KOHONEN NETWORKS AS SHIPS CLASSIFIER

ANDRZEJ ŻAK

Polish Naval Academy
Smidowicza 69, 81-103 Gdynia, Poland
a.zak@amw.gdynia.pl

The paper presents the technique of artificial neural networks used as classifier of hydroacoustic signatures generated by moving ship. In the paper firstly the method of feature extraction from hydroacoustic signatures using calculation of Mel-Frequency Cepstral Coefficients was discussed. Next the method of feature matching using for purpose of object classification basing on hydroacoustic signatures was described. The technique of artificial neural networks especially Kohonen networks which belongs to group of self organizing networks where chosen to solve the research problem of classification. The choice was caused by some advantages of mentioned kind of neural networks for example they are ideal for finding relationships amongst complex sets of data, they have possibility to self expand the set of answers for new input vectors. To check the correctness of classifier work the research in which the number of right classification for presented and not presented before hydroacoustic signatures were made. Some results of research were presented on this paper.

INTRODUCTION

The problem of acoustic signals recognition belongs to a much broader topic in scientific and engineering so called pattern recognition. The goal of pattern recognition is to classify objects of interest into one of a number of categories or classes [6]. Hydroacoustics signals identification or classification is the process of automatically recognizing what kind of object is generating acoustics signals on the basis of individual information included in generated sounds. All signal recognition systems, at the highest level, contain two main modules (figure 1) feature extraction and feature matching. Feature extraction is the process that extracts a small amount of

data from the hydroacoustics signatures that can later be used to represent each object. Feature matching involves the actual procedure to identify the unknown object by comparing extracted features from input sounds with the ones from a set of known stored in some kind of database. Therefore signal recognition systems have to serve two distinguish phases. The first one is referred to the enrollment sessions or training phase while the second one is referred to as the operation sessions or testing phase. In the training phase, each registered object has to provide samples of their sounds so that the system can build or train a reference model for that object. During the testing – operational phase, the input sound is matched with stored reference models and recognition decision is made.
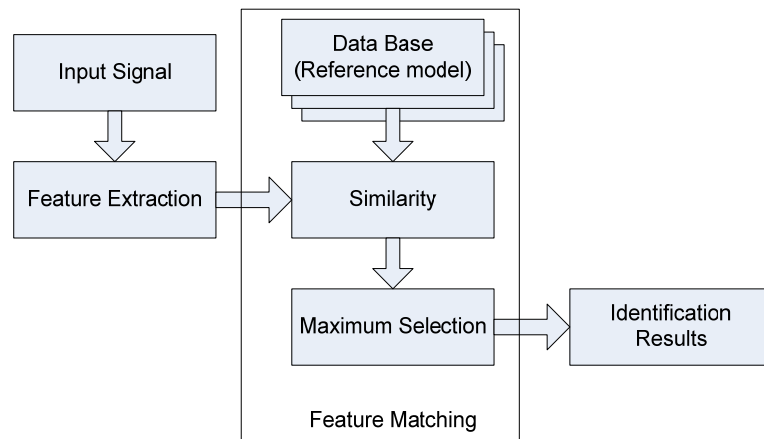

Fig.1 Signal identification

Hydroacoustics signatures recognition is a difficult task and it is still an active research area. Automatic signal recognition works based on the premise that sounds emitted by object to the environment are unique for that object. However this task has been challenged by the highly variant of input signals. The principle source of variance is the object himself. Sound signals in training and testing sessions can be greatly different due to many facts such as object sounds change with time, efficiency conditions (e.g. some elements of machinery are damaged), sound rates, etc. There are also other factors, beyond object sounds variability, that present a challenge to signal recognition technology. Examples of these are acoustical noise and variations in recording environments and changes of environment itself.

As a hydroacoustics signatures in this paper will be understood only sound made by ships in motion. Researches produce that ships has characteristic for them components of spectra in frequency range from about 5 Hz to 2 kHz. Moreover hydroacoustics signals are quasi-stationary so short-time spectral analysis is one of the common way to characterize this kind of sounds.

Hydroacoustic signatures have the great significance because its range of propagation is the widest of all physics field of ship. Controlling and classification of acoustic signature of vessels is now a major consideration for researchers, naval architects and operators. The advent of new generations of acoustic intelligence torpedoes and depth mines has forced to a great effort, which is devoted to classify objects using signatures generated by surface ships and submarines. It has been done in order to increase the battle possibility of submarine armament. Its main objectives are to recognize the ship and only attack this one which belongs to opponent.

In the paper the Kohonen Neural Networks were discussed as hydroacoustic signals, generated by moving ship, classifier. As a signal feature extraction method for hydroacoustic signatures, the Mel-Frequency Cepstral Coefficient (MFCC) were described.

1. SIGNAL FEATURE EXTRACTION

The purpose of signal feature extraction module is to convert the sound waveform to some type of parametric representation for further analysis and processing. This is often referred as the signal-processing front end. A wide range of possibilities exist for parametrically representing the signals for the sound recognition task, such as Linear Prediction Coding (LPC), Mel-Frequency Cepstrum Coefficients (MFCC) [8], and others. Mel-Frequency Cepstrum Coefficients method will be discussed in this paper.

A block diagram of the structure of an MFCC processor is given on figure 2. First step of MFCC processor is the frame blocking. In this step the continuous sound is blocked into frames of N samples, with adjacent frames being separated by M where $M < N$. The first frame consists of the first N samples. The second frame begins M samples after the first frame, and overlaps it by $N - M$ samples. Similarly, the next frames are created so this process continues until all the sound is accounted for within one or more frames.
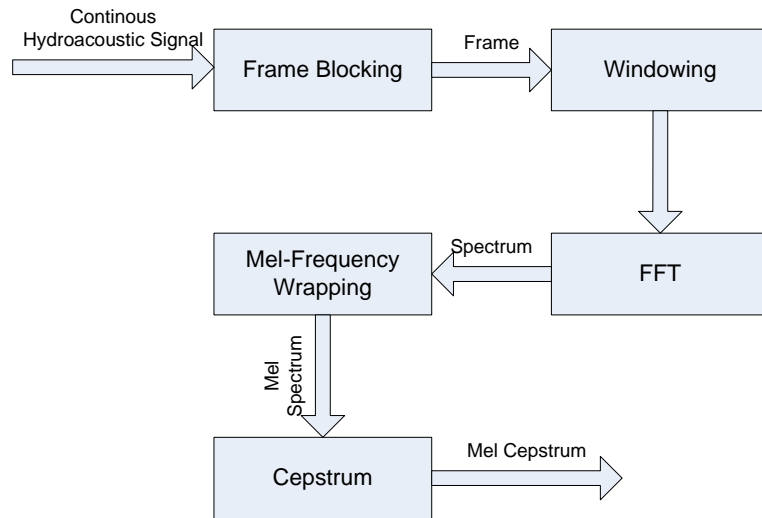

Fig.2 Block diagram of the MFCC processor

The next step in the processing is to window each individual frame so as to minimize the signal discontinuities at the beginning and end of each frame. The concept here is to minimize the spectral distortion by using the window to taper the signal to zero at the beginning and end of each frame. If we define the window as: $w(n), \quad 0 \le n \le N - 1$, where N is the number of samples in each frame, and assume that used window is Hamming window, then the result of windowing is the signal:

$$y(n) = x(n)(0.54 - 0.46\cos(\frac{2\pi n}{N-1})), \quad 0 \le n \le N-1 \tag{1}$$

The next processing step is the Fast Fourier Transform, which converts each frame of N samples from the time domain into the frequency domain. The FFT is a fast algorithm to implement the Discrete Fourier Transform (DFT) which is defined on the set of N samples, as follow:

$$X_n = \sum_{k=0}^{N-1} x_k e^{-2\pi jkn/N}, \quad n = 0,1,2,\ldots,N-1 \tag{2}$$

Next step in MFCC processor is the Mel-frequency Wrapping. Each tone with an actual frequency f, measured in [Hz], a subjective pitch is measured on a scale called the 'mel' scale. The mel-frequency scale is a linear frequency spacing below *1000 [Hz]* and a logarithmic spacing above *1000 [Hz]*. Therefore we can use the following approximate formula to compute the mels for a given frequency f in *[Hz]*:

$$mel(f) = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right) \tag{3}$$

In final step, we convert the logarithmic mel spectrum back to time. The result is called the mel frequency cepstrum coefficients (MFCC). The cepstral representation of the sound spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis. Because the mel spectrum coefficients, and so their logarithm, are real numbers, we can convert them to the time domain using the Discrete Cosine Transform (DCT). Therefore if we denote those mel power spectrum coefficients that are the result of the last step are S_k, $k = 1,2,\ldots,K$, we can calculate the MFCC's, as [8]:

$$c_n = \sum_{k=1}^{K} \log(S_k) \cos(\frac{n(k-0.5)\pi}{K}), \quad n = 1,2,\ldots,K \tag{4}$$

Note that we exclude the first component, c_0 from the DCT since it represents the mean value of the input signal which carried little object specific information.

2. CLASSIFICATION METHOD – KOHONEN NEURAL NETWORK

Kohonen neural network, also known as The Self-Organizing Map (SOM) is a computational method for the visualization and analysis of high-dimensional data, especially experimentally acquired information [3, 4].

One of the most interesting aspects of SOMs is that they learn to classify data without supervision. With this approach an input vector is presented to the network and the output is compared with the target vector. If they differ, the weights of the network are altered slightly to reduce the error in the output. This is repeated many times and with many sets of vector pairs until the network gives the desired output. Training a SOM however, requires no target vector.

For the purposes of this paper the two dimensional SOM will be discussed. The network is created from a 2D lattice of 'nodes', each of which is fully connected to the input layer.

SOM does not need a target output to be specified unlike many other types of network. Instead, where the node weights match the input vector, that area of the lattice is selectively optimized to more closely resemble the data for the class the input vector is a member of. From an initial distribution of random weights, and over many iterations, the SOM eventually settles into a map of stable zones. Each zone is effectively a feature classifier, so the graphical output can be treated as a type of feature map of the input space.

Training occurs in several steps and over many iterations [4]:
- Each node's weights are initialized.
- A vector is chosen at random from the set of training data and presented to the lattice.
- Every node is examined to calculate which one's weights are most like the input vector. The winning node is commonly known as the Best Matching Unit (BMU).
- The radius of the neighborhood of the BMU is now calculated. This is a value that starts large, typically set to the 'radius' of the lattice, but diminishes each time-step. Any nodes found within this radius are deemed to be inside the BMU's neighborhood.
- Each neighboring node's (the nodes found in step 4) weights are adjusted to make them more like the input vector. The closer a node is to the BMU, the more its weights get altered.
- Repeat step 2 for N iterations.

To determine the best matching unit, one method is to iterate through all the nodes and calculate the distance between each node's weight vector and the current input vector. The node with a weight vector closest to the input vector is tagged as the BMU.
There are many methods to determine the distance [5] but the most popular Euclidean distance is given as:

$$d(x, w_i) = \|x - w_i\| = \sqrt{\sum_{j=0}^{N} (x_j - w_{ij})^2} \tag{5}$$

where: x is the current input vector; w is the node's weight vector.

Each iteration, after the BMU has been determined, the next step is to calculate which of the other nodes are within the BMU's neighborhood. All these nodes will have their weight vectors altered in the next step. A unique feature of the Kohonen learning algorithm is that the area of the neighborhood shrinks over time. This is accomplished by making the radius of the neighborhood shrink over time.

To do this the exponential decay function can be used as follow:

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\lambda}\right) \quad t = 0, 1, 2, \ldots \tag{6}$$

where: σ_0 denotes the width of the lattice at time t_0; λ denotes a time constant; t is the current time-step (iteration of the loop).

Every node within the BMU's neighborhood (including the BMU) has its weight vector adjusted according to the following equation [1]:

$$w_{ij}(t+1) = w_{ij}(t) + \Theta(t)\eta(t)(x_j(t) - w_{ij}(t)) \tag{7}$$

where: t represents the time-step; η is a small variable called the learning rate, which decreases with time.

The decay of the learning rate is calculated each iteration using the following equation:

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\lambda}\right) \quad t = 0,1,2,\ldots \tag{8}$$

In equation 11, not only does the learning rate have to decay over time, but also, the effect of learning should be proportional to the distance a node is from the BMU. Indeed, at the edges of the BMUs neighborhood, the learning process should have barely any effect at all. Ideally, the amount of learning should fade over distance similar to the Gaussian decay according to the formula:

$$\theta(t) = \exp\left(-\frac{dist}{2\sigma^2(t)}\right) \quad t = 0,1,2,\ldots \tag{9}$$

where: $dist$ is the distance a node is from the BMU; σ is the width of the neighborhood function as calculated by equation (6).

3. THE RESULTS OF RESEARCH

During research the five ships were measured on the Polish Navy Test and Evaluation Acoustic Ranges which schema was presented on figure 3. Ships No. 1 was minesweeper project 206FM, ship No. 2 was minesweeper project 207D, ship No. 3 was salvage ship project 570, ship No. 4 was minesweeper project 207P, and ship No. 5 was racket corvette project 1241RE.

The recordings were carried out by means of the array of hydrophones. Several hydrophones were strung in a line along the bottom in shallow water. The depth was about *10 m*. During the ship measurements, the average see wave height was less than *1 m* and wind speeds less than *5 m/s*, so the ambient noise level was low. At the time of the measurements the sound velocity profile was typical for the summer. This curve was smooth with gradually decreasing gradient without mixed layers. The ship under test was running at a constant speed and course during cross over hydrophones. The array of hydrophones was mounted about 1 m above sea bottom on tripod. The bottom-mounted hydrophones range is very useful for measuring the noise of surface ships. What more when they are used bottom-fixed hydrophones the irrelevant low-frequency wave-induced noise is also eliminated [7].

All of investigated ships were measured at the similar hydrological and metrological conditions. Every ship was measured with few, various speed of crossing.

Data form hydrophones were recorded on digital recorder designed by crew of Radiolocation and Hydrolocation Department of Polish Naval Academy. This system has possibility to simultaneous recording in 16 channels with resolution of 16 bits and sampling frequency up to 250 kHz per channel. As a sensors of acoustic field of moving ship were used hydrophones produced by Reson model TC4032. This hydrophones has omnidirectional characteristic in horizontal directivity so they were positioned parallel to the plane of sea bottom. Other parameters which cause that these sensors are proper to acquire data for classification systems are: high sensitivity equal -170 dB re 1V/μPa, preamplifier gain of 10 dB and broad

usable frequency band from 5 Hz to 120 kHz. Mentioned above digital recorder has possibility to direct connections of hydrophones TC4032.

The best solutions to detect a ship are the discrete components in the low frequency part of the ship's noise spectrum and that only narrow band filters can be used. This must be done because there are no components discrete lines at frequencies range grater than 200 Hz in the modern submarines and surface warships. In the Baltic's shallow waters an the conditions under which the measurements were made, the area of optimal frequencies for the propagation of sound lies in the band from several Hz up to 2 kHz [2].
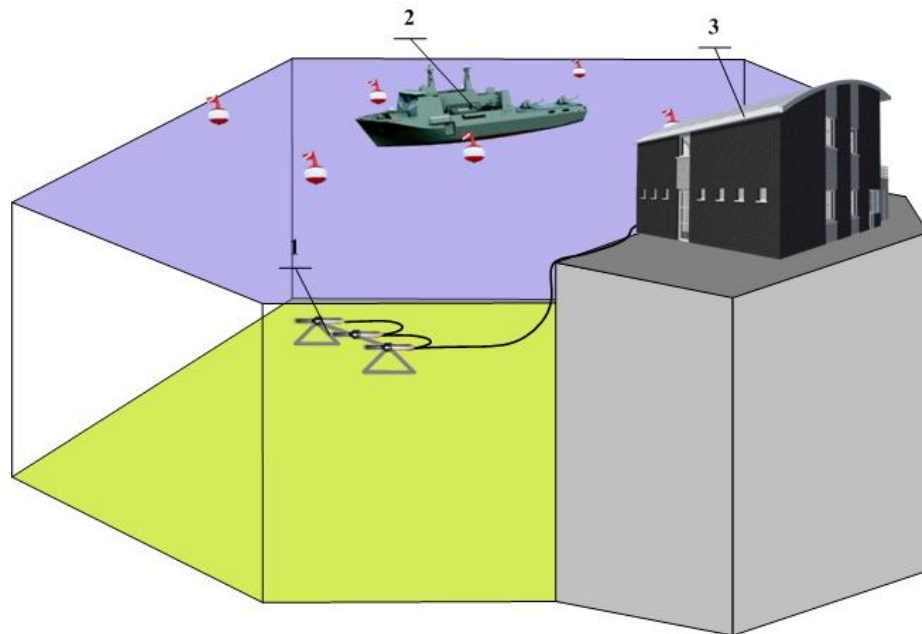

Fig.3 Schema of hydroacoustic range during measurements; 1) sensors of acoustic signatures – array of hydrophones, 2) measured ship, 3) station with mounted hydroacoustic measuring system

Used Kohonen network has two dimensional architecture. For this case because of speed of learning, possibilities to classify data and possibilities to generalize the knowledge it seems that follows values are the best: number of neurons: 30x30 neurons map, beginning size of area of neighborhood: 4, beginning learning rate: 0.25 and method to determine the distance: Euclidean distance.

Data presented as input for Kohonen network were transformed according the method presented in second section. So, Mel-Frequency Cepstral Coefficients calculated from recorded signals become as the input signal for feature matching subsystem it means Self Organizing Maps – Kohonen networks.

After about 35 000 cycles of neural network learning, was obtained the map of memberships for every presented ship as it is shown on figure 4. All areas activated by signals generated by considered ships were clearly separated. The example results of classifier work out after learning process was presented on figure 5 A. These results were received for data which where presented during neural network learning process. To find out if the building classifier is properly configured and learned some data which weren't presented before were calculated. The example results were presented on figure 5 B.

The table 1 shows number of correct classification of presented data relatively to the type of ship. The number of correct answer is presented as percent of all answers. The research was made for data which were presented during learning process and data which weren't presented before.
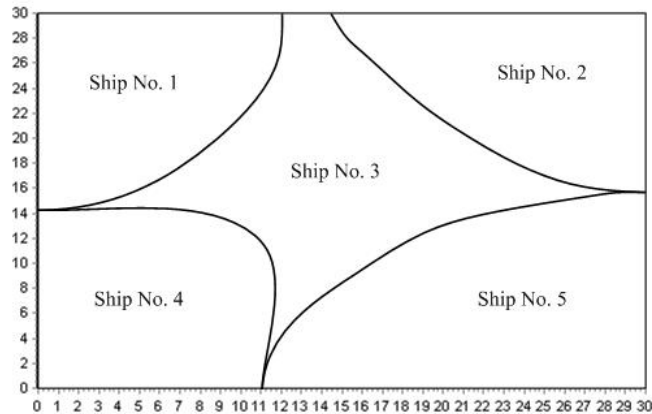

Fig.4 The map of partition for area of activation for researched ships

Tab.1 The number of correct classifications

Ship no. Data	1	2	3	4	5
presented before	95.2%	94.0%	96.1%	94.8%	93.5%
not presented before	78.2%	80.1%	76.8%	77.4%	78.9%

4. CONCLUSION

As it is shown on results the used Self-Organizing Map is useful for ships classification based on its hydroacoustic signature. Classification of signals that were used during learning process, characterize the high number of correct answer (above *90 %*) what was expected. This result means that used Kohonen network has been correctly configured and learned. Presentation of signals that weren't used during learning process, gives lowest value of percent of correct answer than in previous case but this results is very high too (about *75 %* of correct classification). This means that neural network has good ability to generalize the knowledge. It should be noticed that phase of preparing data it means feature extraction has influence on results of classification. Comparison with previous results of research when input data for neural networks has been calculated using Discreet Fourier Transform [9, 10] it were received *10 %* increase in correct classification using method of Mel-Frequency Cepstral Coefficients.

Presented case is quite simple because it not take into account that object sounds change with time, efficiency conditions (e.g. some elements of machinery are damaged), sound rates, etc. It doesn't consider the influence of changes of environment on acquired hydroacoustic signals.

Therefore these cases should be investigated in future research. More over in future research the influence of network configuration on the quality of classification should be checked.

The aim of presented method is to classify and recognize ships basing on its acoustic signatures. This method can found application in intelligence submarine weapon and in hydrolocation systems. From the point of ship's passive defense view it is desirable to minimize the range of acoustic signatures propagation. In other hand it is important to deform and cheat the similar system of our opponents by changing the "acoustic portrait" of own ships. Noise isolation
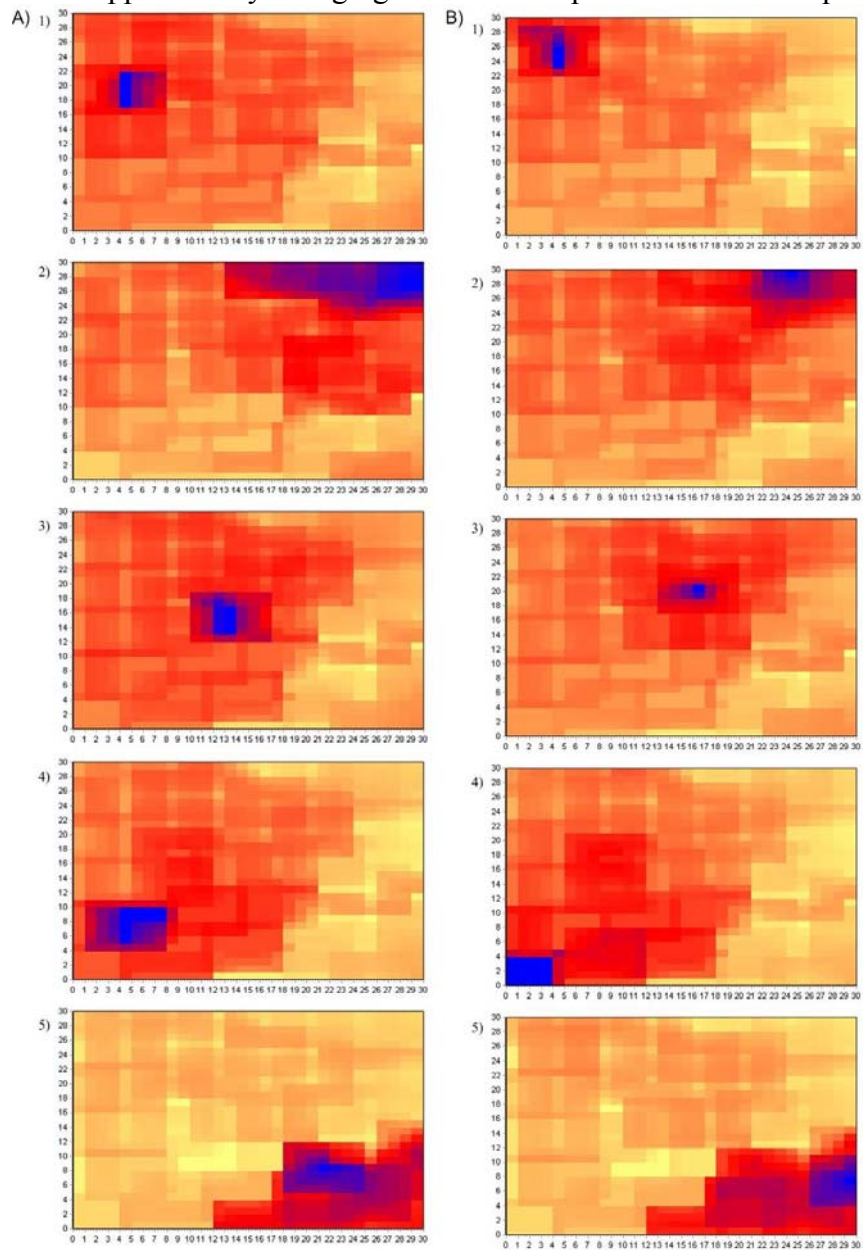

Fig. 5. The results of classifier work out: A) maps of memberships for data which were presented during learning process; 1) for ship no. 1, 2) for ship no. 2, 3) for ship no. 3, 4) for ship no. 4, 5) for ship no. 5. B) maps of memberships for data which weren't presented during learning process; 1) for ship no. 1, 2) for ship no. 2, 3) for ship no. 3, 4) for ship no. 4, 5) for ship no. 5.

systems for vessels employ a wide range of techniques, especially double-elastic devices in the case of diesel generators and main engines. Also, rotating machinery and moving parts should be dynamically-balanced to reduce the noise. In addition, the equipment should be mounted in special acoustically insulated housings (special kind of containers). One of the method to change the hydroacoustic signatures is to pump the air under the hull of ship. It cause the offset of generated by moving ship frequency into the direction of high frequency, the same the range of propagation become smaller.

REFERENCES

1. J.C. Fort, SOM's mathematics, Neural Networks, 19: 812–816, 2006.
2. I. Gloza, S. J. Malinowski, Underwater Noise Characteristics of Small Ships. Acta Acoustica United with Acustica, vol. 88 pp. 718-721, 2002.
3. S. Haykin, Self-organizing maps, Neural networks - A comprehensive foundation, 2nd edition, Prentice-Hall, 1999.
4. T. Kohonen, Self-Organizing Maps, Third, extended edition. Springer, 2001
5. S. Osowski, Neural Networks, Publishing House of Warsaw University of Technology, 1996.
6. K. Stąpor, Automatic object classification, Publishing House EXIT 2005.
7. R. J. Urick, Principles of Underwater Sounds, McGraw-Hill, New York 1975.
8. A. Żak, Creating patterns for hydroacoustics signals, Hydroacoustics Vol. 8, pp. 265- 270, Gdynia 2005.
9. A. Żak, Ship classification basing on hydroacoustic signals, Polish Naval Academy Research Journal, no 169 K/1 year XLVIII, pp. 417-424, Gdynia 2007.
10. A. Żak, Ships Classification Using Hydroacoustic Signatures, Proceedings of the 6th WSEAS Internatiuonal Conference on Computational Inteligence, Man-Machine Sytsems and Cybernetics, pp. 111-115, Spain 2007.