Shipping Automatic Recognition System Based on Object Feature Extraction

Zhang Hongxin[*], Duan Kanghong and Zhang Xiaobo

North China Sea Marine Technical Support Center of State Oceanic Administration, Qingdao City, Shandong Province, 266000, China

Abstract: During the study on shipping object recognition, we propose a scheme based on improved canny also extract the moving object in video monitoring. The lifting wavelet transformation is adopted to provide smoothing detection for the image edges and the smoothed images are analyzed and evaluated in this paper. We also use improved Gauss function for traditional edge detection to smooth cubic b-spline lifting wavelet. It adopts continuous frame difference method to process the motivating area and obtains edge information by Canny detection. Then an improved Hu invariant moment algorithm is proposed to extract the invariant moment features of object images, which is also used to make automatic recognition for moving objects. The experimental results demonstrate the invariant moment values of the samples, acquired by Canny operator splitting, have obvious clustering effect. So it can offer more accurate recognition for shipping images.

1. INTRODUCTION

Moving object recognition is an important link for dynamic tracking of shipping on inland and it belongs to the key research in the field of computer vision. But the accuracy of moving object recognition is closely related to moving object detection. The correction of moving object detection is directly affecting moving object classification, target tracking and further behavior understanding. In actual implementation of video monitoring object tracking motion, it usually demands that it can detect and segregate the objects real-timely. The main task of object tracking is extracting moving object from video sequence. The common moving target detection algorithms include frame difference, background subtraction, optical flow and motion energy [1]. According to difference of actual application requirement, different detection algorithms are acquired in compromise among reliability, real-time and accuracy. The operation process of optical flow method and motion energy is very complicated and it has higher hardware requirement for computer. So it is very difficult to satisfy the requirement of real-time processing. Heeling error has strong adaptability and good robustness in dynamic environment, so it is suited to different dynamic environments. However, all related feature pixel points cannot be completely extracted and the cavity is easily produced inside the moving entity.

Many scholars and researchers propose to use artificial intelligence and pattern recognition to identify the detected objects to acquire accurate target recognition for ships. Srinivas, *et al.* [2] adopt the image processing and neural network technology to propose a static photos-based method to classify the ships. The photos features are extracted by image processing and these features are taken as the input of multi-level feed-forward neural network. Guo *et al.* in literature [3] discuss multi-sensor information fusion and AFNN-based object recognition in shipping classification. This method can be applied to the command and control system such as civil aviation, military affairs, etc. Toumi *et al.* [4] chooses the power spectrum characteristics extraction and bi-spectrum feature extraction algorithms. After the features are extracted, they use back propagation (BP) neural network model. Then the BP neural network is established and the network is trained to provide experiment results for recognition. Under complicated background of video surveillance, Zhang *et al.* [5] propose the recognition method based on inter-frame difference method and moment invariant features. In moving object detection algorithms, the definite pixel area is performed data analysis at first. The moving signals will be generated by data difference between frames to capture the moving objects in scenes. In the object recognition method, the moment invariant feature is extracted and the vectors used for standardization. Then the genetic wavelet neural network is adopted as mode discriminator to make automatic recognition of moving targets. PING *et al.* [6] study the accurate detection and tracking of moving ships in inland river navigation. Since traditional feature detection tracking method of singular gray cannot effectively describe the ship objects, they propose a shipping detection tracking method based on feature fusion and fuse gray feature to describe the moving objects. The examples of filtering algorithm are also used to establish observation probability model for the detection and tracking of moving ships.

From above researches it shows that the significance of moving object recognition has higher implementation value. Based on traditional target recognition method, image processing, neural network method and artificial intelligence, the ships on Inland River can be recognized to get accurate recognition effects. Then, the tracking algorithm is designed to track the recognized ships real-timely and to provide reliable basis for further behaviors of ships. Therefore, this paper introduces object edge detection and integrates the object edge information to the difference method. It not only removes the outliers but also offers more coherence for the object edge. It can provide complete detection objects and effectively avoids the interruption of noise and other factors. According to obtained edge features of the object images, an improved Hu invariant moment method is proposed to extract the invariant moment features of ships. Meanwhile, the effects of our improved algorithm are verified by theory and practice.

2. VIDEO IMAGE PRETREATMENT

Video detection is an effective means for navigational ships base surveillance on inland. It is an object detection technology with short distance and non-contact. So it has characteristics of high image resolution, visual observation, convenient materials transmission, convenient recycling, etc [7]. The image information of ship objects are detected to be gathered and these information are processed to describe the location and qualitative of objects. The CCTV video devices are usually installed at the front end of ocean ships, maritime patrol boat and the nearest coast-based high point of river surface. Thus, the quality of collected ship images is affected by multiple factors such as light, water traces, raindrop, heavy fog, etc. Meanwhile, due to the performance influence of video equipment, the obtained image sequence usually contains some noise. So the binaries image will have some isolated points and small gaps, which will influence the moving objects detection. Therefore, the original image of collected ships is filtered at first before calculation, so the images can be adjusted to reveal moving objects in image for further image segregation. The common image filter denosing methods include median filter, mean filter, wavelet filter, etc [8]. People will apply median filter to process isolated noise points in image sequence usually, but the median filter process has noise dispersion and detail preservation. Thus, this paper proposes a two-dimensional wavelet transformation filter method to process the image noise.

The shipping image can be seen as an energy constraint two-dimensional signal fundamentally, whose noise is eliminated by two-dimensional wavelet transformation. This paper proposes an image denoising method based on two-dimensional wavelet threshold transformation. The theoretical basis of two-dimensional threshold image denoising is: the wavelet transformation especially orthogonal wavelet transformation has strong correlation in data removing. It concentrates the signal energy in wavelet are in some large wavelet coefficients. However, the energy of image noise is distributed in the whole wavelet area. Therefore, after it is decomposed by wavelet, the wavelet coefficient amplitude will be larger than coefficient amplitude of noise. The wavelet coefficient with large amplitude usually takes signal as

the major part, while the coefficients with little amplitude may be noise in a large part. Therefore, we can preserve the signal coefficient by threshold method to reduce the noise coefficient to zero, to eliminate the noise of object images.

3. SHIPPING RECOGNITION BASED ON OBJECT FEATURE EXTRACTION

3.1. Image Denoısıng Based on B-Splıne Wavelet Transformatıon

Under general case, there are two objects for image smoothing: one is strengthening the pseudo-edge of image, including the weak edge; the other is internal smoothing of area aiming at the noise points. So we adopt cubic b-spline lifting wavelet in this paper. It can smooth the internal areas and strengthens the edge of image, which eliminates the conflict between above two problems. The cubic b-spline function and its first derivative are smoothing function and wavelet function [9]. Let $\beta^n(x)$ denote n_{th} symmetrical b-spline function, we get the equations:

$$\begin{cases} h(z) = \dfrac{3}{4} + \dfrac{1}{2}(z + z^{-1}) + \dfrac{1}{8}(z^2 + z^{-2}) \\ g(z) = \dfrac{5}{4}z^{-1} - \dfrac{5}{32}(1 + z^{-2}) - \dfrac{3}{8}(z + z^{-3}) - \dfrac{5}{32}(z^2 + z^{-4}) \end{cases} \tag{1}$$

By the lift of $h(z)$ and $g(z)$ the polynomial matrix $p(z)$ is acquired:

$$p(z) = \begin{bmatrix} 1 & \dfrac{1}{4}(1 + z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (1 + z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -\dfrac{3}{16}(1 + z^{-1}) \\ 3 & 1 \end{bmatrix} \begin{bmatrix} \dfrac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \tag{2}$$

The task of our system is determining the smoothing filter for an image with cubic b-spline parameters. We take account the effect of demand and noise in the image to adjust specific parameters. The convolution template in a 3×3 domain is:

$$\begin{bmatrix} 1/36 & 1/9 & 1/36 \\ 1/9 & 4/9 & 1/9 \\ 1/36 & 1/9 & 1/36 \end{bmatrix}$$

This paper lifts the cubic b-spline with the mode of parallel ranks. Because it has fast running speed and relative occupation on RAM. The computation for nonlinear lifting is small. The procedures in detail:

Step 1: Spiltting. The original image x_0 is split into three detailed high frequency parameters $y^{(1)}$, $y^{(2)}$, $y^{(3)}$ and one approximate low frequency parameter x :

$$\begin{cases} x(m,n) = x_0(2m, 2n) \\ y^{(1)}(m,n) = x_0(2m+1, 2n) - x_0(2m, 2n) \\ y^{(2)}(m,n) = x_0(2m, 2n+1) - x_0(2m, 2n) \\ y^{(3)}(m,n) = x_0(2m+1, 2n+1) - x_0(2m, 2n) \end{cases} \tag{3}$$
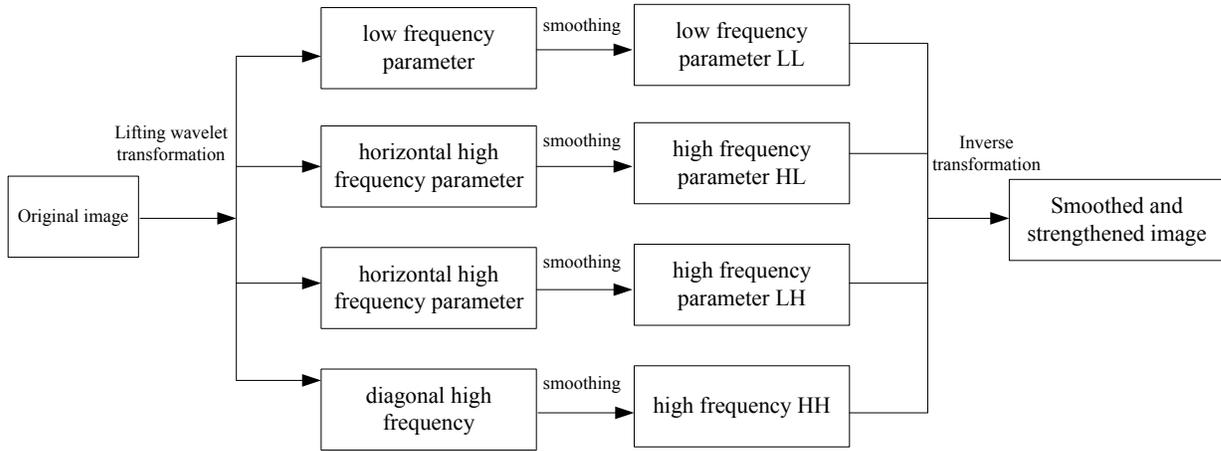
Fig. (1). Image smoothing and strengthening process with cubic b-spline lifting wavelet.

Step 2: Updating. The decision operator C is used to update x with the value which has the strongest relation to x from all directions by y. The definitions of decision operator C and updating operator U are:

$$y^*(m,n) = C(y^{(1)}(m,n), y^{(2)}(m,n), y^{(3)}(m,n)) \tag{4}$$

$$x^*(m,n) = U(x_0(m,n), y^*(m,n)) \tag{5}$$

Step 3: Prediction. We use updated approximate low frequency parameter x^* to predict detailed high frequency parameters $y^{(i)}$. Then the original $y^{(i)}$ will be replaced by predicted error $y^{(i)^*}$, that is, $y^{(i)^*}(m,n) = P(x^*(m,n), y^{(i)}(m,n)), i = 1,2,3$. The definition of prediction operator P is:

$$P(x^*(m,n), y^{(i)}(m,n)) = \\ y^{(i)}(m,n) - C(y^{(1)}(m,n), y^{(2)}(m,n), y^{(3)}(m,n))/2 \tag{6}$$

The inverse process of wavelet lifting has three steps: filter, update and integration. When an image is decomposed by lifting wavelet it will be divided into four parts of parameters: low frequency contour, vertical high frequency, horizontal high frequency and diagonal high frequency. To better extract the maximum edge information for the image, we first smooth each parameter. Then inverse transformation of the lifting wavelet is performed to each parameter. So the strengthened images after smoothing can meet the demand for edge detection. The process in detail is described as Fig. (**1**).

3.2. Invariant Moment Feature Extraction of Shiping Images

If one ship image is looked as a two-dimensional density distribution $f(x,y)$ and it is the Gray distribution function. The definition of its two-dimensional $(p+q)$ order geometric moments m_{pq} [10] is:

$$m_{pq} = \iint_\xi x^p y^q f(x,y)\, dx\, dy \tag{7}$$

ξ is the pixel space area of light function, that is, the domain range of $f(x,y)$. Assuming $f(x,y)$ is piecewise and continuous bounded function and its value is not zero in limited area on the plane, the moments of all the orders will be determined only by $f(x,y)$, according to uniqueness theorem [11], and vice versa. m_{pq} has relation to the location of coordinates. It has not translation invariance so it can not be taken as features. To satisfy the translation invariance we can move the coordinate origin of moment function to the center of image. Then the geometric features computed relative to light center can be defined as the center distance:

$$\mu_{pq} = \iint_\xi (x - x_0)^p (y - y_0)^q f(x,y)\, dx\, dy \tag{8}$$
$$p,q = 0,1,2,...,n$$

x_0, y_0 are the light centroid of gray image $f(x,y)$. $x_0 = g_{10}/g_{00}$, $y_0 = g_{01}/g_{00}$. Based on the definition of the center distance we can get the normalized center distance:

$$\mu_{p,q} = \frac{\mu_{p,q}}{\mu_{00}^{1+(p+q)/2}} = \frac{\mu_{p,q}}{g_{00}^{1+(p+q)/2}} \tag{9}$$

From above derivations we find that the geometric moments and centre distance can be used to describe the form of image, and the centre distance has no relation with the translation. Then the invariant moment invariants are created as:

$$\begin{cases} G_1 = \mu_{20} + \mu_{02} \\ G_2 = (\mu_{20} - \mu_{12})^2 + 4\mu_{11}^2 \\ G_3 = (\mu_{30} + \mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\ G_4 = (\mu_{30} - 3\mu_{12}) + (\mu_{21} + \mu_{03})^2 \\ G_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + 3\mu_{03})] \\ +3(\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ G_6 = (\mu_{20} - \mu_{12})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ +4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\ G_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\ -(\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{cases} \tag{10}$$

(**a**) Original figure

(**b**) Noising image

(**c**) Gauss smoothing with canny operator

(**d**) Cubic b-spline lifting wavelet smoothing
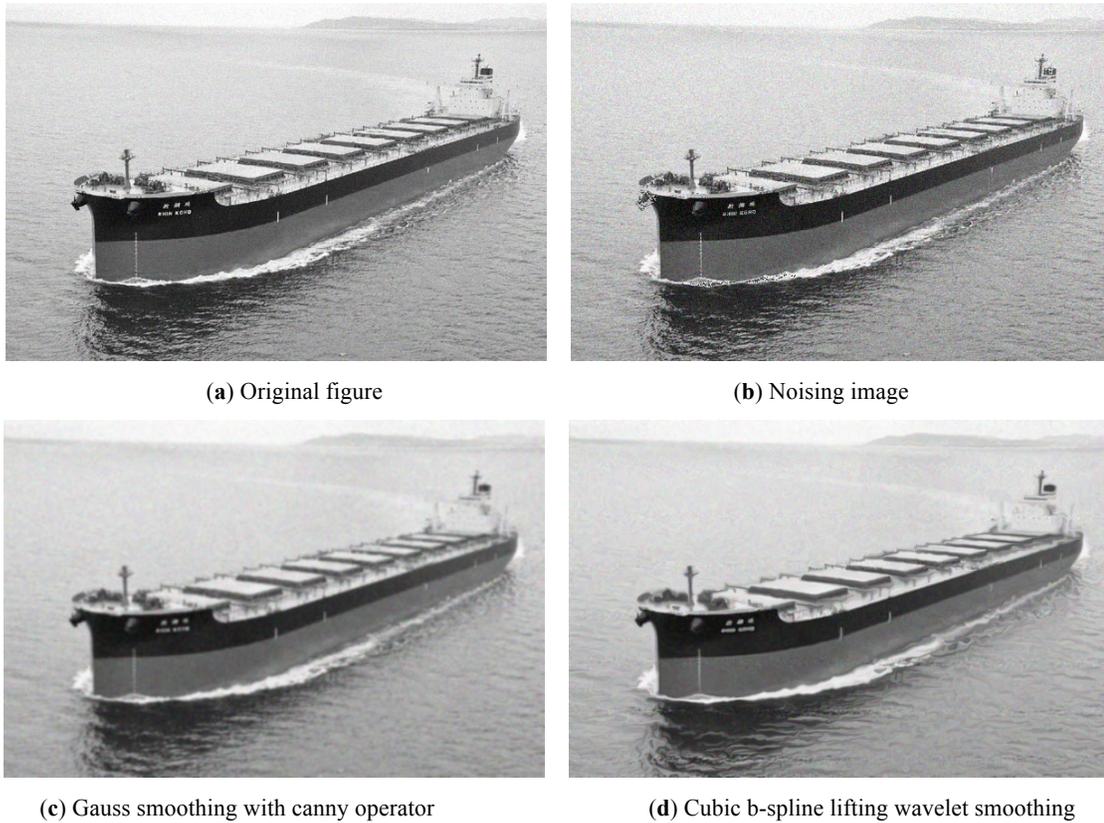
Fig. (2). Cubic b-spline lifting wavelet and Gauss smoothing effects.

These seven moment invariants have invariance in the translation, scaling, mirroring and rotation. But it has large dynamic changing range at the same time. It may cause negative value when extracting the features of object images. For convenience of the moment of objects and sequent neural network classification, we adopt logarithmic values to shrink the dynamic range. Thus the moment invariant in actual is:

$$h_k = |1g|G\|\ (k=1,2,3,...,7) \tag{11}$$

In actual implementation, acquired object images have difference not only in location and rotating states, but also in change scale and contrast. So aiming at acquiring the moment value as close as possible, after rotation, scale change or contrast change of the object images, we assume two images $f_1(x,y)$ and $f_2(x',y')$ have difference in the factors mentioned above. Their relation can be believed as the following equations:

$$f_1(x,y) = af_2(x',y') \tag{12}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \lambda \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \tag{13}$$

a is the contrast change factor of image; λ is scale change factor of image; β is the angle of rotation; a and b respective

ly denote the displacement in x-axis direction and y-axis direction. We compute seven invariant moments of $f_1(x,y)$ and $f_2(x',y')$ by logarithm and construct new invariant moments with equation 13: $M_1 = \sqrt{h_1}/h_1$, $M_2 = (h_1 + \sqrt{h_2})/(h_1 - \sqrt{h_2})$, $M_3 = h_3/h_2h_1$, $M_4 = h_4/h_3$, $M_5 = \sqrt{h_5}/h_4$, $M_6 = h_6/h_4h_1$, $M_7 = h_7/h_5$.

Then we get 7 new invariant moments with the improved algorithms. They have invariance in translation, scale and rotation under discrete conditions. So they can be taken as the features of shipping images for object recognition classification and matching.

4. EXPERİMENTS RESULTS ANALYSİS

4.1. The Denoising Effect Comparison

To verify the accuracy and reliability, we select random images of shipping series in MATLAB for simulations. The motivation background of the series of images is relatively simple with shrinking size of the motion objects. There is almost no other motivating background except the interference of water ripples, which provides ensuring to extract motivating objects effectively. The edge detection with improved canny operator is performed on the image and the results are shown in Fig. (**2**).

Table 1. Image edge dection and norm comparison.

Edge Detection Operator	Norm 1	Norm 2	Norm3	Times (ms)
Canny	50.3093	0.0039	0	256.598643
Wavelet	52.6398	0.0031	0	273.699881
Improved method	54.2210	0.0015	0	260.120453

From above figures we can intuitively see that the edge detection method based on the integration of cubic b-spine lifting wavelet and improved Canny operator has positioning accuracy for detection results and its denoising ability is strong. Our method makes full use of the feature of lifting wavelet transformation in multi-resolution analysis of noise suppression, and keeps the advantage of Canny operator for smoothing connection. By the edge detection on multiple images it is found that the algorithm in this paper shows better performance in denoising ability and edge positioning effect.

From the statistics of specific edge points of images as shown in Table **1**, we can see that the number of edge detection points of our algorithm is more than that of the scheme with traditional wavelet and Canny edge detection. The same results can be acquired by large amount of experiments. From the comparison of three norms in Canny, we find our algorithm really has improvement in SNR and positioning. Though the time consumption is not reduced greatly, it is almost the same as Canny algorithm. Therefore, the detected image edge with lifting wavelet and improved Canny algorithm has rich detailed information, it can depict the information of original image completely and keeps better consistency.

4.2. Clustering Analysis of Object Image Features

Since the categories of inland ships are various, this paper select 4 common types of ships for study from the gathered shipping video series. They are container ships, bulk carriers, passenger ships and boats. We select 5 pieces of each type and the total number is 20. First we adopt the recognition method based on edge detection mentioned above to process given samples. The shipping objects are extracted as the processing objects for the next moment feature value extraction. To verify the effectiveness of algorithm and extract accurate moment data, we choose the moment value of original images and extracted images of the samples. Then the clustering analysis is performed in selected 4 types of ships. If extracted features of object images can differentiate each object with other objects, satisfying the system demand, the selected feature value is effective. The clustering analysis on moment feature value of object images is shown in the following figures (We only list parts of the results for abbreviation). The four broken lines respectively denote the feature value distribution of four types of ships.

Similarly we can acquire the other four clustering results of improved algorithm. From the comparisons of above figures we know that the data by direct extraction for invariant moment of the original images are disorganized and has no rules to follow, which cause big difficulty to the next classification (Fig. **3**). While the curves in improved algorithm show the invariant moments have better ability in differentiation. They can differentiate the invariant moments of different types of images by data pretreatment and image segmentation. If the curves are far away it means the differentiation ability is strong and it is conducive to classification; If the curves are close and there is a intersection among them it means the features are closed and the differentiation ability is weak, tending to cause recognition errors. By the analysis on the clustering ability of 7 features, it is found that though there exists intersection and isolated features have drastic jumps, most of the features can differentiate the four kinds of objects waiting to be recognized. It verifies the features have better clustering ability. Analyzed in actual cases, the categories of inland ships are various and some ships are not stipulated to engage some kind of loading task, so it is hard to differentiate the type of ships just by contour and appearance. That will cause overlap of feature data representing different ships, influencing the accuracy of recognition system. When such cases emerge, it still needs artificial assistant recognition to judge the types of ship.

CONCLUSION

For the problems of AIS and VTS on object recognition for shipping, this paper uses the video as assistant means to make up the defects and improve the recognition accuracy on shipping object. At first, aiming at the deficiency of traditional moving objects detection, we further study the advantages and disadvantages of wavelet to process images, it combines theoretical basis of classical Canny operator and adopts the integrated cubic b-spline lift wavelet to improve the Canny operator edge detection algorithm. Then, the improved Hu invariant moment is adopted to extract features of object ship images. We make comparisons between the invariant moment of original images and the invariant moments of improved Canny operator segregation images. By the analysis on real data of shipping, the results prove the effectiveness of the improved method in this paper. It has high recognition accuracy on moving objects and better classification effects, which is helpful for dynamic tracking of ship objects.
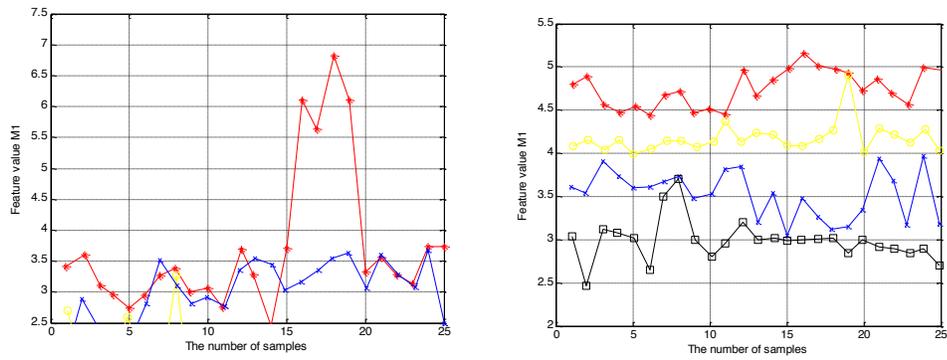
(**a**) Feature value M1
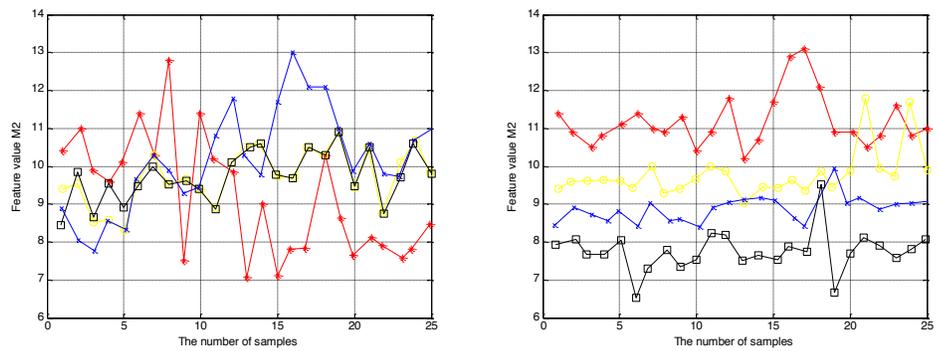
(**b**) Feature value M2
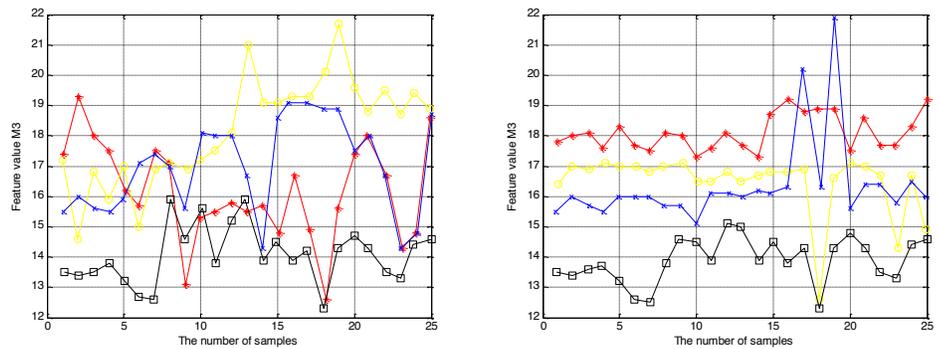
(**c**) Feature value M3

Fig. (3). Clustering analysis results of the feature value (the left are of original image and the right are of improved separations).

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

REFERENCES

[1]   W. Zhao, and S. Qın, "Comparative study on detection methods for video motion targets", *Science & Technology Review*, vol. 27, no. 10, pp. 64-69, 2009.

[2]   U. Srinivas, V. Monga, and R. G. Raj, "Automatic target recognition using discriminative graphical models," In: *Proceedings of International Conference on Image Processing*, Brisbane, QLD, Australia, 2011, pp.33-36.

[3]   Y. Guo, J. Wan, and M. Lu, "A parts-based method for articulated target recognition in laser radar data", *Optik*, vol. 124, no. 17, pp. 2727-2733, 2013.

[4]   A. Toumi, B. Hoeltzener, and A. Khenchaf, "Using Watersheds segmentation on ISAR image for automatic target recognition", In: *Proceedings of 2nd International Conference on Digital Information Management*, Lyon, France, 2007.

[5]   Y. Zhang, J.B. Gong, and J. W. Tian, "Learning complex cell features with cooperating pooling operation for object recognition", *Electronics Letters*, vol. 48, no. 17, pp. 1058-1059, 2012.

[6]   Y. Pıng, "Ship type identification based on neural network and D-S evidence theory", *Modern Electronics Technique*, vol. 13, no. 30, pp. 44-46, 2009.

[7]   X. Yin, C. Wang, and H. Zhang, "Vessel recognition with high resolution TerraSAR-X image based on structure feature", *Journal of Image and Graphics*, vol. 17, no. 1, 106-113, 2012.

[8]   J. Yang, M. Zhang, and Q. Pan, "Edges and details keeping filtering algorithm for color images", *Computer Engineering and Design*, vol. 31, no. 7, pp.1516-1544, 2010.

[9]   Y. Zhao, and Z. Quan, "An efficient wavelet wiener denoising algorithm," *Journal of Beijing University of Posts and Telecommunications*, vol. 27, no. 4, pp. 41-45, 2004.

[10]   X. Li, X. Feng, and Y. Ge, "Movable state recognition from ship-radiated noise", *Acta Acustica*, vol. 27, no. 5, pp. 443-448, 2002.

[11]   D. Kumlu, and J. B. Keith, "Autonomous ship classification using synthetic and real color images", In: *Proceedings of International Society for Optical Engineering*, Burlingame, CA, United states, 2013.