Iterative infrared ship target segmentation based on multiple features

Zhaoying Liu\textsuperscript{a,c,d}, Fugen Zhou\textsuperscript{a,c}, Xiaowu Chen\textsuperscript{b}, Xiangzhi Bai\textsuperscript{a,b,c,e}, Changming Sun\textsuperscript{d}

\textsuperscript{a} Image Processing Center, Beihang University, Beijing 100191, China
\textsuperscript{b} State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China
\textsuperscript{c} Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, China
\textsuperscript{d} CSIRO Computational Informatics, Locked Bag 17, North Ryde, NSW 1670, Australia

A B S T R A C T

This paper presents an efficient method for ship target segmentation in infrared (IR) images. It consists mainly of two procedures: iterative image segmentation and ship target selection. First, based on the intensity distribution of an IR image, we design a global background subtraction filter (GBSF) to suppress the background, and an adaptive row mean subtraction filter (ARMSF) to enhance the target. After iteratively applying these two filters, we can obtain a proper threshold for image segmentation. Second, based on the geometric properties of the ship target, we construct four shape features and a selection criterion to identify the real target and remove the non-target regions. Experimental results demonstrate that the proposed method can effectively segment ship targets from different backgrounds in IR images. The advantage of the proposed method over the others in the previous literatures is validated in both visual and quantitative comparisons, especially for IR images with low contrast and uneven intensities.

1. Introduction

IR imaging systems detect relative differences of thermal radiations emitted from objects. The intensity of the emitted radiation depends on two factors: temperature and emissivity. Therefore, IR imaging systems could work during daytime and at night [1]. They are employed for both military and civil applications, such as long-range detection, automatic target recognition (ATR), video surveillance, and marine searching [2]. In recent years, IR imaging systems used for maritime surveillance have attracted much attention, including ship target detection, segmentation, tracking, and recognition [2,3]. Many researchers focus on IR dim ship targets which are usually small and lack of shape features [4]. For most of the present ATR systems, shape features of the segmented target are very important for target recognition. Therefore, precise IR ship target segmentation is desired, and this is the motivation of our work.

IR imagery may alleviate several problems in computer vision, such as the presence of shadow and sudden illumination changes. However, there are challenges with IR image itself. First, an IR image of sea surface records the total radiation of the ship target and background, including atmosphere and surface waves [3,5]. The intensity distribution is complex, and usually, the ship target is much smaller than the background. Second, heat exchange between the ship target and the background leads to a fuzzy outline of the target. Third, the ship target has uneven intensities because of the heating of the engine and chimney. In addition, limited by the imaging distance and IR imaging technology, IR images are characterized by low signal-to-noise ratio (SNR) and low target-to-background contrast [6]. All these challenges make ship target segmentation in IR images difficult.

To deal with the problem, many methods have been proposed. These methods could be roughly divided into four categories: methods based on thresholding [7], active contour models [8,9], mean-shift [10], and neural network [11,12]. Among the existing methods which are based on neural networks, Hopfield neural network (HNN) is commonly exploited for image segmentation. One identified problem of HNN is that it is difficult to choose penalty parameters. Although some criteria have been proposed to solve the problem, they are time-consuming [12].

Mean-shift is a robust feature-space analysis method, and it has been applied for image segmentation, clustering and tracking [13]. The segmentation is based on local regional merging. This may wrongly merge the target pixels into its neighborhood background when the contrast is low, or eliminate the target when its size is small [10].

The active contour model is defined based on curve evolution and geometric flows. The Chan–Vese model is one of the models which are effectively used for image segmentation [8]. Assuming that each image region is statistically homogeneous, the Chan–Vese model does not perform well for images having inhomogeneous
intensities [9]. To overcome this weakness, a novel active contour model driven by local image fitting energy (AC-LFE) has been proposed [9]. Using local image information as constraints, AC-LFE works well for images with inhomogeneous intensities. Nevertheless, the performance of AC-LFE can be affected by different initializations.

The thresholding method is often used in IR image segmentation because of its simplicity and efficiency. A survey on the thresholding method can be found in [14]. The well-known methods are Otsu’s [15], minimum error thresholding [16], entropy-based thresholding [17], and fuzzy c-means clustering [18]. One disadvantage of these classical methods is that spatial information in the image is not considered. In fact, the data in an image are inherently correlated [19]. To compensate for this shortcoming, local spatial information has been introduced, for example, local adaptive methods based on edge [19] and gradient information [20], 2D Otsu [21], 2D maximum entropy [22,23], and spatial fuzzy c-means [24]. These methods are less prone to noise, leading to more effective segmentation in images with low SNR. However, they are sensitive to the ratio between the pixel numbers in target and background regions. When the ratio is close to one, these methods perform well. When the target is much smaller than the background, these methods may wrongly classify the target as background. In addition, when the contrast is low between the target and the background, it is also hard to find a proper threshold for these methods.

In this paper, to segment the ship targets in IR images which are characterized by different sizes, low contrast, fuzzy outline, and uneven intensities, we propose a segmentation method based on multiple features, including the intensity properties and shape features. This method consists of two stages. First, we propose a global background subtraction filter (GBSF) to suppress the background, and an adaptive row mean subtraction filter (ARMSF) to enhance the target. After iteratively applying these two filters, a threshold is obtained to separate the target and the background. Then, we develop a target selection method based on shape features, including four shape descriptors and a selection criterion. Experimental results demonstrate that our method can effectively segment ship targets from different sea backgrounds. Comparison results indicate that our method achieves a better performance with lower misclassification error (ME) and lower relative foreground area error (RAE).

The rest of this paper is organized as follows. Section 2 presents the detailed procedure of our method. Section 3 demonstrates the main experimental results and discussions. Finally, Section 4 gives conclusions.

2. Segmentation algorithm

The flowchart of the proposed method is illustrated in Fig. 1. Details will be given in the following subsections.

2.1. Global background subtraction filter

Suppose the IR image is a gray scale image \(x\), which is expressed as:

\[ x = f + b, \quad x, f, b \in IR^N, \]

where \(f\) is the foreground or the target, \(b\) is the background, \(IR^N\) is the N-dimensional Euclidean space, where \(N=2\) for gray scale image. The purpose of ship target segmentation is to obtain the exact ship target and remove the background. As a complex background will increase the difficulties of ship target segmentation, it is necessary to initially suppress the background.

From sample IR images in maritime environment (Fig. 2), we can find three key characteristics. First, the ship target is small with the background occupying a large part of the image. Second, notable differences exist in the intensity distribution of different IR images, and it is difficult to model the distributions of the IR images. Third, in a single image, pixel intensities of the background concentrate in a narrow range because the thermal properties of the background are similar except for regions with some strong sea waves. According to these characteristics, we propose a GBSF to suppress the background. As shown in Fig. 3, pixel intensities of the background form a significant peak in the histogram. A histogram analysis, which is an effective way for intensity distribution description [25], is applied to describe the intensity properties of the IR image. The idea of GBSF is to estimate the background by finding the peak of the histogram, and then to subtract it from the image. Thus, we can suppress most of the background and obtain a sparse foreground image, noted as \(f_{se}\), and expressed as

\[ f_{se} = \begin{cases} x - b_e, & x \geq b_e, \\ 0 & \text{otherwise}, \end{cases} \]

(2)

where \(b_e\) represents the estimated intensity of background pixels, and is determined as

\[ b_e = k_{\text{max}} + ext_{\text{min}}. \]

(3)

In this expression, \(k_{\text{max}}\) is the pixel intensity corresponding to the peak of the histogram. Note that as the intensities of the background vary in a certain range, the background pixels occupy a certain width around the largest peak location in the histogram. To find the proper threshold, we extend \(k_{\text{max}}\) to the nearest valley location on the right hand side. The width of the extension is denoted as \(ext_{\text{min}}\), as shown in Fig. 3(b). The \(k_{\text{max}}\) and \(ext_{\text{min}}\) are determined using the following equations:

\[ k_{\text{max}} = \{(k^*) \text{(hist}(k^*) = \max) / \text{hist}(k), \quad k \geq \text{Mean}l\}. \]

(4)

\[ ext_{\text{min}} = \min \{\text{ext} \in N (dh(k_{\text{max}} + ext) < 0, \ dh(k_{\text{max}} + ext + 1) > 0)\}, \]

(5)

\[ dh(k) = \text{hist}(k + 1) - \text{hist}(k) \text{ for } 0 \leq k < 255; \quad dh(255) = 0, \]

(6)

where \(\text{hist}(k)\) is the histogram; \(k\) is the gray level, \(0 \leq k \leq 255\); \(dh\) \(k\) is the difference of \(\text{hist}(k)\) elements; \(\text{Mean}l\) is the average intensity value of the image.

![Flowchart of the proposed method.](Image 321x546 to 534x734)
For the tested IR images, the intensity of the background is generally greater than \( \text{Mean}_I \). To suppress the background sufficiently, we limit \( k_{\text{max}} \) to the range of \([\text{Mean}_I, 255]\). The results of GBSF show that the complex background can be well suppressed (see Fig. 3(c)).

### 2.2. Adaptive row mean subtraction filter

Using GBSF, we can effectively suppress the pixel intensities of the background. However, GBSF treats all pixels in the target and background regions equally. This will lead to the suppression of the ship target as well (see Figs. 3(c) and 4(a)). Thus, we propose an ARMSF to enhance the target, which is defined as

\[
E_{\text{ARMSF}}(R_j) = \begin{cases} 
R_j - \lambda_j \times mR_j, & R_j \geq mR_j \\
0, & \text{otherwise}
\end{cases}, \quad j = 1, ..., H. \tag{7}
\]

where \( R_j = [R_{j1}, R_{j2}, ..., R_{jW}] \) is the pixel intensity vector of the \( j \)th row; \( mR_j \) is the mean value of \( R_j \); \( H \) and \( W \) are the height and width of the image, respectively; \( \lambda_j = [\lambda_{j1}, \lambda_{j2}, ..., \lambda_{jW}] \) is the weight vector used to control the values subtracted from each pixel; \( E_{\text{ARMSF}}(R_j) \) is the output of ARMSF.

In Eq. (7), the weight vector \( \lambda_j \) is important for the performance of ARMSF. Our purpose is to enhance the target and suppress the background. Therefore, a large weight is expected for the background and a small one for the target. In the tested IR images, the target is brighter than its surrounding background, and appears as a salient region with a fuzzy outline, while the background appears without any significant saliency. Therefore, to distinguish the target from the background, we use region saliency and gradient information when calculating the weight \( \lambda_{ij} \) for each pixel at the \( i \)th column and \( j \)th row, expressed as

\[
\lambda_{ij} = \frac{1}{(e^{\text{sgn}(a1) - mR_j})}
\times e^{\text{sgn}(a2) - mR_j} + e^{\text{sgn}(a3) - mR_j}), \quad i = 1, ..., W; j = 1, ..., H, \tag{8}
\]

where \( \text{sgn}(a) \) is the sign function with \( \text{sgn}(a) = -1 \), if \( a < 0 \); and \( \text{sgn}(a) = 1 \), if \( a > 0 \); otherwise \( \text{sgn}(a) = 0 \); \( f_{ij} \), \( s_{ij} \), and \( g_{ij} \) represent the pixel value at position \((i, j)\) in the image \( f_s \), salient map \( S \), and gradient image \( G \), respectively; \( mR_j \) represent the mean pixel value in the \( j \)th row of the image \( f_s \), salient map \( S \), and gradient image \( G \), respectively.

The salient map \( S \) is calculated using the image signature as shown in Fig. 4(b) \([26]\). The gradient image \( G \) is calculated using the Sobel operator, as shown in Fig. 4(c). Associated with the intensity, region saliency, and gradient information, the weight \( \lambda_{ij} \) is adaptively adjusted for each pixel. For background pixels with a low intensity, low saliency and small gradient, the \( \lambda_{ij} \) is high; for target pixels with a high intensity, high saliency and large gradient, the \( \lambda_{ij} \) remains small. Therefore, we can enhance the ship target as shown in Fig. 4(d).

### 2.3. Iterative image segmentation method

Using GBSF and ARMSF, we can effectively suppress the background and enhance the target. Iteratively applying these two filters, ship targets can be separated from the background. We present an iterative method for IR image segmentation. The whole procedure is presented in Fig. 5, which contains the following steps:

1. **Step 1**: Given an image \( x \) with size \( H \times W \), set the maximum iteration number \( N_m \) and the initial segmentation region area \( F_{\text{Area}} = H \times W \).
2. **Step 2**: Perform GBSF on \( x \) to obtain the estimated gray level of the background \( T_b = b_o \), and the sparse foreground image \( f_s \).
3. **Step 3**: Calculate the salient map \( S \) and the gradient image \( G \) of image \( f_s \).
4. **Step 4**: Calculate the weight \( \lambda_{ij} \) and perform ARMSF on \( f_s \) to obtain the enhanced image \( Efx \).
Step 5: Segment the image $Efx$ using $T_b$ as a threshold, and obtain a binary image $BW$, where the target pixels are set to 1, the background pixels are set to 0. Then calculate the area of $BW$, denoted as $F_{Area}$.

Step 6: Calculate the difference of the current segmentation area and the previous one, denoted as $D=F_{Area}-F_{Area0}$. The ratio of the area difference to the image size, $D/(H \times W)$, is used for evaluating the iterative segmentation result. We assume that when the area change is small, a stable result is reached. That means, if $D/(H \times W) < T_s$, or the iteration number reaches the maximum number $N_m$, then stop and output $BW$ as the segmentation result; otherwise, set $F_{Area0}=F_{Area}, x=Efx$, and go to Step 2 to continue the iteration.

The values of $T_s$ and $N_m$ are determined based on the size of the ship target. In the IR image dataset, the sizes of the IR images range from $320 \times 256$ to $720 \times 576$ pixels. The size of the ship targets is greater than one percent of the image size, so we set $T_s=0.01$. In the experiments, for most of the IR images, stable segmentation result can be achieved after two to six iterations. Therefore, we set the maximum iteration number $N_m=6$.

2.4. Ship target selection based on shape features

After iterative segmentation, ship targets are successfully separated from the background. In the following, we develop a target selection method based on shape features to extract the ship targets. Although different types of ships have different geometric designs, they still have something in common, especially the shape of the outline. Different types of ship targets are segmented manually to fully display the shape features, as presented in Fig. 6, in which four intuitive common geometric properties could be observed. First, a ship target is a connected region with a certain size; second, the main body of the ship target has a rectangular shape; third, the ship target appears as a long narrow geometry; fourth, the upper-part of the ship target is narrow and the lower part is wide.

According to these observations, we use the following four features to identify the target: (i) Area ($Area$), which measures the size of the target; (ii) compactness ($Cp$), which describes the shape regularity; (iii) ratio of width to height ($R_{wh}$), where, for a ship target, the bounding box of the ship target is close to a rectangle, and the width is usually greater than the height; (iv) ratio of the upper part to bottom part ($R_{ub}$), where, for a ship target, the upper part is much narrower than the main body of the ship.
These features are expressed as follows:

\[
\text{Area} = \sum_{i=1}^{n} \sum_{j=1}^{w} f_{\text{seg}}(i,j),
\]

(9)

\[
C_p = 2\sqrt{\pi \cdot \text{Area}/P},
\]

(10)

\[
R_{\text{wh}} = w/h,
\]

(11)

\[
R_{\text{ub}} = TL/BL = (Trx-Tlx)/(Rbx-Lbx).
\]

(12)

Given a segmented region \(f_{\text{seg}}\), Area and \(P\) are the area and perimeter, respectively; \(w\) and \(h\) are the width and height of the smallest rectangle containing the region; \(TL\) and \(BL\) represent the length of the upper part and the bottom part, respectively; \(Rbx\) and \(Lbx\) represent the horizontal coordinates of the right bottom point and left bottom point, respectively. As the superstructures (the part above the main deck) of ships are different, we take the middle part of the superstructures as the upper part. That means, we take the \(TLy\) and \(LBy\) as the vertical coordinates of the right bottom point and left bottom point, respectively. Then the middle row of \(TLy\) and \(LBy\) is calculated, which is denoted as \((TLy+LBy)/2\); and the horizontal coordinates of the left point and right point of row \((TLy+LBy)/2\), denoted as \(Trx\) and \(Tlx\), are taken as the horizontal coordinates of the top left point and top right point. After that, the \(TL\) is calculated using \(Trx\) and \(Tlx\), as shown in Fig. 7.

Finally, we intend to remove the background, preserving only the target regions utilizing all the following criteria:

\[
t_{e1} < \text{Area} < t_{e2}, \quad C_p > t_c, \quad t_{w1h} < R_{\text{wh}} < t_{w2h}, \quad R_{\text{ub}} < t_{ub}.
\]

(13)

The thresholds in expression (13) are chosen experimentally, depending on the targets.

For our IR images, the smallest ship is larger than 200 pixels, and the largest target occupies less than 30% of the whole image, therefore we set \(t_{e1}=200\), \(t_{e2}=H \times W \times 0.3\). That means we discard regions smaller than \(t_{e1}\) or greater than \(t_{e2}\).

The ship target is a connected region with a regular shape, the compactness is between 0.1 and 0.6, so we set \(t_c=0.1\). That means we remove regions with compactness less than \(t_c\).

The images for ship targets are usually taken from the side, therefore, the outline appears as a narrow rectangle. That means the width is greater than the height, so we set \(t_{wh1}=1\). As there are images containing a long coastline, we set an upper limit of \(R_{\text{wh}}\), that is \(t_{wh2}=4\), to remove such backgrounds. Regions with \(R_{\text{wh}}\) smaller than \(t_{wh1}\) or greater than \(t_{wh2}\) are removed.

As illustrated in Fig. 7, the upper part of the ship target is much narrower than the bottom part. That means \(R_{\text{ub}}\) is much smaller than 1. For some small targets, the length of the upper part and the bottom part is similar, and \(R_{\text{ub}}\) is in the range of [0.8, 0.9]. We set \(t_{ub}=0.9\), and discard regions with \(R_{\text{ub}}\) greater than \(t_{ub}\). The final result can be found in Fig. 8(h).

3. Experimental results

Experiments were carried out on an IR image dataset supplied by co-researchers. The IR images were taken by a fixed ship-borne camera, which are mainly used in maritime surveillance. The dataset consists of 200 IR images that have different ship targets and backgrounds. The sizes of the IR images range from 320 × 256 to 720 × 576 pixels. The images were processed using an AMD Phenom (TM) II X4960T processor, 3.00 GHz CPU, with 4.00 GB RAM.

3.1. IR image processing results

We tested the proposed method using the parameters given in Section 2. Some experimental results are shown in Figs. 8–12.

Fig. 8 shows the ship target segmentation results of the first image in Fig. 2. The ship target is small, bright, and appears as a salient region with an unclear outline, the sea background is rough with strong waves. Fig. 8(a) is the original IR image. Fig. 8(b) and (c) are the results of GBSF and ARMSF, respectively. Fig. 8(d)–(g) show the segmentation results after one to four iterations. Fig. 8(h) is the final segmented ship target. Fig. 8(b) shows that the complex sea background is effectively suppressed using GBSF. Fig. 8(c) shows that the ship target is enhanced using ARMSF, and the contrast between the ship target and the background is greatly improved. Fig. 8(d)–(g) show that the ship target is gradually separated from the background after four iterations. Then, with the proposed target selection method, we successfully removed the background and obtained the ship target (Fig. 8(h)).

Fig. 9 shows the segmentation results of an IR image with a high contrast. Besides the ship target, sea and coastal background, there are also some vegetation and ripples. These objects would increase the difficulty of ship target segmentation. Fig. 9(a) is the original IR image. Fig. 9(b) and (c) present the results after using GBSF and ARMSF, respectively. Fig. 9(d)–(g) show the segmentation results after one to four iterations. Fig. 9(h) is the final segmented ship target. From Fig. 9 we can see that although the ripples are spatially connected to the ship target, we still can effectively remove them using the proposed iteration strategy.
Finally, followed by the target selection method based on shape features, we successfully removed other objects such as the coastline and the vegetation, as shown in Fig. 9(h).

Fig. 10 shows the experimental results of an IR image with an uneven background, the ship target is surrounded by a bright background. Fig. 10(a) is the original IR image. Fig. 10(b) and (c) show the results after using GBSF and ARMSF, respectively. Fig. 10(d) and (e) are the segmentation results of the first and second iteration, respectively. Fig. 10(f) is the final ship target after region selection. It can be seen from Fig. 10(b) and (c) that after using GBSF and ARMSF, there are still parts of the bright background which have not been suppressed. As these bright background pixels are far away from the target, they can be separated from the target (Fig. 10(e)). Using the target selection method, the...
remaining bright backgrounds are successfully removed and the ship target is obtained (Fig. 10(f)).

Fig. 11 shows the segmentation results of an IR image with a low contrast. A small ship target is embedded in a large sea background. The ship target appears as a bright region with great saliency, while the background region is uniformly distributed without great saliency. Fig. 11(a) is the original IR image. Fig. 11 (b) and (c) represent the results after using GBSF and ARMSF, respectively. Fig. 11(d)–(g) show the segmentation results of the first to third iteration. Fig. 11(g) is the final ship target segmentation result. It can be seen from Fig. 11(b) and (c) that by applying GBSF and ARMSF, we can effectively suppress the sea background.

Fig. 9. Segmentation results of ship target with high contrast: (a) original IR image, (b) the background suppression result using GBSF, (c) the target region enhancement result using ARMSF, (d)–(g) the first to fourth iteration thresholding results, respectively and (h) the final ship target segmentation result.
and enhance the target region. Fig. 11(d)–(f) demonstrate that by the use of iterative processing, the ship target is separated from the background. After three iterations, the area of the segmented region changes very little compared with the previous iteration. This means that a stable state is reached. Finally, we successfully remove all the background noises and extract the ship target (Fig. 11(g)).

Fig. 12 shows the segmentation results of an IR image having three ship targets. Fig. 12(a) is the original IR image. (b) and Fig. 12 (c) are the results after using GBSF and ARMSF, respectively. Fig. 12 (d)–(g) show the segmentation results of the first to fourth iteration. Fig. 12(h) is the final ship targets segmentation result. Fig. 12(b) and (c) demonstrate that after applying GBSF and ARMSF, the background can be effectively suppressed, while the three ship targets are well preserved. Fig. 12(d)–(g) show that by iteratively processing, all the three ship targets are gradually separated from the background. A stable segmentation result is reached after four iterations. Finally, using the target selection method, we successfully remove the background and obtain all the three targets (Fig. 12(h)). From Fig. 12 we can see that although the three ship targets have different sizes, their intensity distribution and geometric features are similar. In our method, except for the threshold $T_0$, all the other parameters are not limited by the sizes of the ship targets. Therefore, our method can successfully segment all the three ship targets.

The above experimental results show that our method is effective and robust. Provided that the ship target has great visual saliency and with available shape information, our method can achieve promising results under different sea backgrounds. In addition, our method can also work well for IR images with multiple ship targets.

3.2. Comparison results

To verify the effectiveness of the proposed method, comparisons were performed with nine existing image segmentation approaches: the Otsu’s (Otsu) [15], the maximum entropy (Entropy) [17], the fuzzy c-means method (FCM) [18], the minimum error thresholding method (MinError) [16], the 2D entropy (2D entropy) [22], the spatial fuzzy c-means (SFCM) [24], the mean-shift (MS) [13], the Chan–Vese active contour model (C-V) [8], and the AC-LFE [9].

The control parameters of each method are chosen and tuned manually according to a number of experiments. Firstly, we test all the IR images with default parameters suggested in the corresponding references. Then, based on the characteristics of each IR image (e.g., the size of the target, the contrast and the corresponding references), we adjust the parameters to obtain the possibly best result. The parameters of each method are as follows. In the 2D entropy method, the neighborhood size $n$ is in the range of $\{3, 5, 7, 9\}$. In the SFCM method, the window size is in the range of $\{5 \times 5, 7 \times 7, 9 \times 9\}$, the iteration stopping threshold $r=5$, the parameters $p, q$ are set as $p=1, q=1$ or $p=0, q=2$. The bandwidth of MS method $h_b$ is in the range of $\{0.05, 0.07, 0.09\}$. The initialization type of the C-V method is “whole”, which automatically divide the image into small chessboard with a size of $9 \times 9$ pixels as initials. In the AC-LFE method, the center and the radius of the initial contour are interactively selected according to the position and the size of the ship target. The comparison results are presented in Figs. 13 and 14. We can clearly see the performances of each method from the figures. It needs to be noted that the selection of the control parameters is difficult. It could not be guaranteed that the parameters chosen are the best ones. However, in our experiments, we have tuned many parameters and tried our best to find the optimal one among them. Therefore, the comparison is reasonable.

The Otsu method assumes that the image to be thresholded contains bimodal histogram (the foreground and the background), and then calculates the optimum threshold to separate those two classes. The Otsu method could successfully segment IR images with high contrasts (for example, images of the third and fourth columns in Fig. 13), and some IR images with uneven backgrounds (e.g., images of the first and second columns in Fig. 13). However, the Otsu method could not work well for IR images with low

![Figure 10](image-url) Segmentation results of ship target with uneven background: (a) original IR image, (b) the background suppression result using GBSF, (c) the target region enhancement result using ARMSF, (d), (e) the first and second iteration thresholding results, respectively and (f) the final ship target segmentation result.
contrasts (e.g., images in Fig. 14), as they generally contain a unimodal histogram. In addition, for IR images with small ship targets, they generally contain a unimodal histogram or a bimodal histogram formed by different backgrounds (e.g., with sea and sky backgrounds). This will also lead to misclassification. For example, in the image of the fifth column in Fig. 13, the ship target is small, thus the pixels of the sky and sea background form a bimodal histogram. Using the Otsu method, the ship target is wrongly classified as sky background.

The MinError method usually assumes that the target and background pixels are normally distributed. It uses gray level histogram as an estimation of the probability density function. Consequently, for IR images with uneven background and IR images with small ship targets (e.g., images of the first, third, and fourth columns in Fig. 13, and images of the first, second, and fifth columns in Fig. 14), this method does not work well. However, as the MinError method is based on histogram analysis, it performs better than the Otsu method for images with low contrasts (e.g., the images of the third and fourth columns in Fig. 14).

The Entropy and 2D entropy methods are both based on the maximum entropy principle. These methods assume that the best threshold should maximize the total entropy of an image. The Entropy method uses 1D histogram to estimate the probability distribution, and then calculates the entropies with the distributions. The threshold is obtained by maximizing the total entropies of the target and the background. The 2D entropy method is based on 2D histogram of the image. It takes into account the local average intensity values, using both gray level distribution and spatial information. The performances of the two entropy-based methods are comparable. The Entropy method obtains a better result than the 2D entropy method for images with uneven backgrounds, for example, for images of the second and fourth columns in Fig. 13. On the other hand, the 2D entropy method performs better than the Entropy method for images with low contrasts and images with strong sea surface waves (e.g., the images of the fifth column in Fig. 14). However, these two methods are both sensitive to the pixel percentages of the ship target and the background. When the ship target is much smaller than the background, they both fail (e.g., the images of the second and the fifth columns in Fig. 14).

The FCM method is an unsupervised method commonly applied for image segmentation. The segmentation result is reached by iteratively minimizing a cost function. The cost function usually depends on the distance between the pixels and the cluster centers. The SFCM method incorporates spatial information by considering the cluster distribution in the neighborhood. It is
more robust to noise than the FCM method. However, for images with low contrasts, data points of the target and part of the background are similar in the feature space. In this case, neither the FCM method nor the SFMC method could successfully classify the target and the background. As can be seen from Fig. 14, both the FCM and the SFMC methods fail to segment almost all the images listed in Fig. 14. Nevertheless, by the use of the spatial information, the SFMC method performs better than the FCM method for images with uneven backgrounds (e.g., images in the first, second, and the third columns in Fig. 13).

The MS method is a feature-space analysis method based on local regional merging. This method may wrongly merge the target into its neighboring background when the intensities are inhomogeneous, or eliminate a target with a small size. Therefore,
for IR images with homogeneous intensities around the ship target, the MS method can achieve satisfactory results (e.g., images of the second and third columns in Fig. 14). For images where the intensities vary a lot or the target is too small, the MS method does not work well (e.g., image of the fifth column in Fig. 13, and image of the first column in Fig. 14).
The methods using active contour models (C–V, AC-LFE) are based on curve evolution and geometric flows, which have been successfully used for image segmentation. From Figs. 13 and 14 we can see that, due to the incorporation of local image information constraints, the overall performance of the AC-LFE method is better than the C–V method. Especially for images with low...
That is, $\text{ME} \in [0, 1]$, smaller value indicating better segmentation result.

### Table 1: Quantitative comparison results.

<table>
<thead>
<tr>
<th>Methods</th>
<th>ME Max</th>
<th>ME Min</th>
<th>ME Average</th>
<th>RAE Max</th>
<th>RAE Min</th>
<th>RAE Average</th>
<th>Average time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu</td>
<td>0.9730</td>
<td>0.0015</td>
<td>0.2947</td>
<td>0.9987</td>
<td>0.0065</td>
<td>0.8434</td>
<td>0.0779</td>
</tr>
<tr>
<td>MinError</td>
<td>0.9296</td>
<td>0.0009</td>
<td>0.5028</td>
<td>0.9989</td>
<td>0.0090</td>
<td>0.8040</td>
<td>0.1831</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.9726</td>
<td>0.0002</td>
<td>0.2645</td>
<td>0.9986</td>
<td>0.0015</td>
<td>0.7598</td>
<td>0.1599</td>
</tr>
<tr>
<td>2D entropy</td>
<td>0.9940</td>
<td>0.0001</td>
<td>0.3954</td>
<td>0.9997</td>
<td>0.0009</td>
<td>0.7742</td>
<td>0.8667</td>
</tr>
<tr>
<td>FCM</td>
<td>0.8067</td>
<td>0.0042</td>
<td>0.4790</td>
<td>0.9989</td>
<td>0.0035</td>
<td>0.9060</td>
<td>4.0084</td>
</tr>
<tr>
<td>SFCM</td>
<td>0.7841</td>
<td>0.0178</td>
<td>0.3826</td>
<td>0.9987</td>
<td>0.0008</td>
<td>0.8951</td>
<td>4.1363</td>
</tr>
<tr>
<td>MS</td>
<td>0.9710</td>
<td>0.0005</td>
<td>0.4954</td>
<td>0.9985</td>
<td>0.0381</td>
<td>0.8623</td>
<td>0.2082</td>
</tr>
<tr>
<td>C-V</td>
<td>0.9468</td>
<td>0.0083</td>
<td>0.3829</td>
<td>0.9986</td>
<td>0.1177</td>
<td>0.8689</td>
<td>4.8487</td>
</tr>
<tr>
<td>AC-LFE</td>
<td>0.6104</td>
<td>0.0004</td>
<td>0.0713</td>
<td>0.9709</td>
<td>0.0066</td>
<td>0.5737</td>
<td>54.4607</td>
</tr>
<tr>
<td>Our method</td>
<td>0.9637</td>
<td>0.0002</td>
<td>0.0259</td>
<td>0.9894</td>
<td>0.0034</td>
<td>0.3594</td>
<td>0.9435</td>
</tr>
</tbody>
</table>

**Equations**

The area of the segmented target is $A_T$ and the area of the reference target is $A_o$. From Eq. (15) we can see that a smaller RAE means a better segmentation. Since there is no ground-truth available, the reference target for the segmentation evaluation is provided by manual segmentations, which were made by the researchers in our group. Comparison results in terms of the value of RAE, ME, and the average processing time are listed in Table 1.

From Table 1 we can see that our method achieves better results with the lowest ME and RAE values on average. In other words, our method can segment most of the targets with a lower misclassification error, and preserve the contour of the ship target.

In terms of the processing time, our method is slower than the non-iterative methods, like the Otsu, MinError, Entropy, 2D entropy, and MS methods. However, our method is much faster than the FCM, SFCM, C-V, and AC-LFE methods. The reason is that most of the IR images have complex backgrounds, needing several iterations to achieve a satisfactory result. Fortunately, our method is mainly used for target feature analysis which can then be used for target recognition. For such applications, the requirements of the processing time are not so important compared with the accuracy. We can see from Table 1 that although our method is slower than the non-iterative methods, the time needed is still less than 1 s on average, and the accuracy is much better. Hence, our method can meet the requirement of such applications.

### 4. Conclusions

In this paper, we have proposed an effective method for ship target segmentation in IR images. There are two main contributions. First, based on the intensity properties of the ship targets and the backgrounds, we proposed an iterative segmentation method using GBSF and ARMSF. This can effectively separate ship targets from backgrounds. Second, based on the shape features, we developed a target selection method to identify the ship targets. Experimental results demonstrated that by exploring multiple information of the IR ship target image, the proposed method effectively segments ship targets with different sizes from different backgrounds in the IR images. In addition, the comparison results validate that the proposed method performs better than the other nine methods, especially for IR images with low contrasts and uneven intensities.

### Acknowledgements

This work was performed under the sponsorship of National Natural Science Foundation of China (61271023), R&D Program (2012BAH07B01), Program for New Century Excellent Talents in University (NCET-13-0020) and China Scholarship Council. Part of this paper has been presented in an international conference [29]. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions which are very helpful in improving the quality of the paper.
References


