VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

MARINE VESSEL RECOGNITION BY ACOUSTIC SIGNATURE

N. Leal[1], E. Leal[1] and G. Sanchez[2]
[1]Faculty of Engineering, Universidad Autónoma del Caribe, Barranquilla, Colombia
[2]Faculty of Engineering, Universidad del Magdalena, Santa Marta, Colombia
E-Mail: nleal@uac.edu.co

ABSTRACT

Automatic recognition of an acoustic signature in underwater environments is an important and active field with multiple applications, one of which is vessel recognition. When a vessel moves through the sea, its engine and the cavitation generated by its propellers produce an acoustic wave of unique characteristics that allow for its individual identification. The problem of identification involves several variables, such as ambient noise, biological noise, and even noise produced by its own machinery, which means that the signal produced, is complex to treat. This paper presents a method based on Fourier transform and digital signal processing to extract a set of features allowing for automatic ship classification (by type). Computational intelligence techniques such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) are used for the classification stage. Results showed that the vessel recognition system has accuracy close to 92%.

Keywords: acoustic signature, vessel recognition, neural networks, support vector machines, fourier transform.

INTRODUCTION

Underwater acoustic signal identification is an important field in areas dealing with coastline surveillance, security in the seas, war strategies and tactics, among others. Acoustic signals produced by a ship are caused mainly by its machinery and the cavitation produced by the propellers, combined with the marine environment, which can include noise from other ships, biological noise such as that produced by dolphins and whales, and even rainfall. Such signals, however, are unique for each type of vessel, and therefore can be used for identification and tracking purposes.

Underwater acoustic signals are detected, recorded and classified by using sonars, which are an important military tool that helps determine if a ship is a friend or enemy. The task of ship classification by sonar is sometimes carried out by a sonar operator that must perform a series of analysis; therefore, the necessary time for vessel recognition increases considerably and depends on operators' skills.

Recent works aimed at developing automatic identification systems of acoustic signals approach the issue as a stationary process, given its variability and short duration [1-3].

Initially, techniques used for automatic ship detection have been based on the extraction of features in the frequency domain by using Fourier fast transform (FFT) [4-10]. An example of this is found in [4], where an omnidirectional hydrophone is used; the power spectral density is extracted from the signal, and the information is used to train a feed forward neural network. Similarly, in [8] sonar information is processed by extracting the amplitude spectrum and transforming it by means of principal component analysis. The objective is to separate the sets or types of ships to be classified, and such information is used as input for a neural network serving as non-linear classifier. With a similar approach, [9] proposes a process that uses neural networks with supervised and unsupervised learning, comparing their performance given the same spectral information as input.

Other works make use of an auto-regressive model. In [11], spectral information is used to build an ARMA (Auto Regressive Moving Average) model. In [12], an autoregressive model is created, from which the poles are extracted periodically, in order to differentiate the source emitting the noise signal to finally classify the different signals by means of statistical classifiers.

Given the results of the application of Fourier transform, interest has broadened to include Wavelet transform [7]. In [7], two pre-processing algorithms are used: the first one, known as Two-Pass Split-Window (TPSW), separates the windows in two steps, extracting information about mean power spectral density from input data; the second one uses Wavelet transform with the same data. For the classification stage, four different classifiers are used, so that a comparative analysis among them is carried out.

Neural networks have a particular and special place in the classifying stage in many of the proposed strategies [8-10], [13-15]. In [15] a Kohonen neural network was used with smoothed spectral data and the differences between components of *k* order as input. Other works with neural networks include [16-19].

Different approaches can be found in [20], which uses fractal techniques for acoustic signature identification and classification.

Two different schemes for the identification of acoustic signatures are addressed in [21]; in both, feature extraction is carried out by using Mel-frequency cepstral coefficients and Linear Predictive Coding (LPC) derived from cepstral coefficients, which have been broadly used for speech recognition. In [22], a robust algorithm is presented to detect the arrival of certain types of boats when there is background noise. The algorithm performs an analysis of the acoustic signature against an existing database of acoustic signals recorded and processed. In [23], the problem of noise in acoustic signals is analyzed, not from the statistical viewpoint, as static noise (white noise), but from the viewpoint of fractal analysis.

This article presents an underwater acoustic signal classifier that allows for the automatic identification of vessel types. Two classifiers were implemented to carry out the classification of the acoustic signal emitted by vessels: a back propagation neural network and a Support Vector Machine. The main feature of the method proposed in this paper, that makes it different from others found in the literature, is the processing and extraction of information from the signal. Information extraction for signal processing is based on processing techniques involving spectral analysis, Gaussian windowing, estimation and reduction of the noise contained in the signal by using Fourier transform.

METHODOLOGY

The stages of the proposed method include a processing stage, followed by a feature extraction stage, and finally vessel classification. The procedure is shown in Figure-1.

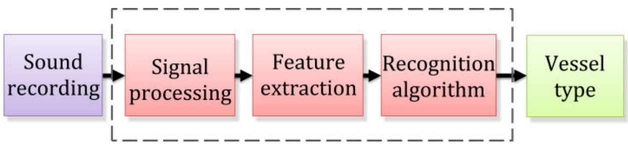


Figure-1. Block diagram of the system identification.

Signal pre-processing

The use of Fourier discrete transform to analyze the noise signal coming from a vessel is inadequate due to the complexity of the signal. When dealing with a signal of finite length, the output spectrum shows spurious frequency bands around the fundamental frequency hindering differentiation. Therefore, it is useful to determine the instant frequencies of the signal at different points in time, providing a dynamic understanding of signal features regardless of its length in time [6].

The signal must be split into smaller samples in order to carry out proper frequency analysis. One option is to multiply the signal through a straight pulse; the equivalence on frequency of this pulse signal, however, gives as a result a Sine function capable of adding detrimental information to the spectrum.

This work proposes the use of a windowing of the form$f(x) = ae^{\frac{-(x-b)^2}{c^2}}$, for any real constant$a > 0, b, c$. If$c^2 = 2$the result of the function is an eigenvalue of Fourier transform, that is, the result of Fourier transform of a signal multiplied by a Gaussian window is another function multiple of the original signal.

Feature extraction

In order to extract the relevant features of the processed signal, three methods have been applied: peak detection, energy analysis, and smoothed signal analysis. The feature vector is formed by the value of the highest peak of the original signal, with eight energy values corresponding to the energy of the first eight 500-Hz segments and the central frequency of the smoothed signal (see Figure-2).

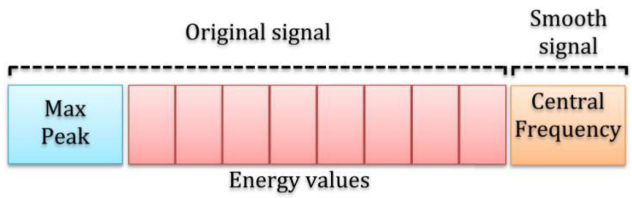


Figure-2. Signal features vector.

a) Peak detection

Peaks provide information on the signal's topological features and indicate special behaviors in the phenomenon they represent; for instance, depending on the scenario, they can represent highest demand, bandwidth increase, price increase, consumption increase, among others. In this context, a peak is defined as high values with a strong tendency rapidly followed by strong drops, generating a narrow base [24]. Algorithms for peak detection have been developed in different areas: bioinformatics, spectrometry, astrophysics, image processing, among others. In this work, we follow the procedure described in [24] for peak detection.

Let $T = x_1, x_2, \ldots, x_N$be an univariate, uniformly-sampled signal containing N values, and x_i be the i-th value of T; an S function, or peak function, is estimated, which associates a positive $S(i, x_i, T)$ value to the i-th value of T. Thus, point x_i is a peak if $S(i, x_i, T) \geq \beta$, where β is a predefined threshold. Where S is defined as follows:

$$S(k, i, x_i, T) = \frac{\left(x_i - \frac{\sum_{j=1}^{k} x_{i-j}}{k}\right) + \left(x_i - \frac{\sum_{j=1}^{k} x_{i+j}}{k}\right)}{2} \quad (1)$$

where k is a positive integer that determines the size of the neighborhood of points coming before and after each x_i value.

VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

b) Energy analysis

Energy estimation for short segments is a simple and effective classification parameter for acoustic signals with or without voices [25]. In general, energy estimation is defined as the sum of the squares of the amplitudes froma segment. In order to characterize the acoustic signal, and since most of the information is found in the band of lower frequencies at 4 kHz, the spectrum was divided into eight 500-Hz segments. In this way, the energy of each individual segment is calculated to form a vector which is a discrete representation of the signal frequency information and comprises a set of individual features.

For a discrete signal with N samples T(N), the energy of short segments measured at point *n* may be identified in three ways [26], [27]:

$$E = \sum_{i=1}^{N} T(i)^2 \tag{2}$$

$$E = \sum_{i=1}^{N} \log T(i)^2 \tag{3}$$

$$E = \sum_{i=1}^{N} |T(i)| \tag{4}$$

In this work, the traditional definition established by equation (2) was adopted.

c) Smoothing

Smoothing allows obtaining an approximate representation that keeps the general structure and the main features of the original signal as well as the tendency and the cyclical components while avoiding random variations. In general, taxonomy of smoothing methods comprises two groups: average methods and exponential methods. The drawback of average-based methods in general is assuming that there is no tendency in the data. Because of this, in this work, the exponential method to smooth the original signal was used, defined in this way:

$$\hat{x}_{i+1} = \alpha x_i + (1-\alpha) s_i \tag{5}$$

where, $\hat{x}_i$is the smoothed value, α is the smoothing constant $0 < \alpha < 1$, the value x_i is the i-th value of the data set and s_i the last smoothed value.

Once the smoothed version is obtained, a single value is adopted, corresponding to the maximum point of the smoothed function which coincides with the central frequency of the spectrum.

Classification algorithm

Once pre-processing is done and after extracting the features, they are entered into a classifier, which will allow identifying the type of vessel. Classification techniques shall be robust regarding the presence of ambient noise, the noise introduced by the capturing device, the capturing depth of the signal, water temperature changes and the noise produced by biological organisms, among others.

In this work, classification techniques have been selected by means of Artificial Neural Networks (ANN) and Support Vector Machines (SVM). These techniques allow classifying non-linear and noisy data.

Back propagation artificial neural network

Neural networks with backward propagation or *back propagation* use hyperplanes for complex region separation in the n-dimensional space, by assigning each input pattern to a determined region. For a three-layer perceptron, the input units distribute the inputs to the hidden layer units. Theequationsthat describe this distribution are:

$$x_{hj} = \sum_{i=1}^{N_I} w_{ij} x_i \tag{6}$$

$$y_{hj} = f(x_{hj}) \tag{7}$$

$$x_{ok} = \sum_{j=1}^{N_j} w_{kj} x_{hj} \tag{8}$$

$$y_{ok} = f(x_{ok}) \tag{9}$$

Where x_i is the value of the *i-th* input unit, w_{ij} is the associated weight between the *j-th* hidden neuron and the *i-th* input unit. N_I is the total number of inputs, x_{hj} is the total input to the *j-th* hidden neuron, $f((\cdot))$is the transference function, y_{hj} is the output value of the *j-th* hidden neuron, w_{kj} is the associated weight between the *k-th* output neuron and the *j-th* hidden neuron. N_j is the total number of hidden neurons, x_{ok} is the total input at the *k-th* output neuron, and y_{ok} is the output value of the *k-th* output neuron. Typically, the transition function is a sigmoid function, which can be unipolar and continuous (see Equation 10) or bipolar continuous (see Equation 11).

$$f(x) = \frac{1}{1+e^{-x}} \tag{10}$$

$$f(x) = \frac{1-e^{-x}}{1+e^{-x}} \tag{11}$$

The total network error, denoted by *E*, is defined as:

$$E = \frac{1}{2}\sum_{c=1}^{N_c}\sum_{k=1}^{N_k}\left(y_{ok,c} - t_{k,c}\right)^2 \tag{12}$$

where $y_{ok,c}$ is the current value of the *k-th* output neuron for the *c-th* input pattern, $t_{k,c}$ is the desired value of the *k-th* output neuron for the *c-th* input pattern, N_k is the total number of output neurons, N_c is the total number of input patterns. To minimize the error E regarding the weights, the gradient descent method is applied, the weights are updated according to:

$$\Delta w_{kj}(s+1) = -\eta \sum_{c=1}^{N_c} \frac{\partial E}{\partial w_{kj}} + \beta \Delta w_{kj}(s) \quad (13)$$

$$\Delta w_{ji}(s+1) = -\eta \sum_{c=1}^{N_c} \frac{\partial E}{\partial w_{ji}} + \beta \Delta w_{ji}(s) \quad (14)$$

$$\frac{\partial E}{\partial w_{kj}} = (y_{ok} - t_k) y_{ok} (1 - y_{ok}) y_{hj} \quad (15)$$

where$s$ represents the current iteration of the evaluation and $s+1$ is the following iteration of the evaluation, η is the learning rate, and β is the moment factor.

Support vector machine classifier

Support Vector Machines are used for classification and regression. SVMs use machine learning theories to maximize the prediction accuracy while avoiding data overfitting.

Let us consider a linearly separable training data set, composed by N vectors called patterns, of the type $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$where $x_i \in R^n; i = 1, \dots, l$. Each scalar $y_i \in \{+1, -1\}$. Since it is a separable problem, there is a hyperplane defined by its vector. An SVM is a binary classifier that assigns a y_i tag to the x input vector of each class. The main idea of SVMs is to build a decision hyperplane between two linearly-separable classes and to maximize the distance interval between the two classes (see Figure-3).

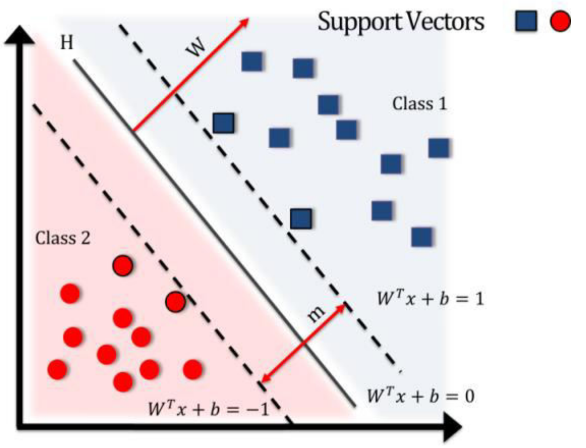


Figure-3. SVM classification scheme, H is the classification hyperplane, W is the normal vector to the hyperplane, m is the minimum distance between positive and negative hyperplanes.

The mathematical model of the given classification hyperplane $m = \frac{2}{\|w\|}$ is to maximize $\frac{2}{\|w\|}$. In a mathematical sense, this is equivalent to minimizing $\frac{1}{2}\|w\|^2$ and is convenient from a computational standpoint. To this end, Lagrange multipliers are used, in such a way that a Lagrangian function $\tilde{L}(\alpha)$ must be built (see equation 16).

$$\tilde{L}(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i^T \cdot x_j) \quad (16)$$

Finally, W can be calculated by using the terms α, $w = \sum_{i=1}^{n} \alpha_i y_i x_i$, with $b = \frac{-1}{2} w \cdot [X_r + X_s]$, where X_r and X_s are a pair of support vectors, one of each class.

Ideally, the model should produce a hyperplane that separates completely the data from the classes, grouped in two categories. However, a separation that completely meets these criteria does not always happen. If it does, the model cannot be generalized to other data and will correspond to overfitting. Thus, the model includes the parameter C which controls the compensation between training errors and rigid margins, which creates a soft margin that admits some classification errors. The problem of the classification hyperplane is reformulated in the following way:

$$\begin{aligned} &min\left\{\frac{1}{2} W \cdot W + c \sum_{i=1}^{n} \xi_i\right\}; \\ &\text{subject to } y_i(w^T x_i + b) \geq 1 - \xi_i; \quad (17) \\ &i = 1, \dots, n; \ \xi_i \geq 0, i = 1, \dots, n \end{aligned}$$

If the data from the classes do not allow to do a linear separation in the input space, we consider the input vector mapping **x** in a greater dimension space R^m, called feature space$T$, that is provided with a scalar product and which allows to do a linear separation in R^m; but it represents a non-linear hyperplane R^n, by using the Lagrange multipliers we have

$$\tilde{L}(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\emptyset(x_i) \cdot \emptyset(x_j)\right) \quad (18)$$

The calculation of $\emptyset(x_i) \cdot \emptyset(x_j)$ is computationally expensive because m is much greater than n. The solution is to use the *kernel* functions, which cause that the scalar product appears in the space of features T but by making the calculation in the input space, which causes that it is not necessary to know $\emptyset(x_i)$since$k(x_i, x_j) = \emptyset(x_i) \cdot \emptyset(x_j)$. Therefore, $\tilde{L}(\alpha)$is defined as:

$$\tilde{L}(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (19)$$

The most common *Kernel* functions are:

ARPN Journal of Engineering and Applied Sciences

$$K(x_i, x_j) = (x_i \cdot x_j) \tag{20}$$

$$K(x_i, x_j) = [(x_i \cdot x_j) + c] \tag{21}$$

$$K(x_i, x_j) = e^{-\frac{1}{2\sigma^2}\|x_i - x_j\|^2} \tag{22}$$

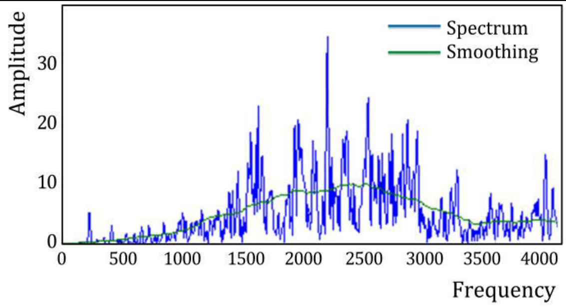
$$K(x_i, x_j) = \tanh(b(x_i \cdot x_j) + c) \tag{23}$$

RESULTS

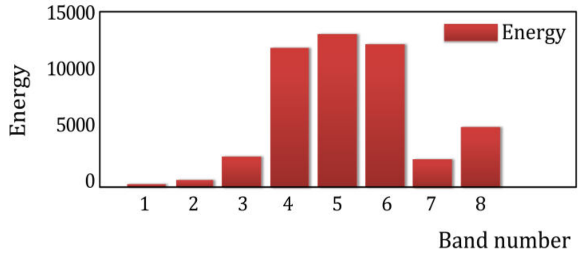
The tests were performed in a PC with an Intel Core i7 2.2 GHZ processor and 8GB RAM memory. The test applications were developed in MATLAB 7.0. A neural network with the features shown in Table 1 was built; these features were obtained after having tested different settings based on the number of samples and the components of these samples. An average time of 47 seconds for training the network and an average time of 1 second for signature recognition was used.

Signal pre-processing

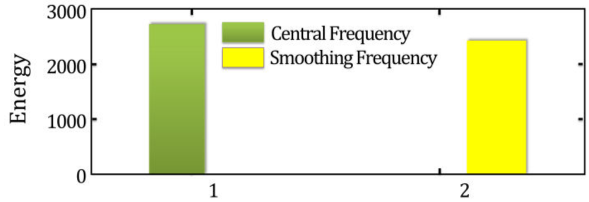
Figure-4(a) presents an example of an analyzed spectrum and the result of the smoothing stage. Figure-4(b) shows the estimated energy in each one of the 8 bands that subdivide the original signal and Figure 4(c) presents the central frequency and the smoothed average frequency.



(a)



(b)



(c)

Figure-4. (a) Signal spectrum, (b) Higher power spectral density frequency peaks and (c) Central frequency and smoothed average frequency.

The method was tested with a data set of 110 samples coming from 2 types of vessels and a data set of biological noises, distributed in the following way: 42 samples from go-fast boats, 26 from merchant vessels and 42 from biological noise. With the above samples, the average success percentage attained was 92%. Table-2 shows the success percentages for each sample subset.

Table-1. Building neural network configurations.

Feature	Value
Number of layers	2
Neurons in layer 1	20
Neurons in layer 2	30
Number of inputs	10
Layer 1 activation function	tansig
Layer 1 activation function	tansig
Learning algorithm	TRAINLM
Number of classes to identify	4
Number of training patterns	97
Number of test patterns	33

Table-2. Precision percentage during recognition.

		% accuracy	
Source	**Samples**	**RNA**	**SVM**
Biologic (Dolphins)	42	90	90
Merchant	26	92	92
Go-fast boat	42	81	71
Total	110	87	84

CONCLUSIONS

This work describes a method for acoustic signature recognition of marine vessels. It was shown that the combined use of Fourier transform with neural networks and support vector machines attains an accuracy percentage of 92%. Results showed that with the computational implementation of the proposed method the recognition of a vessel from its acoustic signature takes one second approximately, which constitutes a great improvement over the time used by a human worker (around one minute).

As future work, we plan to perform the development of a system that combines different classifiers and learning techniques, to increase recognition accuracy.

ACKNOWLEDGEMENTS

The authors thank two Colombian navy institutions: the Escuela Naval de Suboficiales de Barranquilla-ENSB- and the Flotilla de Submarinos de la Escuela Naval de Oficiales Almirante Padilla-ENAP-for all the support received from scuba divers and sonar operators, and for the acoustic-signal samples delivered.

REFERENCES

[1] Urick R.J. 1996. Principles of underwater sound. 3rd Edition, Editorial McGraw-Hill, New York, USA.

[2] Regazzoni C., Tesei A., Tacconi G. 1994. A comparison between spectral and bispectral analysis for ship detection from acoustical time series. Acoustics, Speech and Signal Processing. 2(1): 289-292.

[3] Pflug L.A., Ioup G.E., Ioup J., Jackson P. 1997. Variability in higher order statistics of measured shallow-water shipping noise. IEEE Signal Processing Workshop on Higher-Order Statistics, Banff. pp. 400-404.

[4] Streilein W.W., Gaudiano, P., Carpenter G.A. 1998. A neural network for object recognition through sonar on a mobile robot. Intelligent Control (ISIC). Held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), Intelligent Systems and Semiotics (ISAS), pp. 271-276, 14-17 September.

[5] Lathi B.P. 1998. Signal Processing and Linear Systems. New York: Oxford University Press.

[6] Fulop Sean A. and Fitz, Kelly. 2006. A Spectrogram for the Twenty-First Century. Acoustics Today. pp. 26-32.

[7] Chen C., Lee J., Lin M. 1998. Classification of underwater signals using wavelet transforms and neural networks, Math. Comput. Model. 27(2): 47-60.

[8] Soares-Filho W., Seixas J.M., Caloba L.P. 2002. Enlarging neural class detection capacity in passive sonar systems. IEEE Int. Symp. On Circuits and Systems, Scottsdale. 3(1): 105-108.

[9] Howell B.P., Wood S., Koksal S. 2003. Passive sonar recognition and analysis using hybrid neural networks. IEEE Oceans, San Diego. 4: 1917-1924.

[10] Kang C., Zhang X., Zahang A., Lin H. 2004. Underwater acoustic targets classification using welch spectrum estimation and neural networks, Adv. Neural Netw. 3173: 930-935.

[11] Eom K., Wellman M., Srour N., Hillis D., Chellappa R. 1997. Acoustic target classification using multiscale methods. Sensors and Electron Devices Symp. University of Maryland, College Park, MD.

[12] Huang J., Zhao J., Xie Y. 1997. Source classification using pole method of AR model. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Munich. 1: 567-570.

[13] Baran R.H., Coughlin J.P. 1991. A neural network for target classification using passive sonar. Proc. Conf. on Analysis of Neural Network Applications. pp. 188-198.

[14] Meister J. 1993. A neural network harmonic family classifier, J. Acous. Soc. Am. 93(3): 1488-1495.

[15] Park K.-C., Lee P.-H., Park J.-N., Yoon J.-R. 2006. Neural networks for automatic classification of ship-radiated noise, Jpn J. Appl. Phys. Part 1. 45(5): 4859-4861.

[16] Chen C-H., Lee J-D., Lin M-C. 2000. Classification of Underwater Signals Using Neural Networks. Tamkang Journal of Science and Engineering. 3(1): 31-48.

[17] LI Si C, YANG D-s., JIN L-p. 2009. Classifying ships by their acoustic signals with a cross-bispectrum

algorithm and a radial basis function neural network Journal of Marine Science and Application. 8(1): 53-57.

[18] Chung K., Sutin, A., Sedunov, A., Bruno M. 2011. DEMON: Acoustic Ship Signature Measurements in an Urban Harbor Advances in Acoustics and Vibration, Hindawi Publishing Corporation. 2011(1): 1-13.

[19] Lobo V., Moura P.F. 1995. Ship noise classification using Kohonen Networks. Engineering applications of artificial neural networks International conference; 1st, Engineering applications of artificial neural networks, in Proc. of EANN 95, pp. 601-604.

[20] Yang S., Li Z., Wang X. 2002. Ship recognition via its radiated sound: The fractal based approaches Acoustical Society of America. 112(1).

[21] Kuçukbayrak M., Gunes O., Grade Jr., ARICA N. 2009. Underwater Acoustic Signal Recognition Methods. Journal of Naval Science and Engineering. 5(3): 64-78.

[22] Averbucha A., Zheludeva V., Neittaanmäkib P., Wartiainenb P., Huomanc K., Jansonc K. 2011. Acoustic Detection and Classification of River Boats. Applied Acoustics. 72(1): 22-34.

[23] Chen S., Zhang H. 2011. Detection of Underwater Acoustic from Ship Noise Based on WPT Method. Chaos-Fractals Theories and Applications (IWCFTA), 2011 Fourth International Workshop on. pp. 324-327.

[24] Palshikar G. 2009. Simple algorithms for peak detection in time-series. In: Proceedings of 1st IIMA International Conference on Advanced Data Analysis, Business Analytics and Intelligence, Ahmedabad, India.

[25] D. Enqing L. Guizhong, Z. Yatong, c. Yu. 2002. Voice Activity Detection Based on Short-Time Energy and Noise Spectrum Adaptation, 6th International Conference on Signal Processing.

[26] L.R. Rabiner and M.R. Sambur. 1975. An Algorithm for Determining the Endpoints of Isolated Utterances, Bell Syst. Tech. J. 54: 297-315.

[27] Yang Xingjun. 1995. Digital Processing of Speech Signal. Publishing House of Electronics Industry.