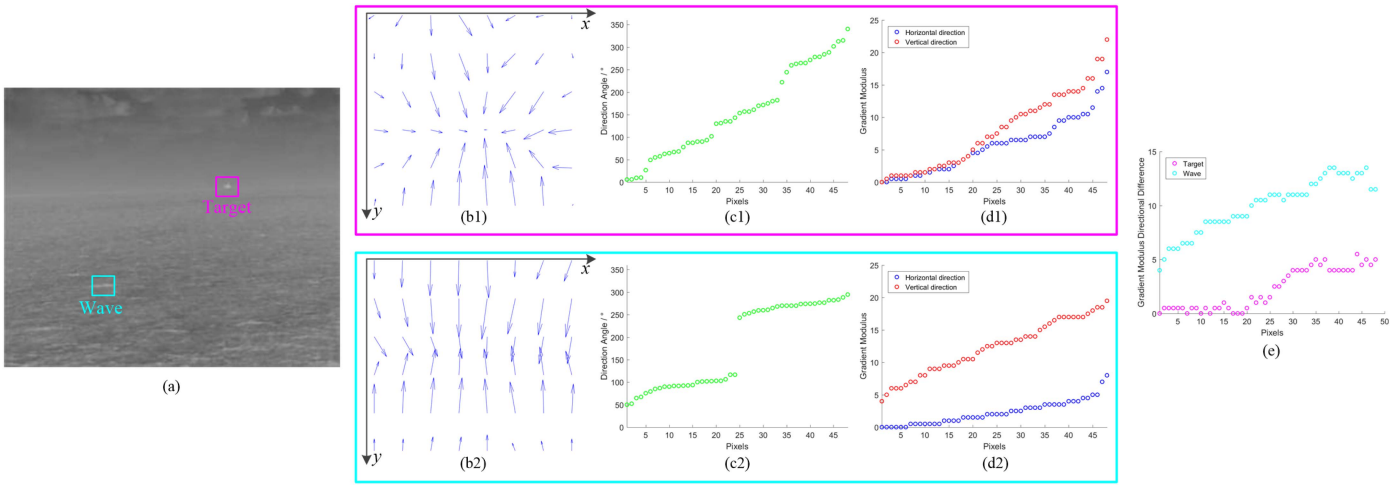

Fig. 2. Gradient characteristics for different types of regions. (a) Raw image. (b1) and (b2) GVF of target and wave, respectively. (c1) and (c2) Direction angle of target and wave, respectively. (d1) and (d2) Gradient modulus in two directions (horizontal and vertical) of target and wave, respectively. (e) Gradient modulus difference in horizontal and vertical directions.

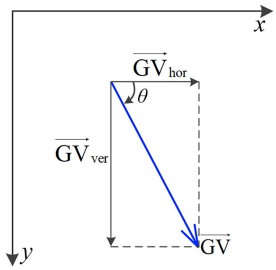

Fig. 3. Decomposition for gradient vector. $\overrightarrow{\mathrm{GV}}$ represents a gradient vector. θ is the direction angle of the $\overrightarrow{\mathrm{GV}}$. $\overrightarrow{\mathrm{GV}}_{\mathrm{hor}}$ and $\overrightarrow{\mathrm{GV}}_{\mathrm{ver}}$ denote subgradient vector of the $\overrightarrow{\mathrm{GV}}$ in horizontal and vertical directions, respectively.

also be expressed as follows:

$$\begin{cases} |\overrightarrow{\mathrm{GV}}| = \sqrt{|\overrightarrow{\mathrm{GV}}_{\mathrm{hor}}|^2 + |\overrightarrow{\mathrm{GV}}_{\mathrm{ver}}|^2} \\ \theta = \tan^{-1} \dfrac{\overrightarrow{\mathrm{GV}}_{\mathrm{ver}}}{\overrightarrow{\mathrm{GV}}_{\mathrm{hor}}} \end{cases} \tag{3}$$

where θ is the angle between $\overrightarrow{\mathrm{GV}}$ and the clockwise direction of the x-axis, so θ ranges from 0° to 360°. $\overrightarrow{\mathrm{GV}}_{\mathrm{hor}}$ and $\overrightarrow{\mathrm{GV}}_{\mathrm{ver}}$ can be calculated by (2). Form Fig. 2(c1) and (c2), we can see that direction angles of target are evenly distributed in all directions, however, direction angles of wave mainly concentrate on the vertical direction (i.e., 90° and 270°). Furthermore, it can be seen from Fig. 2(d1) and (d2) that gradient moduli of target pixels in horizontal and vertical directions are approximate, but there are significant diversities of module values for wave in two directions, which can also be seen from Fig. 2(e). Module value differences of wave in two directions are generally greater than that of target. Moreover, horizontal gradient modulus of wave is obviously small. Given the diverse gradient vector characteristics between target and wave, we present GVDM and GMHLD to achieve real target enhancement and background clutter suppression.

B. ST and Neighborhood Estimation

To improve the computing efficiency of the target detection algorithm, we first extract STs coarsely and adaptively estimate corresponding neighborhoods for subsequent processing. Initially, we normalize the gradient modulus of the infrared image to obtain the gradient modulus map (GM) as follows:

$$\mathrm{GM} = \mathrm{norm}\left(|\overrightarrow{\mathrm{GV}}|\right) \tag{4}$$

where $\mathrm{norm}(\cdot)$ is the normalization operator. Then, a simple adaptive threshold is used to binarize GM and extract STs as

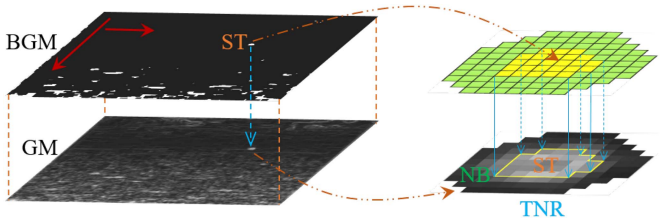

Fig. 4. Illustration of matching process and relationship between ST, NB, and TNR.

follows:

$$T_b = \text{mean}(\text{GM}) + k \cdot \text{std}(\text{GM}) \tag{5}$$

where $\text{mean}(\cdot)$ and $\text{std}(\cdot)$ denote the mean and standard deviation of GM. k is a constant and set to 3 empirically, which is discussed in Section III-B. The binary GM is denoted as BGM.

After the aforementioned process, we employ (6) to capture the target neighboring region (TNR) and corresponding NB for ST

$$\begin{cases} \text{TNR} = \text{ST} \oplus \mathcal{D}_r = \{(x,y) | [ST \cap (\mathcal{D}_r)_{x,y}] \neq \varnothing\} \\ \text{NB} = \text{TNR} - \text{ST} = \{(x,y) | (x,y) \in \text{TNR}, (x,y) \notin \text{ST}\} \end{cases} \tag{6}$$

where $\oplus$ is dilation operator, $\mathcal{D}_r$ is a disk-shaped morphological structure element, and its radius is r pixels. r is a positive integer and set to 4 empirically in this article, see more discussion in Section III-B.

Then, STs in BGM are utilized to guide corresponding GM. To be more specific, we traverse STs in BGM from top to bottom and left to right, and match positions and shapes of STs and corresponding NBs to GM. To make it more intuitive, the matching and neighborhood acquisition process is illustrated in Fig. 4. Thereinto, the small yellow squares correspond to pixels of the ST, small green squares represent pixels of the corresponding NB, and they collectively constitute TNR. After that, we employ the information of STs with NBs and TNRs in GM to measure gradient vector distribution and calculate horizontal local dissimilarity.

C. Gradient Vector Distribution Measure

Gradient vector consists of direction angle and gradient modulus, and they represent the high-frequency characteristics of images synergistically. According to the description for gradient vector in Section II-A, we comprehensively analyze the direction angle and modulus of the gradient in detail to further filtrate STs strictly.

1) Direction Angle Analysis: According to the gradient vector direction angle characteristics of target and heavy wave analyzed in Section II-A, direction angle distribution of target is homogeneity and ranging from 0° to 360° while direction angles of heavy wave diverge in two ranges. Therefore, an SHT based on Kolmogorov–Smirnov (K-S) test and absolute difference standard deviation (ADSD) is proposed to analyze direction angle distribution in TNR of each ST.

K-S test is a test of goodness-of-fit based on cumulative distribution function (CDF) and usually used to test whether a observed sample obeys the hypothetical theoretical distribution function [46]. For an ST, assume that there are N pixels in corresponding TNR, and the theoretical sample is expressed by $A_t = \{360 * i/N, i = 1, 2, \ldots, N\}$. We represent observed sample in ascending sort as A_o. Then, the CDF of theoretical sample and observed sample can be denoted as $F_t(x)$ and $F_o(x)$, respectively

$$F_t(x) = \frac{\#\left\{A_t^i \leqslant x, i = 1, 2, \ldots, N\right\}}{N} \tag{7}$$

$$F_o(x) = \frac{\#\left\{A_o^i \leqslant x, i = 1, 2, \ldots, N\right\}}{N} \tag{8}$$

where $\#\{\cdot\}$ represents the number of individuals that satisfy corresponding criteria in the sample, and $x \in \{A_t, A_o\}$. Thus, K-S statistic is the maximal absolute deviation between $F_t(x)$ and $F_o(x)$ as follows:

$$D_{\text{KS}} = \max\left(\left|F_t(x) - F_o(x)\right|\right). \tag{9}$$

The critical value can be calculated by

$$D_{\text{crit},\alpha} = \sqrt{-\frac{1}{2.6}\ln\frac{\alpha}{2}} \cdot \sqrt{\frac{1}{N}} \tag{10}$$

where α is the significance level and related to the confidence coefficient, which can be represented as 1-α. In this article, at a given confidence coefficient of 95% (i.e., 1-α = 0.95). Then, α = 0.05, and we can get the corresponding critical value $D_{\text{crit},0.05}$ by (10). According to K-S test theory, if $D_{\text{KS}} < D_{\text{crit},0.05}$, then A_o obeys distribution of A_t. That is, direction angle distribution of the ST is homogeneity.

ADSD is also introduced to improve the robustness of the SHT. ADSD is used to measure the degree of dispersion for the absolute difference between A_o and A_t, which can be calculated by the following formulas:

$$\text{AD}_i = \left|A_o^i - A_t^i\right|, i = 1, 2, \ldots, N \tag{11}$$

$$\text{ADSD} = \sqrt{\frac{1}{N-1}\sum\nolimits_{i=1}^{N}\left(\text{AD}_i - \overline{\text{AD}}\right)^2}. \tag{12}$$

Obviously, the smaller the value of ADSD, the more homogeneous the direction angle distribution. In this article, if ADSD $\leqslant$16, direction angle distribution of the ST is homogeneity empirically.

Therefore, D_{KS} is combined with ADSD to make the SHT. When $D_{\text{KS}} < D_{\text{crit},0.05}$ or ADSD $\leqslant$16 is satisfied, the corresponding ST is more likely to be a target. The specific decision rule can be described as

$$\text{SHT}_{\text{ST}} = \begin{cases} 1, & \text{if } D_{\text{KS}} < D_{\text{crit},0.05} \vee \text{ADSD} \leqslant 16 \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

2) Gradient Modulus Analysis: Linear regression classification (LRC) model can effectively deal with the recognition problem, such as face recognition [47]. Fig. 5 shows the scatter diagrams of the gradient modulus distribution for representative targets and waves. By observing the changing trend of gradient modulus, we find that it conforms to the exponential change law, which demonstrates that there is a log-linear relationship between the gradient moduli and pixels. Thus, we introduce

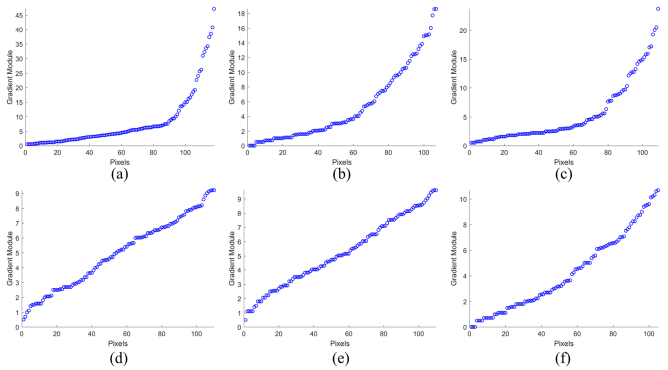

Fig. 5. Examples of gradient modulus distribution for pixels in corresponding TNR. (a)–(c) Modulus distribution of targets. (d)–(f) Modulus distribution of waves.

log-linear regression to infrared small target detection, assuming that the regression function of gradient modulus admits the following formula:

$$\ln \mathbf{y}^T = \beta_0 + \beta_1 \mathbf{x}^T \tag{14}$$

where $\mathbf{x} = \{1, 2, \ldots, N\}$ is the pixel sequence and N is pixel number in TNR. $\mathbf{y} = \{y_1, y_2, \ldots, y_N\}$ represents the gradient modulus sequence arranged in ascending order. β_0 and β_1 are coefficients that can be estimated using the least-squares estimation method [48], [49], [50]. Then, (14) can be expressed as $\hat{y} = A \cdot B^x$, where $A = e^{\beta_0}$ and $B = e^{\beta_1}$. We now calculate distance between original and predicted gradient modulus

$$d(\beta_0, \beta_1) = \|\mathbf{y}^T - A \cdot B^{\mathbf{x}^T}\|_2. \tag{15}$$

Then, the optimal estimation of β_0 and β_1 can be effectively handled by minimizing $d(\beta_0, \beta_1)$, i.e,

$$\arg \min_{\beta_0, \beta_1} \|\mathbf{y}^T - A \cdot B^{\mathbf{x}^T}\|_2, \quad \text{s.t. } A = e^{\beta_0}, B = e^{\beta_1}. \tag{16}$$

Therefore, the specific LRC rule is

$$\text{LRC}_{\text{ST}} = \begin{cases} 1, & \text{if } \beta_0 < \beta_{0,\text{crit}} \wedge \beta_1 > \beta_{1,\text{crit}} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

where $\beta_{0,\text{crit}}$ and $\beta_{1,\text{crit}}$ are numerical constants and determined experimentally. The adaptive critical values of $\beta_{0,\text{crit}}$ and $\beta_{1,\text{crit}}$ are ln2.3 and ln1.02, respectively.

Finally, the gradient direction angle distribution of ST in corresponding TNR is combined with distribution of gradient modulus to make a comprehensive final decision as follows:

$$\text{ST is } \begin{cases} \text{preserved, if } \text{SHT}_{\text{ST}} = 1 \wedge \text{LRC}_{\text{ST}} = 1 \\ \text{removed, otherwise.} \end{cases} \tag{18}$$

Consequently, the preservation or removal of ST is concluded to achieve GVDM. The procedure of GVDM is described in Algorithm 1. Furthermore, we introduce GMHLD to enhance saliency of real targets and suppress residual clutter.

D. Gradient Modulus Horizontal Local Dissimilarity

Considering the similarity of target in horizontal and vertical directions and diversity of wave in two directions for gradient modulus, we present GMHLD to enhance target and suppress residual clutter simultaneously.

Algorithm 1: Gradient Vector Distribution Measure.

Input: GVF, BGM, structure element radius r.
Output: ST extraction result GVDM_{ST}.

```
Calculate the number of STs found in BGM (N_st);
for k = 1 : N_st do
    Generate the TNR of ST k by (6);
    Match ST and TNR to GVF and obtain the gradient
    modulus and direction angle of each pixel;
    Sort the direction angles of all pixels of the ST in
    ascending order and denoted as A_o;
    Construct theoretical sample
    A_k = {360 * i/pixn, i = 1, 2, ..., pixn}, where
    pixn is the number of pixels of ST;
    SHT by (7)–(13);
    Sort the gradient modulus of all pixels of the TNR in
    ascending order and denoted as y;
    Analyze gradient modulus based on LRC model by
    (15)–(17);
    Make the decision by (18);
end for
Obtain the GVDM map.
```

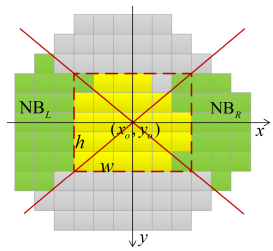

Fig. 6. Example of NB division. Small yellow squares correspond to pixels of RST; small green and gray squares represent pixels of NB. (x_o, y_o) is the center of RST. $2w$ and $2h$ are the width and height of RST, respectively.

GMHLD measures the dissimilarity between refined ST (RST) and horizontal NB, where the RST is calculated according to STs in GVDM as follows:

$$\text{RST} = [\![\text{BGM} \bullet \mathcal{D}_3, \text{GVDM}_{\text{ST}}]\!]_{\text{RG}} \tag{19}$$

where $\bullet$ denotes morphological closing operation, $\mathcal{D}_3$ denotes disk structure element with a radius of 3 pixels, and $[\![\cdot]\!]_{\text{RG}}$ is region growing operation.

In order to achieve GMHLD, NB is divided into four patches according to the shape of RST, and we select two horizontal patches for GMHLD calculation, which is shown in Fig. 6. We build a two-dimensional coordinate system with the origin at the center of RST (i.e., (x_o, y_o)). Thus, the two horizontal patches of NB can be denoted as follows:

$$\text{NB}_L = \left\{ (x, y) \mid (x, y) \in \text{NB}, x \leqslant x_o, \frac{y - y_o}{x - x_o} \leqslant \left| \frac{h}{w} \right| \right\} \tag{20}$$

Algorithm 2: GMHLD and Target Segmentation.

Input: GM, BGM, structure element radius r, GVDM map.
Output: Final detection result *Res*.

```
Generate RST according to (19);
Calculate the number of RSTs (N_rst)
for k = 1 : N_rst do
    Generate the NB of RST k by (6);
    Generate the NB_L and NB_R by (20) and (21);
    Match RST, NB_L, and NB_R to GM followed by
    calculating their mean of gradient modulus,
    respectively;
    Calculate the GMHLD_k by (22);
    Record the salient value of the RST as follows
    SV_RST(k) = GMHLD_k;
end for
Calculate the segmentation threshold T by (24).
Separate the GMHLD map by (25).
Acquire the final detection result Res and target
position.
```

$$\mathrm{NB}_R = \left\{ (x,y) \mid (x,y) \in \mathrm{NB}, x \geqslant x_o, \frac{y-y_o}{x-x_o} \leqslant \left| \frac{h}{w} \right| \right\} \tag{21}$$

where (x, y) represents coordinates of pixels in NB. GMHLD can be calculated by

$$\begin{aligned} \mathrm{GMHLD} = &\left(\overline{\mathrm{GM}}_{\mathrm{RST}} - \overline{\mathrm{GM}}_{\mathrm{NB_L}}\right) \cdot \left(\overline{\mathrm{GM}}_{\mathrm{RST}} - \overline{\mathrm{GM}}_{\mathrm{NB_R}}\right) \cdot \\ &\varepsilon\left[\overline{\mathrm{GM}}_{\mathrm{RST}} - \max(\mathrm{GM}_{\mathrm{NB}})\right] \end{aligned} \tag{22}$$

where $\overline{\mathrm{GM}}_*$ denotes mean of gradient modulus of the corresponding region. $\varepsilon[\overline{\mathrm{GM}}_{\mathrm{RST}} - \max(\mathrm{GM}_{\mathrm{NB}})]$ is a step function which can be expressed as follows:

$$\begin{aligned} &\varepsilon\left[\overline{\mathrm{GM}}_{\mathrm{RST}} - \max(\mathrm{GM}_{\mathrm{NB}})\right] \\ &\quad = \begin{cases} 1, \overline{\mathrm{GM}}_{\mathrm{RST}} > \max(\mathrm{GM}_{\mathrm{NB}}) \\ 0, \overline{\mathrm{GM}}_{\mathrm{RST}} \leq \max(\mathrm{GM}_{\mathrm{NB}}). \end{cases} \end{aligned} \tag{23}$$

Finally, the targets can be separated from residual clutter by an adaptive segmentation threshold. We count the salient value of each RST in enhanced result and denote them as SV_RST; then, the threshold can be calculated as follows:

$$T = \lambda \cdot \mathrm{SV\_RST}_{\max} + (1-\lambda) \cdot \mathrm{SV\_RST}_{\min} \tag{24}$$

where $\mathrm{SV\_RST}_{\max}$ and $\mathrm{SV\_RST}_{\min}$ are the maximum and minimum of SV_RST. λ is a constant ranging from 0 to 1. Finally, the detection result is obtained as follows:

$$\mathrm{Res}(x,y) = \begin{cases} 1, \mathrm{GMHLD}(x,y) \geq \mathrm{T} \\ 0, \text{otherwise} \end{cases} \tag{25}$$

where (x, y) represents the infrared image pixel point. The procedure of GMHLD and target segmentation is summarized in Algorithm 2.

III. Experimental Results and Discussion

In this section, we utilize eight infrared maritime image datasets with strong ocean wave clutter to demonstrate the small target detection performance of the proposed method. Initially, the information of datasets, baseline methods, and evaluation metrics are introduced in this article. Then, we analyze the effects of key parameters by performing experiments on the representative datasets. And ablation experiments are designed to demonstrate the robustness of our method. Finally, qualitatively and quantitatively, we discuss and compare the detection performance of the considered methods. All experiments are conducted on a computer with 16.0-GB RAM and Intel i7-9700 CPU with 3.0 GHz processor, and the code is implemented using MATLAB(X64) R2016a.

A. Experimental Setup

1) Datasets: In the experiment, eight infrared maritime image datasets (1390 images) with different types of targets (more than 2100 targets) and complex background are used to demonstrate the performance of our method. The detailed information of eight datasets (i.e., Datasets 1–8) is listed in Table I. The resolution of Datasets 1 to 7 is 284 × 236, and the resolution of Dataset 8 is 640 × 512. Dataset 1 is disturbed by heavy waves and islands simultaneously, and two small targets have similar contrast. In Datasets 2 and 5, there is a small ship and sea–sky background with heavy wave clutter and cloudy sky. And Dataset 5 has small islands in the distance. Dataset 3 presents a very strong ocean wave background and the target is submerged in ocean wave clutter. In Dataset 4, there are three small targets with obviously different intensity, and the background is sea–sky background with strong ocean waves. Datasets 6 and 7 are synthetic images. The two image datasets are synthesized by real targets from Dataset 3 and real infrared maritime background with heavy waves or/and thick cloud, and the synthesis strategy is similar to that in [31]. Dataset 8 contains two small dim targets and there are islands with high radiation intensity in the background.

2) Baseline Methods: In order to illustrate the effectiveness and robustness of the proposed method, we perform the comparative experiment with the classical and acclaimed methods, namely LIG, LCM, MPCM, NLCD, PSTNN, and the recent, TLLCM, ADMD, FAMSIS, and MSLSTIPT. LIG and FAMSIS consider the LIG simultaneously to detect infrared targets. LCM, NLCD, and TLLCM are methods based on local contrast. MPCM and ADMD employ directional information and multiscale difference to enhance target saliency and suppress background clutter. PSTNN transforms the infrared small target detection task into nonconvex tensor robust principal component analysis model, and it achieves target detection according to the low-rank and sparse property of image. MSLSTIPT introduces multisubspace learning strategy into IPT model, which deeply excavates the spatial–temporal information to improve the robustness on complicated scenes. The parameters of baseline methods are set to their recommended values. Table II shows the parameter settings of considered methods.

TABLE I
DETAILED INFORMATION OF THE EIGHT INFRARED IMAGE DATASETS

Dataset	Frame	Target number	Target size	Target description	Background description
1	150	300	5×2 ∼ 10×3	Two targets with similar contrast.	Sea–sky line, heavy waves, and island with high intensity.
2	200	200	4×3 ∼ 10×5	One small target.	Cloudy sky and heavy waves.
3	160	160	2×2 ∼ 5×3	A small target with low and varying intensity.	Heavy waves with strong edges obscure the target.
4	150	450	2×2 ∼ 7×4	Three targets with obviously different local contrast.	Sea–sky background with heavy wave clutter.
5	200	200	4×2 ∼ 5×4	One small dim target submerged in complicated background.	Sea–sky line, thick cloud, small islands, and heavy waves.
6	130	130	2×2 ∼ 5×3	A small target with low and varying intensity.	Heavy waves with strong edges obscure the target.
7	130	130	2×2 ∼ 5×3	A small target with low and varying intensity.	Sea–sky background, thick cloud, heavy waves.
8	270	540	4×5 ∼ 6×7	Two small dim targets.	Islands with high radiation intensity in the background.

TABLE II
PARAMETER SETTINGS OF CONSIDERED METHODS

Method	Abbreviation	Parameter setting
Local Intensity and Gradient [37]	LIG	local window size: 11×11, k = 0.2
Local Contrast Measure [26]	LCM	cell size: 3×3, 5×5, 7×7, 9×9, k = 3
Multi-scale Path-based Contrast Measure [33]	MPCM	L = 3, patch size: 3×3, 5×5, 7×7
Novel Local Contrast Descriptor [52]	NLCD	K = 4, p = 6, β = 200, window size: 11×11
Partial Sum of Tensor Nuclear Norm [18]	PSTNN	patch size: 40×40, sliding step:40, $\varepsilon = 10^{-7}$, $\lambda = 0.6/\sqrt{n_3 \max(n_1, n_2)}$
Tri-Layer Local Contrast Measure [36]	TLLCM	L = 3, scale size: 5×5, 7×7, 9×9
Absolute Directional Mean Difference [34]	ADMD	scale size: 3×3, 5×5, 7×7, 9×9
Fast Adaptive Masking and Scaling With Iterative Segmentation [38]	FAMSIS	scale factor s = 3, p = 0.2%
Multiple Subspace Learning and Spatial-Temporal Infrared Patch-tensor [8]	MSLSTIPT	patch size: 30×30, L = 6, p = 0.8, $\lambda = 1/\sqrt{n_3 \max(n_1, n_2)}$
Gradient Vector Field Characterization (Proposed)	GVFC	k = 3, r = 4

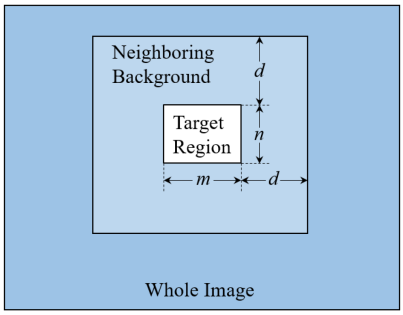

Fig. 7. Illustration for target and NB regions of a small target.

3) Evaluation Metrics: In order to validate the performance of various methods, we introduce local contrast gain (LCG) and background suppression factor (BSF) in this article [17], [51], [52]. LCG is used to measure the target enhancement quality and defined as follows:

$$\mathrm{LCG} = \frac{\left|\overline{\mathrm{TP}} - \overline{\mathrm{NBP}}\right|_{\mathrm{enh}}}{\left|\overline{\mathrm{TP}} - \overline{\mathrm{NBP}}\right|_{\mathrm{raw}}} \tag{26}$$

where enh and raw are short for enhanced result and raw image, respectively. $\overline{\mathrm{TP}}$ and $\overline{\mathrm{NBP}}$ represent the mean gray values of the target pixels and NB pixels, respectively. The LCG is calculated in a local region, as shown in Fig. 7. Assume that the small target size is $m \times n$, we set the range of corresponding NB is $d = \max(m, n)$. Note that LCG is the average value of all the actual targets in the image.

BSF is usually used to evaluate background suppression ability, and it can be calculated by

$$\mathrm{BSF} = \frac{\sigma_{\mathrm{BP,raw}}}{\sigma_{\mathrm{BP,enh}}} \tag{27}$$

where $\sigma_{\mathrm{BP,raw}}$ and $\sigma_{\mathrm{BP,enh}}$ denote the standard deviations of background pixels for raw infrared image and enhanced result, respectively. It should be noted that we only consider the background pixels other than the target region pixels when calculating BSF, and the enhanced results for all the methods stand for the target enhancement and background suppression result, which is prior to final threshold segmentation.

Besides, the receiver operating characteristic (ROC) curve [53] is usually used to comprehensively measure the detection accuracy for target detection method. ROC reflects variation tendency and congruent relationship between detection probability P_d and false-alarm rate P_f. The vertical and horizontal axes of ROC are P_d and P_f, respectively. Their definition can be expressed as follows:

$$P_d = \frac{\text{number of detected actual targets}}{\text{number of total actual targets}} \tag{28}$$

$$P_f = \frac{\text{number of false alarm targets}}{\text{number of total detected targets}}. \tag{29}$$

According to the ROC curve, the value of area under the curve (AUC) can be obtained. And the closer the ROC curve is to the upper left corner, the greater the value of AUC is, which indicates more satisfactory detection performance and accuracy for the corresponding method.

B. Parameters Analysis

To achieve preferable detection accuracy and robustness, two key parameters of the proposed method are discussed in this section. We perform experiments on five real datasets (i.e., Datasets 1–5) to analyze binarization threshold parameter k in (5) and structure element radius r in (6) by controlling the variate technique. And the ROC and AUC curves are conjunctively used to determine the optimal parameters.

1) Binarization Threshold Parameter k: Parameter k controls the number, positions, shapes, and sizes of STs, which also partly affects TNR and NB regions. Therefore, the value of k will affect the robustness of the proposed method. We change

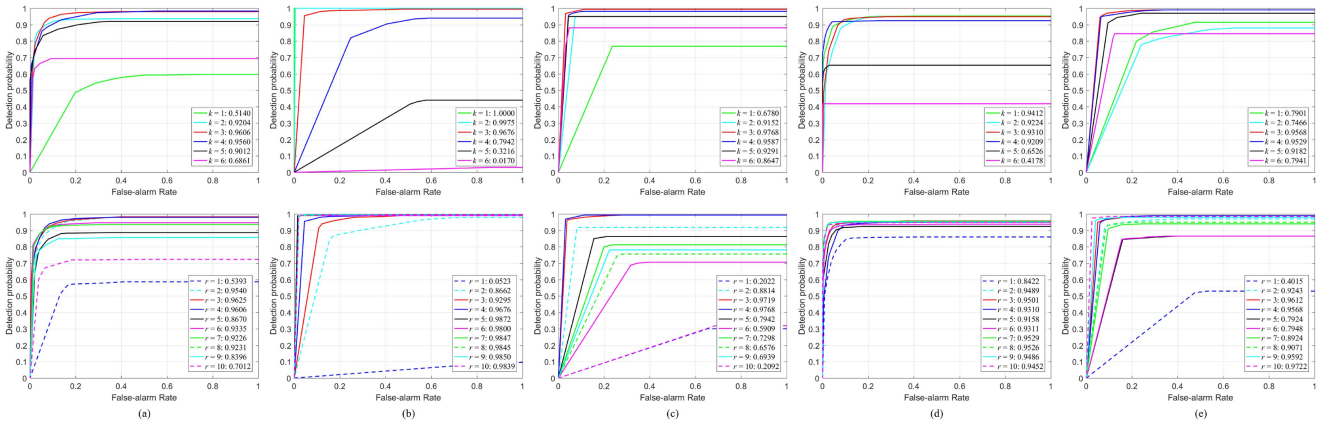

Fig. 8. ROC curves of the proposed methods with different values of parameters k and r on five datasets. (a) Dataset 1. (b) Dataset 2. (c) Dataset 3. (d) Dataset 4. (e) Dataset 5.

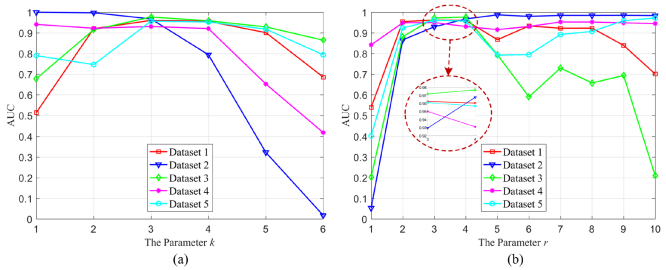

Fig. 9. AUC curves of the proposed methods with different values of parameters k and r on five datasets.

k from 1 to 6 on five datasets, and the interval is 1. In the experiment, structure element radius r is set to 4. The ROC curves are shown in the first row of Fig. 8. And the corresponding AUC values are listed in the legend of each graph. In order to intuitively and clearly observe the effect of k on performance, we show the AUC curves in Fig. 9(a). It can be seen that if k is set to 3, the detection performance is most satisfactory and robust.

2) Structure Element Radius r: Parameter r conducts the shapes and sizes of TNR and NB regions directly. In order to ensure the detection accuracy, we must choose it carefully. If r is too small, the information in NB cannot represent NB characteristics, because the target neighborhood often behaves as a Gaussian, which causes the poor performance. If r is too large, TNR and NB may contain other strong ocean wave edges and cause miss detection. We change r on five datasets from 1 to 10 with the interval of 1. In the experiment, the binarization threshold parameter k is set to 3. The ROC and AUC curves are shown in the second row of Figs. 8 and 9(b), respectively. One can see that higher AUC values are achieved using larger r values for Datasets 2, 4, and 5. However, lower AUC values are obtained with larger r values for Datasets 1 and 3. When r is 3 or 4, our method can achieve higher AUC values for all the datasets robustly. Consequently, taking both detection accuracy and robustness into account, r is set to be 4 for the experiments below.

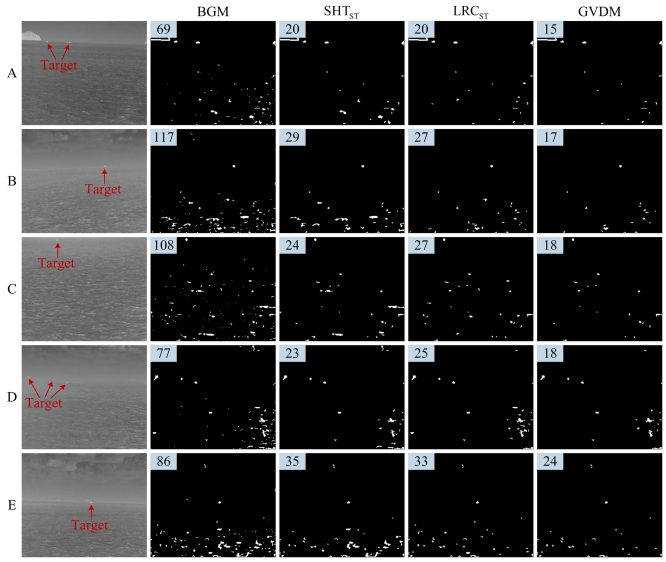

Fig. 10. Numbers of STs in different parts. Column 1 shows the raw representative images. Columns 2 through 5 show the results for BGM, SHT_{ST}, LRC_{ST}, and GVDM. The number of ST is indicated in the upper left of each image.

C. Ablation Experiments

In this section, we design two experiments to analyze the contribution of each step of our method by ablation study.

First, to demonstrate the performance of GVDM intuitively, we count the numbers of STs in BGM, SHT_{ST}, LRC_{ST}, and GVDM for different infrared maritime images with a complicated background as shown in Fig. 10. One can see that SHT and LRC can reduce the number of false-alarm targets to some extent, however, the number of STs in GVDM is the lowest together with well preservation of all the real targets, which indicates that GVDM is conducive to filtrating STs more accurately.

In the second experiment, in order to demonstrate the effectiveness of GVDM and GMHLD, we compare the results of the proposed method with different processing models as shown in Fig. 11. Thereinto, (a1)–(e1) show the processing

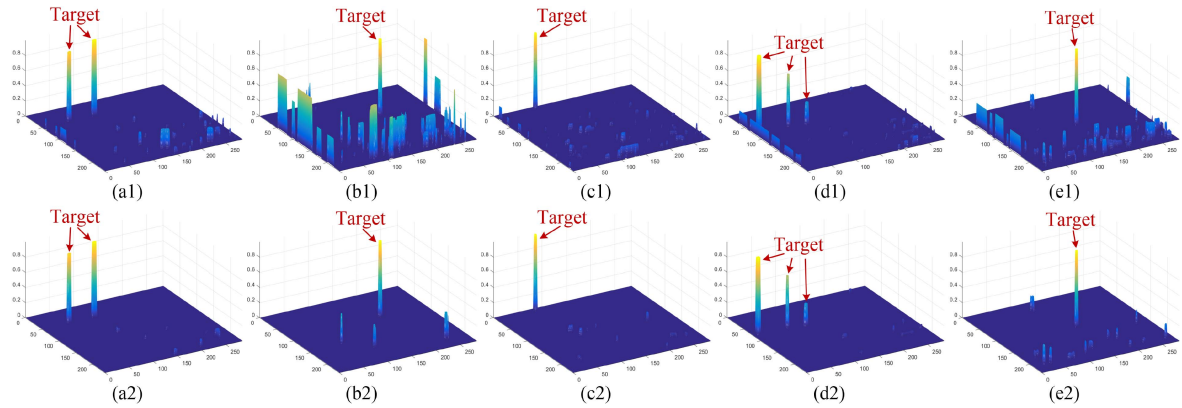

Fig. 11. Results of GMHLD for infrared maritime images A-E in Fig. 10. (a1)–(e1) Enhanced results of the proposed method without GVDM. (a2)–(e2) Enhanced results of the proposed method.

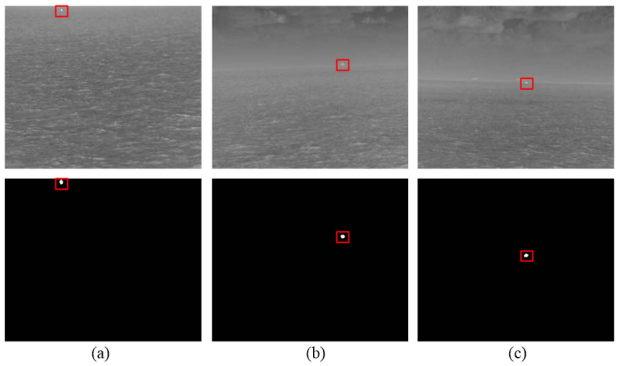

Fig. 12. Single target scenes and corresponding results. Rows 1 and 2 show the original images and detection results, respectively. (a) Scene 1. (b) Scene 2. (c) Scene 3.

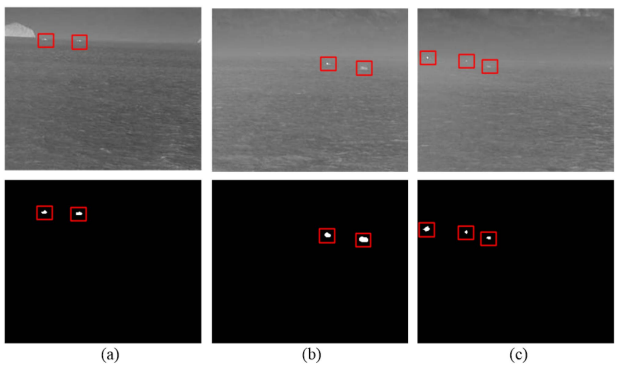

Fig. 13. Multiple target scenes and corresponding results. Rows 1 and 2 show the original images and detection results, respectively. (a) Scene 1. (b) Scene 2. (c) Scene 3.

results of the proposed method without GVDM, that is, we only implement GMHLD on GVF. It can be seen that GMHLD performs well in target enhancement while unsatisfactorily in background clutter suppression. Particularly, some clutter is salient, which even resembles real target. Fig. 11(a2)–(e2) shows the enhanced results of the whole proposed method that contains GVDM followed by GMHLD. One can see that residual clutter is

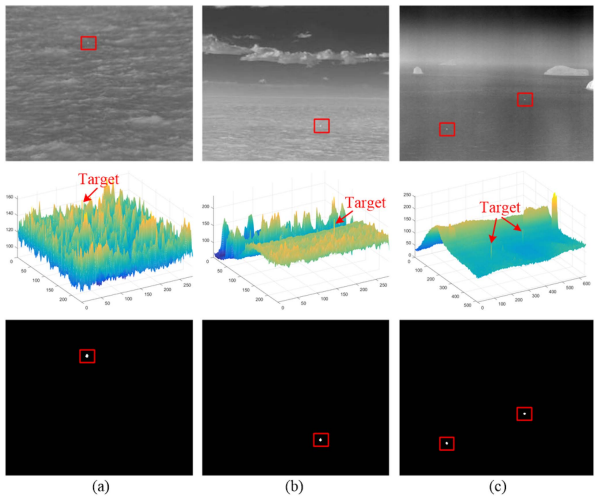

Fig. 14. Complicated scenes and corresponding results. Rows 1, 2, and 3 show the original images, corresponding 3-D gray distributions, and detection results, respectively. (a) Scene 1. (b) Scene 2. (c) Scene 3.

suppressed distinctly and robustly, and real targets are enhanced simultaneously. The comparative results indicate the effectiveness of the integration of GVDM and GMHLD.

D. Validity of the Proposed GVFC Method

In this section, we demonstrate the robustness of the GVFC method combined with the characteristics of images in various scenes.

1) Robustness to Single Target Scene: First, we test the proposed GVFC method on real single target scenes. The representative images and detection results are shown in Fig. 12, in which the real targets are labeled in red rectangular boxes. The results in Fig. 12 validate that the targets are detected successfully and strong ocean wave clutters are well suppressed.

2) Robustness to Multiple Target Scene: Actually, there may be multiple targets in the same scene. Therefore, we demonstrate the validity of the proposed GVFC method on real maritime scenes with diverse targets. Fig. 13 shows the representative images and corresponding results. Thereinto, scenes 1 and 2

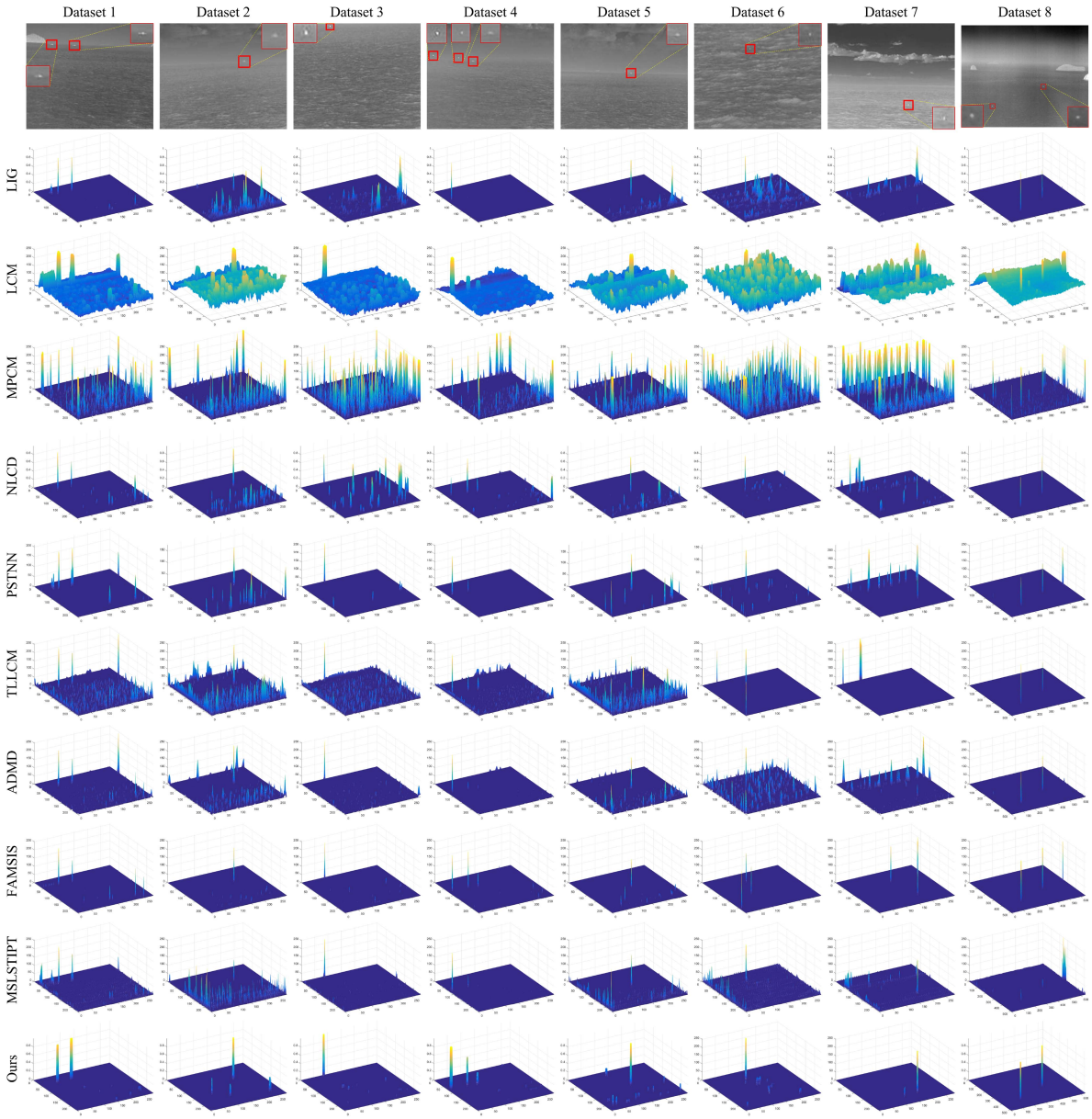

Fig. 15. Enhanced results processed by the considered methods for Datasets 1 to 8.

both contain two targets, the intensity and contrast of three targets in scene 3 are different. The detection results in Fig. 13 demonstrate that the proposed method can effectively enhance targets' saliency and suppress the interferences of island and ocean wave clutters.

3) Robustness to Dim Target and Complex Background Scene: In the marine scene, strong wave edge, cloud, and island are the key factors that affect the detection performance. Thus, we test the robustness of the proposed GVFC method in complicated scenes. Fig. 14 shows the representative images and results. It is worth noting that scenes 1 and 2 are synthetic images according to the real infrared maritime background and target by a similar method introduced in [31]. In scene 1, the small dim target is submerged in heavy waves. In scene 2, there are thick cloud and heavy waves that cause great difficulties to target detection. In scene 3, the intensity of target is low, by contrast, the islands in background have the high radiation intensity, which also makes challenges to small dim target detection. We use 3-D gray distributions to further display the characteristics of each image intuitively, which is shown in the second row of Fig. 14. It can be seen from the third row of Fig. 14 that the proposed method can perform satisfactory detection results, which demonstrates the robustness of the proposed GVFC method to dim target and diverse complicated scenes.

E. Comparison to State-of-the-Art Methods

1) Qualitative Analysis: Fig. 15 shows the raw representative infrared maritime images and 3-D gray distribution of enhanced results on Datasets 1 to 8. For Dataset 1, we can see that

TABLE III
QUANTITATIVE COMPARISON OF 10 METHODS ON THE EIGHT DATASETS

Evaluation metrics	Datasets	LIG	LCM	MPCM	NLCD	PSTNN	TLLCM	ADMD	FAMSIS	MSLSTIPT	Ours
LCG	1	1.2159	***9.5127***	2.2708	0.8733	1.6907	1.8203	1.4754	0.8630	1.3238	**11.8835**
	2	1.6255	***5.6564***	4.6377	3.3159	2.4264	4.6570	3.7648	1.4750	1.0096	**11.9630**
	3	0.0000	***4.2301***	2.3835	0.8049	2.2621	2.1232	1.8171	1.1523	1.9821	**13.5030**
	4	0.9111	***20.9502***	2.6402	2.4701	1.8173	2.2010	1.5198	1.9594	1.8924	**23.4689**
	5	3.4462	***8.6345***	4.3983	2.1351	2.6223	3.0761	2.3453	1.3467	2.4483	**16.7522**
	6	3.9644	***15.0673***	2.9618	2.4217	2.6494	2.4838	4.3008	1.3195	3.7083	**17.1539**
	7	0.7033	***5.4710***	1.9163	2.4690	1.6385	0.0000	1.0287	0.9539	2.8139	**12.4011**
	8	1.5040	***8.5572***	3.7008	1.2352	1.6911	0.6477	1.8028	2.7069	0.3844	**11.4339**
BSF	1	***13.6400***	0.7918	1.1432	6.4465	6.1186	2.5874	4.7417	10.8869	5.6031	**18.3145**
	2	2.0129	0.3884	0.8465	2.9448	4.8302	1.5543	2.4971	**18.0055**	2.5890	***5.6669***
	3	2.6480	0.8656	0.6346	2.2544	14.2172	2.9688	8.4756	**22.8986**	***16.9053***	15.8642
	4	***234.2258***	1.1152	1.0179	5.5540	**Inf**	4.4503	18.2031	15.2248	29.7323	7.1926
	5	4.2981	0.4946	0.9891	4.3555	5.7067	1.5956	3.6365	**13.8033**	3.8189	***6.2848***
	6	2.0624	0.3250	0.5146	***9.7904***	**10.8864**	9.4394	1.9155	7.7370	3.6200	7.8482
	7	9.6571	0.9125	1.6021	12.1319	9.8971	14.5855	7.5533	***18.2768***	15.7336	**67.5323**
	8	***198.3043***	1.1328	9.1089	79.3965	19.8550	**Inf**	128.9804	36.1038	15.7095	**Inf**

* **Inf** means infinity, which indicates the gray values of NB pixels shrink to zero after processing.
* Bold text with underline indicates the best result. Bold italic text indicates the best result.

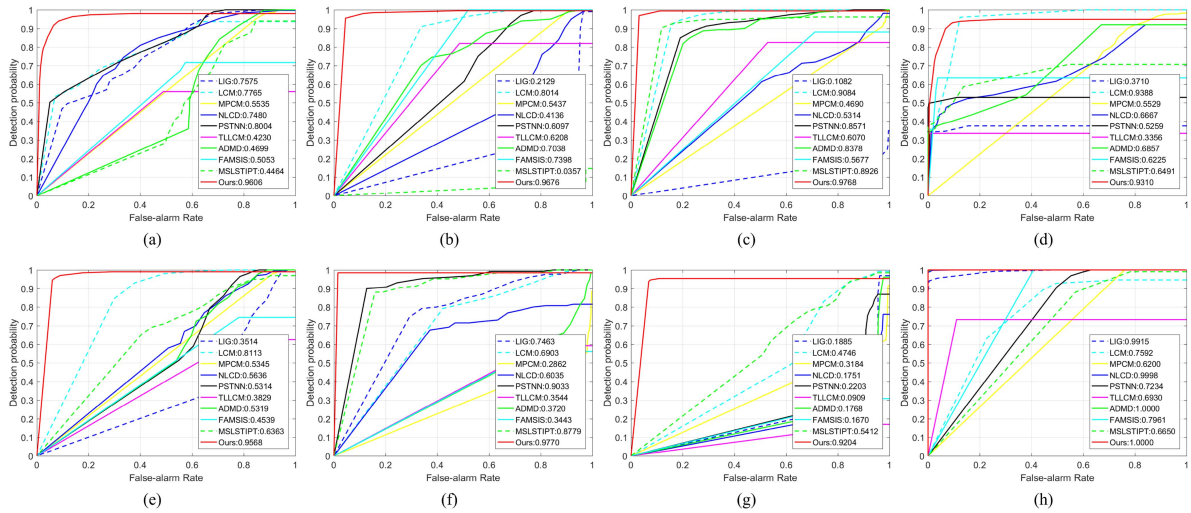

Fig. 16. ROC curves and AUC values of the 10 methods for Datasets 1 to 8. (a) Dataset 1. (b) Dataset 2. (c) Dataset 3. (d) Dataset 4. (e) Dataset 5. (f) Dataset 6. (g) Dataset 7. (h) Dataset 8.

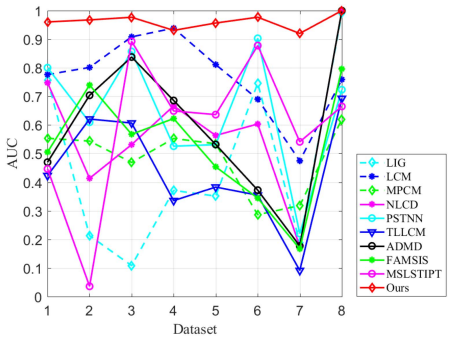

Fig. 17. AUC curves of the considered methods for all the eight datasets.

LCM and MSLSTIPT cannot eliminate the interference of islands. MPCM, NLCD, PSTNN, TLLCM, ADMD, and FAMSIS struggle to suppress heavy wave clutter. LIG can achieve target enhancement and clutter suppression compared with the other baseline methods. It is remarkable that our method performs the best with respect to target enhancement and background suppression. Enhanced results for Dataset 2 demonstrate that FAMSIS and our method outperform the other considered methods. For Dataset 3, LIG misses real target; LCM, PSTNN, ADMD, FAMSIS, MSLSTIPT, and our method achieve the satisfactory performance while MPCM, NLCD, and TLLCM are hard to handle heavy ocean waves with strong edges. For Dataset 4, there is residual strong clutter in enhanced results by MPCM, NLCD, and TLLCM; LIG, PSTNN, ADMD, and MSLSTIPT fail to enhance saliency of all the three small infrared targets. By comparison, FAMSIS and our method enhance all the three targets and suppress background clutter simultaneously. Moreover, FAMSIS and our method perform better in background suppression on Dataset 5 than other baseline methods. For Dataset 6, NLCD, PSTNN, and our method perform well in enhanced result while the other baseline methods are deficient in suppressing the interference of strong ocean wave clutters. For Dataset 7, the thick cloud edge and heavy waves cause great challenge for target detection; however, our proposed method can

TABLE IV
RUNNING TIME (IN SECONDS) OF 10 METHODS ON THE EIGHT DATASETS

Datasets	LIG	LCM	MPCM	NLCD	PSTNN	TLLCM	ADMD	FAMSIS	MSLSTIPT	Ours
1	1.2013	0.0503	0.0399	0.2077	0.2053	4.9673	0.0264	0.0297	3.8335	0.4596
2	1.1508	0.0484	0.0375	0.2074	0.2293	4.6802	0.0253	0.0287	3.9082	0.6217
3	1.1781	0.0483	0.0412	0.1416	0.2232	5.2332	0.0240	0.0270	4.2469	0.6252
4	1.1487	0.0479	0.0374	0.1763	0.1202	4.6225	0.0238	0.0210	4.0188	0.4532
5	1.1654	0.0497	0.0380	0.2028	0.2174	4.7108	0.0249	0.0293	3.8709	0.5889
6	1.1628	0.0496	0.0351	0.1815	0.1913	4.8149	0.0275	0.0295	3.7556	0.6305
7	1.1896	0.0581	0.0355	0.2045	0.2078	4.6287	0.0223	0.0285	3.8456	0.4879
8	7.0883	0.2405	0.2113	0.7241	1.2069	29.5464	0.0641	0.1530	16.4685	0.5610

achieve the preferable performance in both target enhancement and clutter suppression. From the last column of Fig. 15, one can see that LCM, MPCM, PSTNN, FAMSIS, and MSLSTIPT perform poorly in suppressing background clutter relatively. The performance of LIG, NLCD, TLLCM, ADMD, and our method are satisfactory in enhancing target and suppressing background for Dataset 8. On the whole, our method achieves preferable and robust performance on all the eight infrared maritime images with strong ocean waves and interference.

2) Quantitative Analysis: To further illustrate the performance of our method objectively, the LCG and BSF are measured for eight representative images in each dataset. From the definition, the higher the values of LCG and BSF are, the better performance of the corresponding method achieves. And the results by different methods are listed in Table III. It is clear that our method obtains the highest values of LCG, and almost gets higher BSF. That is, our method is superior to others. Furthermore, we provide the ROC curves of various methods on eight datasets to demonstrate the performance in terms of detection accuracy. It can be intuitively seen from Fig. 16 that the ROC curves of our method is closest to the upper left corner. In order to objectively validate the result, we provide the AUC values of various methods in the legend of each graph. For visual convenience, Fig. 17 shows the AUC curves of different methods on all the eight datasets, which intuitively illustrates the robustness and effectiveness of considered methods. It can be seen that our method is superior to the baseline methods for almost datasets. In addition, the running time of different methods on the eight datasets is listed in Table IV. One can see that our method can accomplish infrared maritime target detection in around 0.5 s. And our method can be accelerated through GPU to meet the needs of practical engineering applications.

IV. CONCLUSION

In this article, an effective and robust method using GVFC of infrared image is proposed to detect small maritime targets under intricate scenarios, simultaneously combining gradient property, directionality, and local dissimilarity. GVDM is applied to eliminate strong ocean wave interference, which makes full use of distribution characteristics of gradient direction angle and gradient modulus. Therein, an SHT based on K-S test with ADSD is designed to test the uniformity of direction angle, which conduces to further filtrate RSTs combined with regression analysis of gradient modulus. Furthermore, GMHLD is introduced to achieve satisfactory performance of target enhancement and residual ocean wave clutter suppression. Extensive experiments illustrate that the proposed GVFC achieves preferable and robust detection performance compared with the baseline methods, which presents not only in target enhancement and background suppression but also in detection accuracy. However, in our model, only directional dissimilarity is employed to enhance target saliency, which may be unsatisfactory for enhancing multiple targets with diverse intensity and contrast. More attributes need to be excavated to enhance targets. Furthermore, parallel acceleration will be implemented to achieve real-time performance.

REFERENCES

[1] L. Dong, B. Wang, M. Zhao, and W. Xu, "Robust infrared maritime target detection based on visual attention and spatiotemporal filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 3037–3050, May 2017.

[2] B. Wang, Y. Motai, L. Dong, and W. Xu, "Detecting infrared maritime targets overwhelmed in sun glitters by antijitter spatiotemporal saliency," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5159–5173, Jul. 2019.

[3] G. Paravati, A. Sanna, B. Pralio, and F. Lamberti, "A genetic algorithm for target tracking in FLIR video sequences using intensity variation function," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 10, pp. 3457–3467, Oct. 2009.

[4] P. Zhang, L. Zhang, X. Wang, F. Shen, T. Pu, and C. Fei, "Edge and corner awareness-based spatial-temporal tensor model for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10708–10724, Dec. 2021.

[5] A. Bal and M. S. Alam, "Automatic target tracking in FLIR image sequences using intensity variation function and template modeling," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 5, pp. 1846–1852, Oct. 2005.

[6] H. K. Liu, L. Zhang, and H. Huang, "Small target detection in infrared videos based on spatio-temporal tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8689–8700, Dec. 2020.

[7] P. Du and A. Hamdulla, "Infrared moving small-target detection using spatial-temporal local difference measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1817–1821, Oct. 2020.

[8] Y. Sun, J. Yang, and W. An, "Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3737–3752, May 2021.

[9] D. Pang, T. Shan, P. Ma, W. Li, S. Liu, and R. Tao, "A novel spatiotemporal saliency method for low-altitude slow small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Jan. 2022, Art. no. 7000705, doi: 10.1109/LGRS.2020.3048199.

[10] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.*, vol. 43, no. 6, pp. 2145–2156, Jun. 2010.

[11] H. Zhu, J. Zhang, G. Xu, and L. Deng, "Balanced ring top-hat transformation for infrared small-target detection with guided filter Kernel," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 5, pp. 3892–3903, Oct. 2020.

[12] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Proc. SPIE*, vol. 3809, pp. 74–83, Oct. 1999.

[13] T. Soni, J. R. Zeidler, and W. H. Ku, "Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 327–340, Jul. 1993.

[14] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.
[15] Y. He, M. Li, J. Zhang, and Q. An, "Small infrared target detection based on low-rank and sparse representation," *Inf. Phys. Technol.*, vol. 68, pp. 98–109, Jan. 2015.
[16] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image Vis. Comput.*, vol. 63, pp. 1–9, Jul. 2017.
[17] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3752–3767, Aug. 2017.
[18] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, Feb. 2019, Art. no. 382.
[19] X. Kong, C. Yang, S. Cao, C. Li, and Z. Peng, "Infrared small target detection via nonconvex tensor fibered rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Apr. 2022, Art. no. 5000321, doi: 10.1109/TGRS.2021.3068465.
[20] M. Zhao, W. Li, L. Li, P. Ma, Z. Cai, and R. Tao, "Three-order tensor creation and tucker decomposition for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5000216, doi: 10.1109/TGRS.2021.3057696.
[21] B. Zhao, C. Wang, Q. Fu, and Z. Han, "A novel pattern for infrared small target detection with generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4481–4492, May 2021.
[22] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Attentional local contrast networks for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9813–9824, Nov. 2021.
[23] N. Wang, B. Li, X. Wei, Y. Wang, and H. Yan, "Ship detection in spaceborne infrared image based on lightweight CNN and multisource feature cascade decision," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4324–4339, May 2021.
[24] J. Ma, L. Tang, M. Xu, H. Zhang, and G. Xiao, "STDFusionNet: An infrared and visible image fusion network based on salient target detection," *IEEE Trans. Instrum. Meas.*, vol. 70, Apr. 2021, Art. no. 5009513.
[25] J. Du, H. Lu, M. Hu, L. Zhang, and X. Shen, "CNN-based infrared dim small target detection algorithm using target-oriented shallow-deep features and effective small anchor," *IET Image Process.*, vol. 15, no. 1, pp. 1–15, Jan. 2021.
[26] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
[27] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.
[28] Y. Qin and B. Li, "Effective infrared small target detection utilizing a novel local contrast method," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1890–1894, Dec. 2016.
[29] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.
[30] P. Du and A. Hamdulla, "Infrared small target detection using homogeneity-weighted local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 514–518, Mar. 2020.
[31] D. Liu, L. Cao, Z. Li, T. Liu, and P. Che, "Infrared small target detection based on flux density and direction diversity in gradient vector field," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 7, pp. 2528–2554, Jul. 2018.
[32] R. Lu, X. Yang, W. Li, J. Fan, D. Li, and X. Jing, "Robust infrared small target detection via multidirectional derivative-based weighted contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 7000105.
[33] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, Oct. 2016.
[34] S. Moradi, P. Moallem, and M. Sabahi, "Fast and robust small infrared target detection using absolute directional mean difference algorithm," *Signal Process.*, vol. 177, Dec. 2020, Art. no. 107727.
[35] R. Lu et al., "Infrared small target detection based on local hypergraph dissimilarity measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Dec. 2022, Art. no. 7000405, doi: 10.1109/LGRS.2020.3038784.
[36] J. Han, S. Moradi, I. Faramarzi, C. Liu, H. Zhang, and Q. Zhao, "A local contrast method for infrared small-target detection utilizing a tri-layer window," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1822–1826, Oct. 2020.
[37] H. Zhang, L. Zhang, D. Yuan, and H. Chen, "Infrared small target detection based on local intensity and gradient properties," *Inf. Phys. Technol.*, vol. 89, pp. 88–96, Mar. 2018.
[38] Y. Chen, G. Zhang, Y. Ma, J. U. Kang, and C. Kwan, "Small infrared target detection based on fast adaptive masking and scaling with iterative segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Jan. 2022, Art. no. 7000605, doi: 10.1109/LGRS.2020.3047524.
[39] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Infrared small-target detection using multiscale gray difference weighted image entropy," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 60–72, Feb. 2016.
[40] P. Yang, L. Dong, and W. Xu, "Detecting small infrared maritime targets overwhelmed in heavy waves by weighted multidirectional gradient measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 7002005.
[41] S. Qi, J. Ma, C. Tao, C. Yang, and J. Tian, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 495–499, May 2013.
[42] X. Bai and Y. Bi, "Derivative entropy-based contrast measure for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2452–2466, Apr. 2018.
[43] X. Cao, C. Rong, and X. Bai, "Infrared small target detection based on derivative dissimilarity measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 3101–3116, Aug. 2019.
[44] B. Xiong, X. Huang, and M. Wang, "Local gradient field feature contrast measure for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 553–557, Mar. 2021.
[45] Y. Qian et al., "Infrared small target detection based on saliency and gradients difference measure," *Opt. Quant. Electron.*, vol. 52, Feb. 2020, Art. no. 151.
[46] Y. Xiao, "A fast algorithm for two-dimensional Kolmogorov-Smirnov two sample tests," *Comput. Statist. Data Anal.*, vol. 105, pp. 53–58, Jan. 2017.
[47] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[48] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2106–2112, Nov. 2010.
[49] S. Huang and J. Yang, "Linear discriminant regression classification for face recognition," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 91–94, Jan. 2013.
[50] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.
[51] P. Yang, L. Dong, and W. Xu, "Infrared small maritime target detection based on integrated target saliency measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2369–2386, Jan. 2021.
[52] Y. Qin, L. Bruzzone, C. Gao, and B. Li, "Infrared small target detection based on facet Kernel and random walker," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7104–7118, Sep. 2019.
[53] H. Zhu, H. Ni, S. Liu, G. Xu, and L. Deng, "TNLRS: Target-aware non-local low-rank modeling with saliency filtering regularization for infrared small target detection," *IEEE Trans. Image Process.*, vol. 29, pp. 9546–9558, Oct. 2020, doi: 10.1109/TIP.2020.3028457.

Ping Yang received the B.E. degree in optoelectronics information engineering and the Ph.D. degree in information and communication engineering from Dalian Maritime University, Dalian, China, in 2016 and 2022, respectively.

She is working with PLA Strategic Support Force Information Engineering University, Zhengzhou, China. Her research interests include target detection, image processing, and laser point cloud imaging and error analysis.

Lili Dong (Member, IEEE) received the B.E. degree in mechanical design manufacturing and automation, the M.E. degree in machine-electronic engineering, and the Ph.D. degree in instrument science and technology from the Harbin Institute of Technology, Harbin, China, in 2002, 2004, and 2008, respectively.

From 2005 to 2008, she was a Teaching Assistant with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. From 2008 to 2012, she was a Lecturer with Dalian Maritime University and an Associate Professor from 2012 to 2019. Since 2019, she has been a Professor of Information Science and Technology with Dalian Maritime University. She has authored more than 60 articles and 3 inventions. Her research interests include multispectral target recognition, tunnel lighting, and photoelectric detection.

Wenhai Xu received the B.E. and M.E. degrees in precision instrument and the first Ph.D. degree in imprecision instrument from the Harbin Institute of Technology, Harbin, China, in 1982, 1984, and 1991, respectively, and the second Ph.D. degree in manufacturing machine from the Tokyo Institute of Technology, Tokyo, Japan, in 1993.

From 1986 to 1988, he was a Lecturer with the Harbin Institute of Technology and an Assistant Professor from 1992 to 2001. He has been a Professor with the Harbin Institute of Technology for four years since 2001. He was the Project Director with Cannon Inc., Tokyo, from 1993 to 2003. He was also a Research Scientist with System Engineers Company, Ltd., Yamato City, Japan, from 1995 to 1997. He is currently a Professor of Opt-Electric Information Science and Engineering with Dalian Maritime University, Dalian, China. In the last 10 years, he has directed more than 30 research projects and applied 10 national patents. He has authored more than 100 research papers. His research interests include infrared detection, digital image processing, design of high-resolution optical imaging system, and optelectronic information processing.