Methods for Underwater Sonar Image Processing in Objection Detection

Tang Xinyu, Zhang Xuewu(✉), Xu Xiaolong
College of Internet of Things Engineering
Hohai University
Changzhou, China
tangxinyu7@163.com

Sheng Jinbao, Xiang Yan
Nanjing Hydraulic Research Institute
Nanjing, China

***Abstract*—It is important to improve the integrity and accuracy of sonar image target detection, which is significant for underwater detection. In this paper, a variety of sonar image denoising algorithms and segmentation algorithms are studied, and a denoising algorithm based on fast curve transform is proposed. The image segmentation algorithm based on *k*-means clustering is studied, and the optimal clustering number screening and sonar image subsurface segmentation are realized. The sonar image fast segmentation algorithm based on ICM algorithm and the object contour detection of sonar image based on level set method are realized in Matlab. The results show that the proposed algorithm can improve the noise reduction effect of the sonar image under reverberation interference, and obtain a better image detection effect.**

Keywords-sonar image denoising; underwater target detection; clustering; markov random field; level set method

I. INTRODUCTION

Under the influence of underwater environment noise, ship's own noise and reverberation signal, Doppler shift and propagation loss in the process of acoustic transmission brings the sonar image problems as low contrast, strong speckle noise, fuzzy target contour etc[1]. Low Resolution, poor image quality, less useful information and other issues, seriously affecting the underwater detection and operations. Ensuring the integrity of the partitioned area and the accuracy of the boundary is of great significance for underwater detection and also contributes to the further development of the sonar in the field of underwater target detection.

II. SONAR IMAGE TARGET DETECTION BASED ON CLUSTERING

A. Extract image texture and features based on GLCM

Gray Level Co-occurrence Matrix(GLCM) is a tool for counting image texture and features. It can reflect the information of gray in the direction of the adjacent interval and other aspects. In the direction θ of the straight line, the gray point i of a pixel, the pixel j distance d from the gray scale, and the frequency of i and j occur at the same time is the gray value of the GLCM, which can be written as (i,j) . It usually takes four directions: *0 °, 45 °, 90 °, 135 °*.

Haralick mentioned 14 texture features in [2], in which the four kinds of eigenvectors of energy, contrast, correlation and deficit are irrelevant. They are convenient to calculate and the accuracy of their classification is high. The solution formula is as follows:

$$
\begin{aligned}
S_{Energy} &= \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P(i,j,d,\theta)^{2} \\
S_{Contrast} &= \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} (i-j)^{2} P(i,j,d,\theta) \\
S_{Correlation} &= \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} \frac{ijP(i,j,d,\theta)-\mu_i\mu_j}{\sigma_i\sigma_j} \\
S_{IDM} &= \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} \frac{P(i,j,d,\theta)}{1+(i-j)^{2}}
\end{aligned}
\tag{1}
$$

When a texture feature is extracted, a square window is selected centered on the pixel to be extracted, and the gray level co-occurrence matrix is calculated in the window, and the eigenvalues in the four directions of the matrix are calculated according to (1). Slide the window to the next pixel, repeat the above steps until the full image is calculated.

B. K-means clustering

The *k*-means clustering[3] relies on the distance squared sum of each data point to the center of its decision, and divides each sample into a class according to the minimum of distance. The clusters are continually updated and the categories of each sample are re-divided until the sum of the squares of the distances from each sample to its decision center is minimized. In order to obtain the best segmentation effect, *Silhouette* (*Sil*)[4] is used to measure the clustering effect, which reflects the intraclass tightness and interclass separability of clustering results, which can be written as:

$$
Sil(i) = \frac{b(i)-a(i)}{\max(b(i),a(i))} \tag{2}
$$

Where $a(i)$ is the average non-similarity between a sample i in class C and all other samples in C, and $b(i)$ is the minimum of the average distance from sample i to all samples in each of the other classes. The greater the value of *Sil*, the better the clustering effect.

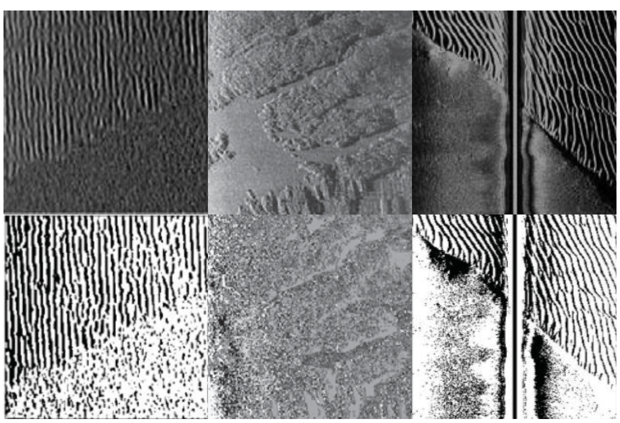
C. Specific steps of seabed segmentation

Step 1. Select the appropriate size of the square window to extract the gray level co-occurrence matrix on the sidescan sonar image, then extract texture features and normalize them, and store them in a 16-dimensional row vector.

Step 2. Cluster the row vector. The range of clusters is usually set to [2,6], and the value of k in this range are clustered in order: Select *k* initial clustering centers randomly. Then classify the samples, calculate the distance from the sample to the center, and update the center point. Determine whether the algorithm can be converged. If there is no change in the category of the sample, stop it. Calculate the average of *Sil* according to (2), then record the value of *k*, clustering centers, distance and other data. Finally,Repeat the above three steps, the number of iterations is usually set to 5. Record the average of *Sil* of the maximum clustering value and clustering results

Step 3. Compare the value of *k* between [2,6]. When the average of *Sil* is the largest, the corresponding *k* is the optimal clustering number. Similarly, the corresponding result is the best clustering result.

Step 4. Generate a pixel matrix of the same size as the original image. Assign the same gray value or the same color to the same pixel and output the current clustering result, the current cluster number (*k*), and the average of *Sil*.



(a) (b) (c)

Fig.1. Clustering segmentation on sidescan sonar images of seabed ((a) and (c)) and coral reefs (b). The clustering number (value of k) is 2,6,2. The upper row are original images, and the lower row are clustering result.

TABLE I. THE AVERAGE OF SIL OF CLUSTERING SEGMENTATION AT DIFFERENT VALUES OF K ON SIDESCAN SONAR IMAGES

k	**Average of *Sil***		
	Fig.1.(a)	Fig.1.(b)	Fig.1.(c)
2	**0.911**	0.901	**0.859**
3	0.820	0.805	0.778
4	0.829	0.850	0.808
5	0.859	0.898	0.715
6	0.845	**0.905**	0.833

III. SONAR IMAGE TARGET DETECTION BASED ON MRF

In 1984, S. Geman and D. Geman proposed the concept of Markov random field [5], which can effectively characterize the spatial information of the image and is widely used in the field of image segmentation. The random field *X* defined on the grid set *S* having the neighborhood system *N* is called a Markov Random Field (MRF).

A. FGMM observation field modeling and parameter estimation

The probability distribution of the Gaussian mixture model is:

$$f\left(y_s \mid x_s\right)=\sum_{m=1}^{K}\pi_m f\left(y_s \mid x_s=m\right) \tag{3}$$

y_s represents the observed value at S, $m=1,2,...,K$ represents the classification mark of the observation position.

$$f\left(y_s \mid x_s=m\right)=\frac{1}{\sqrt{2\pi\sigma_m^2}}\exp\left[-\frac{\left(y_s-\mu_m\right)^2}{2\sigma_m^2}\right] \tag{4}$$

π_m represents the proportion of pixels in the image labeled *m*. The mean value of the region labeled *m* is μ_m, the variance is σ_m^2, and the probability distribution X is a hidden random field, so the observed image data is an incomplete data set. EM algorithm can be used to estimate the parameters, the likelihood function is:

$$\begin{aligned}\ln(\theta \mid y) &= \ln \prod_{s\in S} P\left(y_s \mid \theta\right)\\ &= \sum_{s\in S}\ln\left[\sum_{x_s=m}^{K}\pi_m P\left(y_s \mid \mu_m,\sigma_m^2\right)\right]\end{aligned} \tag{5}$$

The E step of the algorithm is:

$$\tau_s^{(t+1)}=\frac{\pi_m^{(t)} f\left(y_s \mid \mu_m^{(t)},\left(\sigma_m^{(t)}\right)^2\right)}{\sum_{m=1}^{K}\pi_m^{(t)} f\left(y_s \mid \mu_m^{(t)},\left(\sigma_m^{(t)}\right)^2\right)} \tag{6}$$

The M step of the algorithm is:

$$\begin{aligned}&\pi_m^{(t+1)}=\frac{1}{N}\sum_{s\in S}\tau_s^{(t+1)},\ \mu_m=\frac{1}{N\pi_m^{(t)}}\sum_{s\in S}\tau_s^{(t)}y_s\\ &\left(\sigma_m^{(t+1)}\right)^2=\frac{1}{N\pi_m^{(t)}}\sum_{s\in S}\tau_s^{(t+1)}\left(y_s-\mu_m^{(t)}\right)^2\end{aligned} \tag{7}$$

Where τ is the number of iterations and *N* is the number of pixels.

B. Iteration conditions mode(ICM)

ICM is a deterministic algorithm based on local conditional probability. The method of image segmentation is to update the mark by point. Assume that each pixel y_i of the image data $y=\{y_1,y_2,...,y_n\}$ is independent of each other under the given segmentation result *x*, and the conditional distribution of y_i with respect to x depends only on its mark x_i:

$$f(y \mid x) = \prod_{i \in S} f(y_i \mid x_i) \tag{8}$$

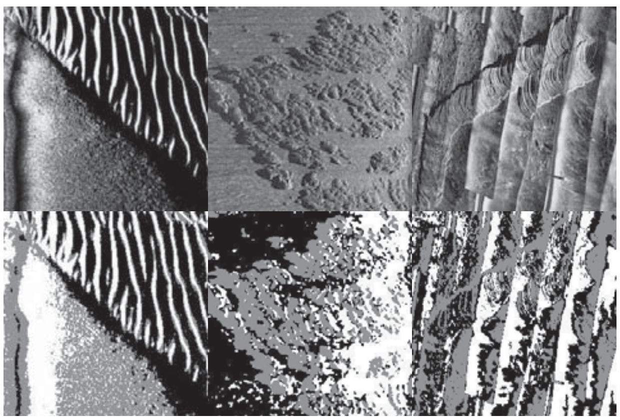
According to the Bayesian formula:

$$P(x_i \mid y, x_{N_i}) \propto f(y_i \mid x_i) P(x_i \mid x_{N_i}) \tag{9}$$

Maximized it, that is, the classification of the pixel mark:

$$\hat{x}_i = \arg\max_{x_i} P(x_i \mid y, x_{N_i}) \qquad (i = 1, 2, \cdots, n) \tag{10}$$

The ICM algorithm relies heavily on the initial segmentation results. In this paper, the *k*-means algorithm is used to initialize the image, and ICM is applied.



(a) (b) (c)

Fig.2. Segmentation results using ICM on sidescan sonar images of sand grain (a), coral reefs (b) and watercourse (c). The upper row are original images, and the lower row are segmentation results.

As can be seen from the experimental results, ICM algorithm can effectively divide the substrate in sidescan sonar images. Different substrates can be distinguished easily by using different gray values to mark different pixels.

IV. SONAR IMAGE TARGET DETECTION BASED ON LEVEL SET

The basic idea of the level set is to regard the curve on the two-dimensional plane as the zero level set of the surface function in the three-dimensional space $\Phi(x,y)=0$. This kind of tracking transforms the evolution of the tracking curve into the problem of solving the numerical partial differential equation, thus avoiding the problem of parameterization. The closed curve hidden in the level set function can also be evolved by updating the level set function. This function can remain active even when the closed curve is split or merged [6].

The sidescan sonar image consists of the target highlight area, the shadow area and the background area, so the three types of segmentation are appropriate. The principle of the piecewise constant in the C-V multiphase level set method[7] is to describe a 2^a-phase with a set of horizontal sets, and the number of horizontal functions decreases from a to $loga$, so that it is possible to effectively avoid occurrence of overlapping or "vacuum" in the area [8].

According to the heaviside function:

$$\{(x,y) \mid H(\Phi(x,y)) \in H(\Phi(\Omega))\} \tag{11}$$

The energy function is:

$$\begin{aligned} F(\Phi, c) = & \iint_\Omega (u_0 - c_{11})^2 H(\phi_1) H(\phi_2) dxdy \\ & + \iint_\Omega (u_0 - c_{10})^2 H(\phi_1)(1 - H(\phi_2)) dxdy \\ & + \iint_\Omega (u_0 - c_{01})^2 (1 - H(\phi_1)) H(\phi_2) dxdy \\ & + \iint_\Omega (u_0 - c_{00})^2 (1 - H(\phi_1))(1 - H(\phi_2)) dxdy \\ & + v \iint_\Omega |\nabla H(\phi_1)| dxdy + v \iint_\Omega |\nabla H(\phi_2)| dxdy \end{aligned} \tag{12}$$

The image to be split is:

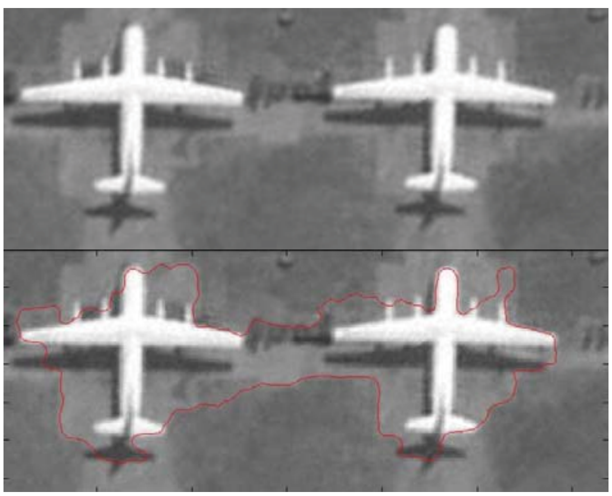
$$\begin{aligned} u = & c_{11} H(\phi_1) H(\phi_2) + c_{10} H(\phi_1)(1 - H(\phi_2)) \\ & + c_{01}(1 - H(\phi_1)) H(\phi_2) \\ & + c_{00}(1 - H(\phi_1))(1 - H(\phi_2)) \end{aligned} \tag{13}$$

Minimized the energy function, the parameters are estimated as:

$$\begin{aligned} c_{11} &= \frac{\int_\Omega I \cdot H_\varepsilon(\phi_1) \cdot H_\varepsilon(\phi_2) dxdy}{\int_\Omega H_\varepsilon(\phi_1) \cdot H_\varepsilon(\phi_2) dxdy} \\ c_{10} &= \frac{\int_\Omega I \cdot H_\varepsilon(\phi_1)(1 - H_\varepsilon(\phi_2)) dxdy}{\int_\Omega H_\varepsilon(\phi_1) \cdot (1 - H_\varepsilon(\phi_2)) dxdy} \\ c_{01} &= \frac{\int_\Omega I \cdot (1 - H_\varepsilon(\phi_1)) \cdot H_\varepsilon(\phi_2) dxdy}{\int_\Omega (1 - H_\varepsilon(\phi_1)) \cdot H_\varepsilon(\phi_2) dxdy} \\ c_{11} &= \frac{\int_\Omega I \cdot (1 - H_\varepsilon(\phi_1)) \cdot (1 - H_\varepsilon(\phi_2)) dxdy}{\int_\Omega (1 - H_\varepsilon(\phi_1)) \cdot (1 - H_\varepsilon(\phi_2)) dxdy} \end{aligned} \tag{14}$$

During the evolution of the level set, there will be some irregular phenomenon, which makes the evolution curve deviate from the real contour of the target, and the level set function gradually away from the distance function. A common solution is to continually initialize the level set function during evolution, But this method is not only time-consuming, the amount of calculation is also great, the problem about when to initialize and how to initialize needs to be more artificial, it is difficult to achieve program automation. To solve this problem, Professor Li Chunming [9] proposed to add the penalty items and data items[10] in the model, to avoid the level set function away from the signed distance function to ensure that the contour curve is close to the real target contour. It solves the problem of initializing the level set function in the evolution process continually, simplifies the computation and improves the efficiency, and the detection effect is superior to the traditional method.

(a) (b) (c)



(d)

Fig.3. Level set segmentation results. Level set segmentation on images of two cells (a), galaxy (b) and two planes (d), sidescan sonar image of crashed plane (c). The upper row are original images, and the lower row are segmentation results. The Number of iterations is 300,500,1500,800.

V. CONCLUSION

In this paper, the classical k-means clustering algorithm is adopted, which is efficient and suitable for large-scale data sets. Combined with the initial segmentation results of *k*-means clustering, the final segmentation is achieved by updating the markings through the iterative conditional model (ICM). The principle and defects of the C-V model segmentation constant four-phase in the level set are discussed in detail, by studying the improved algorithm proposed by Professor Li Chunming, we can extract the outline of the target in the image accurately, which have good adaptability.

ACKNOWLEDGMENT

This paper is supported by the National Natural Science Foundation of China (Grant No.61671202, 61573128, 61273170), the National Key Research and Development Program of China (Grant No.2016YFC0401606), the Fundamental Research Funds for the Central Universities of Ministry of Education of China (Grant No.2015B25214), and the Fundamental Research Funds for the Central Public Research Institutes of China (Grant No.Y716019).

REFERENCES

[1] H.Z. Teng, X.M. Yan, S.Q. Li, et al. Real time enhancement technique of sidescan sonar images. Marine Charting, 2004, 24(2): 47-49.

[2] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classication. IEEE Transaction on Systems, Man, and Cybernetics, 1973, 3(6): 610-621.

[3] Zaki M J, Meira W. Data Mining and Analysis: Fundamental Concepts and Algorithms. New York, NY: Cambridge University Press, 2014.

[4] Campello R J, Hruschka, E R. A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets and Systems, 2006, 157(21), 2858-2875.

[5] S. Geman, D. Geman. Stochastic relaxation Gibbs distribution, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1984, 16: 721-741.

[6] J.N. Chen. Research on Image Segmentation Based on Level Set Method. Yanshan University,2007 :10-20.

[7] T. Chan and L. Vese, “Active contours without edges,” *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[8] S.X. Dong. Medical Image Segmentation Based on Vese - Chan Multiphase Level Set Method. Qingdao University, 2007:37-42 .

[9] C.M. Li, R. Huang, Z. Ding, et al. A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI. *IEEE Trans. Image Process.*, 2011.

[10] C. Li, C. Kao, J.C. Gore, and Z. Ding, “Minimization of region-scalable fitting energy for image segmentation,” *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.