Self-Organizing Maps Approach to Object Localization in Sonar Imagery

Amornrit Puttipipatkajorn, Bruno Jouvencel

Self-Organizing Maps approach to object localization in sonar imagery

Puttipipatkajorn Amornrit, Bruno Jouvencel

LIRMM – University of Montpellier II
161 rue Ada, 34392, Montpellier, Cedex 5, France
putti@lirmm.fr, jouvencel@lirmm.fr

Abstract

The Self-Organizing Map is well-known as the unsupervised classification method. It is employed as classifier in various applications such as image segmentation. The main purpose of this paper is to identify and detect an object of interest on side scan sonar image. This work is performed by two steps. The first one is to split an image into regions of uniform texture using the Gray Level Co-occurrence Matrix Method (GLCM) which is widely used in texture segmentation application. The last one address the unsupervised learning method based on the Artificial Neural Networks (Self-Organizing Map or SOM) used for determining the comparative model of object of interest from an image. To increase the performance of SOM, we propose a penalty function based on data histogram visualization. After a brief review of both techniques (GLCM and SOM), we present our method and some results from several experiments on the real world data set.

1. Introduction

Besides the human interpretation, the high-resolution Side Scan Sonar seems to be the advanced tool for analyzing the sea floor. Three kinds of regions can be visualized: echo, shadow and sea bottom reverberation. The echo information is caused by the reflection of acoustic wave from the object while the shadow area corresponds to a lack of acoustic reverberation behind the object and the remaining is the sea-bottom reverberation area. The only available type of sonar image is the gray level of the pixels corresponding to the acoustic reflectance. Unfortunately, we can not recognize and classify the objects base on single feature, consequently several methods are proposed in order to obtain more in the aspect of second-order information. Many studies do about the performance of various families of computational methods for texture feature extraction, for instances, the 2-dimensions of FFT, the Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length and etc[10]. In addition, a comparative study from several methods show that the GLCM is a excellent statistical tool for extracting second-order texture information from image.

The co-occurrence matrix is used as an estimator of the joint probability density function of gray-level pairs in an image. The matrix is in general symmetric and, when normalized, element values are bounded by [0,1], and the sum of all element value equal to 1. Features extracted from this matrix are so called the second-order statistical feature, for instance, energy, entropy, inverse difference moment, and etc.

The next section of this paper concerns with clustering algorithms based on the Self-Organizing Map (SOM) [1]. This method is frequently employed in various applications such as data mining [2], image segmentation [7] and also pattern recognition system. The SOM is a neural network algorithm based on unsupervised learning. It is an efficient tool for visualizing the multidimensional numerical data. It represents high-dimensional data into low-dimensional grid in 1D or 2D. Several methods to visualize clustering base on the SOM can be found in the literature. The most widely used methods for visualizing the cluster structure of the SOM are distance matrix technique[2][4], especially the unified distance matrix or U-matrix. Unlike U-matrix, data histogram visualization method is to show the number of hits in each map unit. This information can be utilized in clustering the SOM by using zero-hit units to indicate boundary of cluster[2]. The main purpose of this paper is to determine a comparative model of object of interest based on the data histogram visualization method. To find the object location, the penalty function is formed and the object of interest is found when its penalty value is nearest zero.

2. Seabed Recognition and Classification System for detecting an object of interest

The basic seabed recognition and classification system is composed of three stages: pre-processing, feature extraction and classification (Figure 1). The role of the pre-processing module is to remove noise and normalize the pattern.

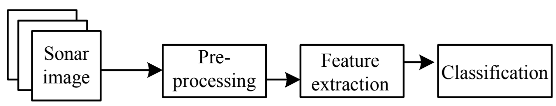


Figure 1.

For this stage, the median filter and histogram equalization are employed. The next stage which will be discussed in the next section is the Gray Level Co-occurrence Matrix (GLCM). The GLCM is used to extract different second-order features of image from the co-occurrence matrix. Finally, these features are trained by the SOM to partition the feature space and then to create a comparative model of object of interest in this case.

3. The Gray Level Co-occurrence Matrix

The good texture feature extraction method should be capable of identifying the major groups of seabed patterns based on their prominent features to give the best information for texture classification. In the scientific literature, one of the most well-known and wildly used techniques is the G*ray Level Co-occurrence Matrix* (GLCM). The GLCM is based on the estimation of the second-order joint conditional probability density function, $P(i,j/d,\delta)$, derived from co-occurrence matrix.

3.1. Co-occurrence matrix

The co-occurrence matrix, *P*, represents the repeated occurrence of pairs of pixels (*i,j*) going from gray level *i* to gray level *j* through distance *d* along direction δ. Let $I_x=\{1,2,...,N_x\}$ and $I_y=\{1,2,...,N_y\}$ be the **X** and **Y** spatial domains, where $I_x \times I_y$ is the set of resolution of square image, and the digital image I contain a finite number of gray-level value $g \in G\{1,2,...,N_g\}$ for every pixels, formally $I: I_x \times I_y \rightarrow G$. Let the distance *d* is the distance between two-pixel positions (x_1,y_1) and (x_2,y_2), which indicated by angular angle δ. The matrix, P, is $N_g \times N_g$ square matrix, where N_g -1 is gray value in the image.

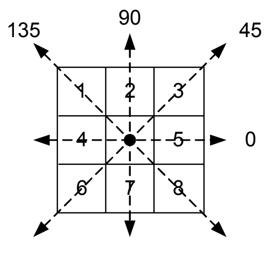


Figure 2.

The co-occurrence matrix allows us to derive four matrix for each given distance: *P*(0,*d*), *P*(45,*d*), *P*(90,*d*) and *P*(135,*d*), as indicated in Figure 2.

3.2 Statistical texture description functions

Prior to calculate the statistical texture descriptors, the matrices are normalized. They approximate the joint probability densities of the co-occurrence gray level.

$$P(i,j) = P(i,j)/N \quad \text{where } N = \sum_i \sum_j P(i,j) \tag{1}$$

From the normalized co-occurrence matrix, a set of textural features is extracted. In our experiments, the most relevant features used are listed below.

Energy:

$$\text{F1} = \sum_i \sum_j P(i,j)^2 \tag{2}$$

Entropy:

$$\text{F2} = -\sum_i \sum_j P(i,j) \log P(i,j) \tag{3}$$

Maximum probability:

$$\text{F3} = \max\{P(i,j)\} \tag{4}$$

Inverse difference moment*:*

$$\text{F4} = \sum_i \sum_j \frac{P(i,j)}{1+(i-j)^2} \tag{5}$$

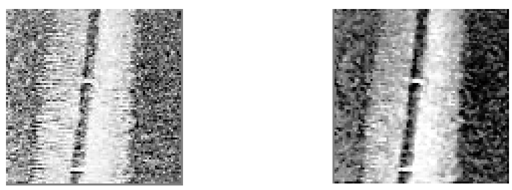
Contrast:

$$\text{F5} = \sum_i \sum_j (i-j)^2 P(i,j) \tag{6}$$

Homogeneity:

$$\text{F6} = \sum_i \sum_j \frac{P(i,j)}{1+|i-j|} \tag{7}$$

In this case, we obtain texture feature vectors, $F=\{F_1,F_2,...,F_6\}$. Each element contains information of image texture calculating from statistical description functions above. For instance, the figure 3 show the images are extracted from real sonar image used in experiment with six statistical descriptors.



(a) (b)

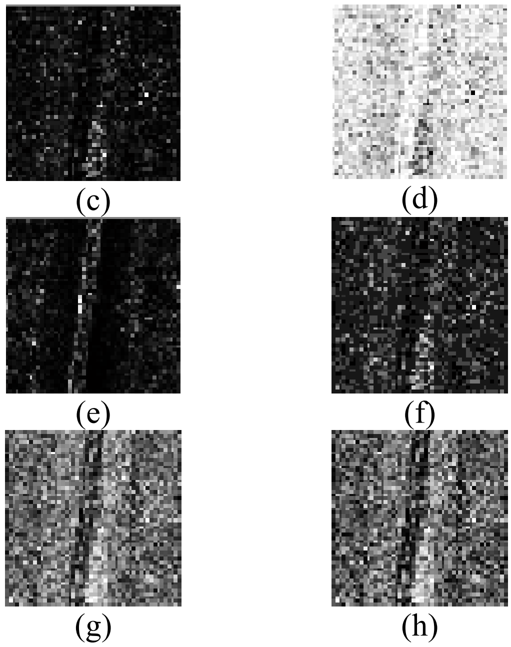


Figure 3. (a) Original image. (b) Image after passing through the median filter and histogram equalization. (c) Energy of image (d) Entropy of image (e) Contrast of image (f) Maximum of image (g) Homogeneity of image (h) Inverse difference moment of image

4. Self Organizing Map (SOM)

One of the most popular of the *Artificial Neural Network* (ANN), the self-organizing map (SOM), is the best one for pattern recognition and classification task. It belongs to the category of unsupervised learning neural networks. The SOM have only two layers of neurons, an input layer and a competitive layer (figure 4). Each node in the input layer is connected to every node in the competitive layer. The nodes in the competitive layer may also be connected to each other in the aspect of various models of connection, such as squared neighboring connection.

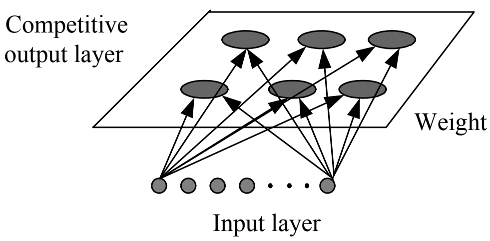


Figure 4.

The model of SOM used in our application is a two-dimensional array of k nodes. Each neuron k is represented by an n-dimensional vector $m_k = [m_{k1},\ldots,m_{kn}]$, where n is the dimension of the input space. On each training step, a data sample x is randomly selected and the best-matching unit (BMU or m_c) is found on the map unit:

$$\|x-m_c\| = \min_k\{\|x-m_k\|\} \tag{8}$$

And then, the vector m_c and its neighbors on the grid are updated by closing to the sample vector:

$$m_k = m_k + \alpha(t)h_{ck}(t)(x-m_k) \tag{9}$$

where t denotes time, $\alpha(t)$ is learning rate and $h_{ck}(t)$ is a neighborhood kernel centered on the winner unit c:

$$h_{ck}(t) = \exp\left(-\frac{\|r_c-r_k\|^2}{2\sigma^2(t)}\right) \tag{10}$$

and

$$\alpha(t) = \frac{\alpha_0}{1+100t/T} \tag{11}$$

where $\|r_c-r_k\|$ is distance between map units of neurons c and k on the SOM grid. In equation (11), α_0 denotes initial learning rate and T is the total iterative time. Both learning rate function $\alpha(t)$ and neighborhood kernel radius decrease monotonically with time.

During the iterative training, the SOM adapt to input data set in such a way that the model vectors which belong to units close to each other on the map unit, are also close to each other in the data space.

4.1. Data histogram visualization method

The data histogram visualization method is to display the number of hits in each map unit. It means that each unit of map unit belongs to a number of the best-matching units of any given input vectors. For instance, if we have 20 input vectors and 2×2 map unit, the first unit possesses 6 the best-matching units, the other ones have 4, 2 and 8 the best-matching units respectively. Finally, a matrix of map unit is formed , shown in Figure 5, and the normalization of this matrix is employed in the next section.

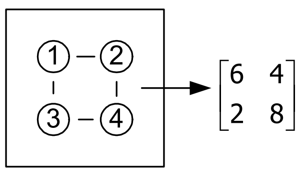


Figure 5. The matrix is derived from the best-matching unit of given inputs

4.2 Penalty function

In this section, we address the penalty function based on the data histogram visualization method. In this paper, we do not directly use the SOM to classify different textures of image, but it is used for evaluating a comparative model of interested object. It means that the SOM model generally contains a number of categories of given input. These categories can be perhaps clustered by using the well-known method such as the U-matrix. But, in this paper, this model is used for evaluating the penalty function as follows:

$$E^k(i,j)=\sum_{i=1}^{n}\sum_{j=1}^{m} I_w(i,j)\otimes\left\{I^k(i,j)-I_w(i,j)\right\} \tag{12}$$

Where, I_w is the SOM model or the comparative model of object of interest, I^k is the one of tested sample, k denotes the index of tested sample and $\otimes$ is product of matrix in term by term. The best-matching sample is found when the value of E is the nearest zero.

5. Methods

The system introduced in this work proceeds in two phase: a training phase and a testing phase. The training phase has a set of labeled images of object of interest or so called training window. Each contains a single object of interest and its environment with a standard size and orientation, the structure of training window shown in Figure 6. During the training phase, the training window is firstly trained by SOM network in order to obtain a comparative model of object of interest. In case of two-dimensional map units, this model is the matrix of probability density such as I_w .

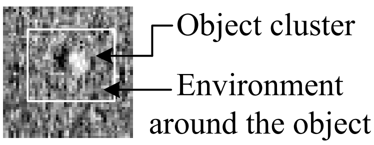


Figure 6. The structure of training window

By means of the trained SOM model, the matrix of probability density of sliding-window, I^k, or testing window is calculated. During this phase, the trained window, I_w, is compared to tested window, I^k, sliding through the image from left to right and top to bottom (figure 7). To find the object location, the penalty value E^k for every k is calculated. The object of interest is located on the tested window k which has the penalty value nearest zero.

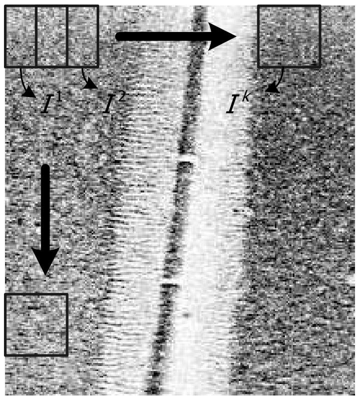


Figure 7. Real SSS image of pipeline

6. Experimental results

The experiment follows the seabed recognition block diagram illustrated in section 2. In training step, the training window I_w containing the object of interest is cut from the testing image I (see Figure 8.1). A size of 5×5 units of SOM grid is selected for the best result from a number of experiments. The first experiment uses the testing image of 50×50 pixels and the training window of 15×15 pixels. The result is shown in figure as follows.

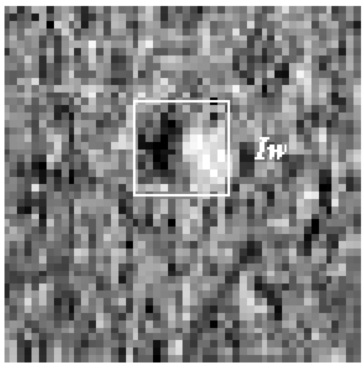


Figure 8.1. The testing image I of 50×50 pixels and training window I_w of 15×15 pixels

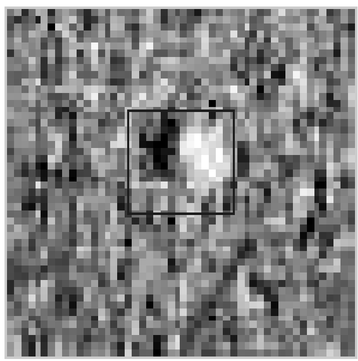


Figure 8.2. The object found by the SOM model

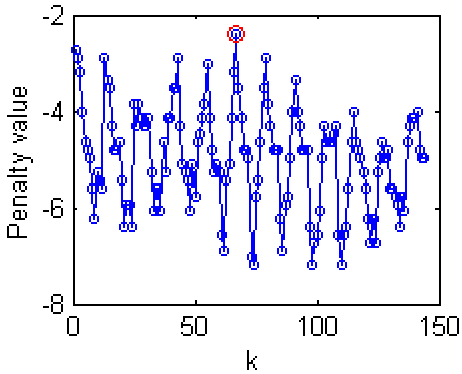


Figure 8.3. The only one best-matching window indicated by the value of k at peak of curve.

The second experiment not only tries to detect the object, but also attempts to find the object which is bigger than the sliding-window as pipeline image shown in Figure 8.4, 8.5 and 8.6 respectively.

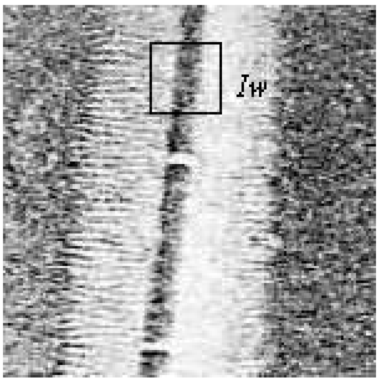


Figure 8.4. The pipeline sonar image I of 200×200 pixels and training window I_w of 30×30 pixels

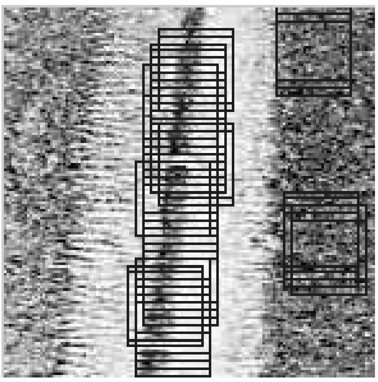


Figure 8.5. The pipeline found by SOM model in tracking mode.

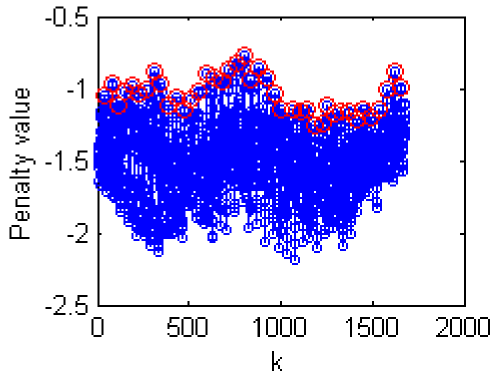


Figure 8.6. The best-matching windows found on each line of image in tracking mode.

7. Conclusion

This paper proposes a technique for approaching to object localization using a comparative model based on the SOM. The experiments show that the SOM performs well in finding an object of interest in real world sonar image. The advantage of this technique is simple and robust. However this method has the high computational time due to co-occurrence matrix calculation. In addition, this technique can not precisely solve the rotation problem of object, because size of sliding-window is always fixed to one of training window in specific direction.

For future work, we will attempt to improve this technique for finding the object lying on an image in different directions and also penalty function to identify more precisely the best-matching window.

References

[1] T. Kohonen. “Self-Organizing Maps”, 3rd edition. Springer-Verlag, Berlin, 2001.

[2] Juha Vesanto and Esa Alhoniemi. “Clustering of the Self-Organizing Map”. IEEE Transaction on Neural Network 11 (2000) 586-600.

[3] Elias Pampalk, Andreas Rauber and Dieter Merkl. “Using smoothed data histograms for cluster visualization in Self-Organizing Maps”. Technical Report OeFAI-TR-2002-29, extended version published in Proceeding of the International Conference on Artificial Neural Networks, Springer Lecture Notes in Computer Science, Madrid, Spain, 2002.

[4] Juha Vesanto. “SOM-Based Data Visualization Methods”. Laboratory of Computer and Information Science, Helsinki University of Technology, Finland, November 1999.

[5] Arthur Flexer. “On the use of Self-Organizing Maps for clustering and visualization”. Intelligent Data Analysis 5 (2001) 373-384, IOS Press.

[6] Vladan M Babovic and Gopakumar R. “Seabed Recognition using neural networks”. D2K Technical Report, 0399-1, March, 1999.

[7] K.C. Yao, M. Mignotte, C. Collet, P. Galerne and G. Burel. “Unsupervised segmentation using a Self-

Organizing Map and a noise model estimation in sonar imagery". Pattern Recognition 33 (2000) 1575-1584.
[8] N.Pican, E. Trucco, M. Ross, D.M. Lane, Y. Petillot and I. Tena Ruiz. "Texture Analysis for Seabed Classification: Co-occurrence Matrices vs Self-Organizing Maps". Ocean Systems Laboratory, Department of Computing and Electrical Engineering, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK.
[9] Ross F. Walker, Paul Jackway and I.D. Longstaff. "Improving Co-occurrence Matrix Feature Discrimination". Proceedings of DICTA'95, The 3rd Conference on Digital Image Computing: Techniques and Applications, 6-8th December, 1995, pp643-648.
[10] W. Kenneth Stewart, Min Jiang and Martin Marra. "A Neural Network Approach to Classification of Side scan Sonar Imagery from a Mid-ocean Ridge Area". IEEE journal of oceanic engineering, VOL. 19, NO. 2, April 1994.