International Symposium on
HYDROACOUSTICS AND ULTRASONICS
EAA Symposium (formerly 13th FASE Symposium)
Gdańsk-Jurata, 12-16 May 1997

Method of shipping noise recognition

Wiesław Kiciński
Naval Academy
Gdynia, Poland

ABSTRACT

Multidimesional approach to shipping noise recognition is described. Using discriminatory analysis methods, classification of underwater noise sources in coastal range is done.

INTRODUCTION

Signal detection in nonstationary noise environment has been described extensively in the literature [1]. In the derivation of desired test procedure, one usually assumes either the model of a known deterministic signal in unknown corellated noise or the model of unknown random signal in white noise. The solution proposed for the former case is an adaptive implementation of the matched filter or the generalized likelihood ratio test, while that for the latter is the normalization technique.

Deterministic signal models and the corresponding test procedures are applied only to nonfluctuating targets. However, most of targets encountered with sonar are fluctuating, so they are better described by stochastic models. Very often the stochastic nature of signals is also attributed to the multipath propagation, especially in an underwater environment. The general idea of the normalization technique is based on comparing the squared amplitudes of the received data with the noise power to determine the presence of a target signal [7]. The normalization technique performs two functions: normalize the power of the received data in order to achieve a constant alarm rate and integrate the useful information in data to enhance the detection performance. These goals can be achieved, indeed, provided that the noise is spatially uncorrelated [6].

Unfortunately, the white noise assumption may not be valid in practice. Application of the conventional normalization technique to these situations will cause two problems: performance degradation and lost of the constant false alarm rate property. The second problem is a severe obstacle to the system implementation since the threshold of the test depends on the noise covariance matrix, which is typically unknown or varying in time.

The primary objective of this paper is to develop methods of multidimensional analysis for shipping noise recognition. The problem is focused on coastal range case, where changes of ambient noise characteristics result from approaching ships or disturbance noise.

*) This work was supported under grant no 8T11B02710

THE PROBLEM

Let

$$\mathbf{x}_N = (x_1, x_2 \dots x_N) \tag{1}$$

denotes the measurement vector of underwater noise.

The purpose of signal processing is to choose one of the following hypotheses:

$$\begin{array}{l} H_0\text{: } \mathbf{x} \text{ contains disturbance noise alone,} \\ H_1\text{: } \mathbf{x} \text{ contains shipping noise plus disturbance} \end{array} \quad (2)$$

Conventional approach to solve the problem stated by (2) lies in defining a test statistic and testing hypothesis on given significance level.

This paper deals with multidimensional analysis of underwater noise for recognition of shipping noise by discrimination of observed data.

The method proposed in this paper is based on a supervised pattern recognition procedure and consist of following stages:
- preprocessing of measured signal for measurement space constituting,
- reduction of measurement space to feature space of underwater noise,
- determining of the form of discriminant function and estimating their parameters,
- constituing of norm space for disturbance noise in field of discriminant variables,
- determining the measure of distance any observation from centroid of norm space,
- testing hypothesis (1) using the measure of distance the observation from norm space as test statistic.

The parameters estimation of underwater noise and testing will be performed by data analysis in the frequency domain using Fourier transform. The main emphasis is put upon multidimensional analysis of observed data for identification of underwater noise nature.

THEORY

Let

$$\mathbf{x_N} = (x_1, x_2 \ldots x_N) \quad (3)$$

be a N dimensional measurement vector. Usually this vector is preprocessed for dimensionality reduction.

Let a modified vector of measurement be determined as

$$\mathbf{y_P} = (y_1, y_2 \ldots y_P) = f(x_1, x_2 \ldots x_N) \quad (4)$$

The problem we shall consider here is to classify the vector $\mathbf{y_P}$ as belonging to one of the two possible classes:
- class 1: the result of measure is the shipping noise,
- class 2: the result of measure are the coastal disturbances.

Broadly speaking two distinct approaches to solving this problem may be traced: Statistical Decision Theory or Discriminant Analysis.

Statistical Decision Theory assumes that the forms for the underlying probability distributions are known, and uses the samples to estimate the values of their parameters.

For two classes case the Bayes minimum error decision rule may be used to assign $\mathbf{y_P}$ to class 1 if [2]:

$$\frac{p(\mathbf{y}|\omega_1)}{p(\mathbf{y}|\omega_2)} \gtrless \frac{P(\omega_2)}{P(\omega_1)} \Rightarrow \mathbf{y} \in \begin{cases} \Omega_1 \\ \Omega_2 \end{cases} \quad (5)$$

where:

$P(\omega_i)$ - probability that an object belongs to class ω_i ($i=1,2$),

$p(\mathbf{y}|\omega_i)$ - class - conditional probability density function.

In many cases neither $P(\omega_1)$ nor $P(\omega_2)$ are known. Moreover, the costs of misclassification may be virtually impossible to assess or may vary.

In many practical situations the form of the probability distributions that governs the pattern vectors is unknown. One approach to this problem is to use the samples to estimate the unknown probabilities and probability densities, and to use resulting estimates as if they were the true values.

Discriminant Analysis approach assumes that the form of discriminant functions is known, and uses the samples (measurement vector) to estimate the values of parameters of the classifier. However, none of them requires knowledge of the forms of underlying probability distributions, and in this sense all of them can be said to be nonparametric. A great deal of interest has been centered around linear methods because of the obvious computational advantages. Fisher used a linear function of the measurements [3]:

$$Y = \sum_{i=1}^{P} a_i y_i \tag{6}$$

where: a_i - linear coefficients, which maximized the square of the difference between the mean values of Y for classes divided by covariance of Y.

The problem of finding a linear discriminant function will be formulated as a problem of minimizing a criterion function. The obvious criterion function for classification purposes is the sample risk, the average loss incurred in classifying the set of design samples.

The matrix form of discriminant function may be written as

$$g(\mathbf{y}) = \mathbf{w}^T\mathbf{y} + w_o \tag{7}$$

where: $\mathbf{w}$ is called the weight vector and w_o is the threshold weight.

Two category classifier implements the following decision rule: decide ω_1 if $g(\mathbf{y})>0$ and ω_2 if $g(\mathbf{y})<0$. Thus $\mathbf{x}$ is assigned to ω_1 if the inner product $\mathbf{w}^T\mathbf{y}$ exceeds the threshold $-\omega_o$. If $g(\mathbf{y})=0$, $\mathbf{y}$ can ordinarily be assigned to either class, but in this chapter we shall leave the assignment undefined.

The equation $g(\mathbf{y})=0$ defines the decision surface that separates points assigned to ω_1 from points assigned to ω_2. When $g(\mathbf{y})$ is linear them this decision surface is a hyperplane.

The proposed procedure of discriminant analysis consist of two steps. The first one is to determine coefficients of classification functions and the second one is to classify the observed data.

NORM SPACE

The norm space for multidimensional subject can be derieved from measuring training set.

Let

$$\begin{aligned} \mathbf{x}_1 &= (x_{11}, x_{12} \dots x_{1N}) \\ \mathbf{x}_2 &= (x_{21}, x_{22} \dots x_{2N}) \\ &\dots\dots\dots\dots \\ \mathbf{x}_k &= (x_{k1}, x_{k2}, \dots, x_{kN}) \end{aligned} \tag{8}$$

denotes the reference data that contain coastal disturbances noise only.

The reference data can be obtain from a'priori identification (observation, scans) of the underwater environment in coastal range. The vector

$$\mathbf{x} = (x_1, x_2 \dots x_N) \tag{9}$$

denotes an N-dimensional measurement vector to be RMS spectrum of disturbance noise.

Let

$$\mathbf{y}_P = (y_1, y_2 \dots y_P)$$

denotes an P-dimensional modified measurement vector to be third octave spectrum of disturbance noise.

For the assumptions made above, the norm point estimator is given by

$$\bar{\mathbf{y}} = (\bar{y}_1, \bar{y}_2 \dots \bar{y}_P) = \left(\frac{1}{n}\sum_i x_{1i}, \frac{1}{n}\sum_i x_{2i} \dots \frac{1}{n}\sum_i x_{Pi}\right) \tag{11}$$

and is the maximum likelihood estimator.

MAHALANOBIS DISTANCE

The Mahalanobis distance is the distance of a case from centroid in multidimensional space. This measure provides an indication of whether or not an observation is an outlier with respect to the norm space of observation. It is computed as follow:

$$D = (\bar{\mathbf{y}} - \mathbf{y})C^{-1}(\bar{\mathbf{y}} - \mathbf{y})^T \tag{12}$$

where:

- $\bar{\mathbf{y}}$ - the vector of means for independent variables,
- $\mathbf{y}$ - the vector of data for independent variables,
- C - covariance matrix.

The Mahalanobis distance has χ^2 distribution with p degree of freedom.

EXPERIMENTAL RESULTS

In this section the experimental results of shipping noise recognize are presented.
The general scheme for multidimensional underwater noise processing is presented in fig. 1.

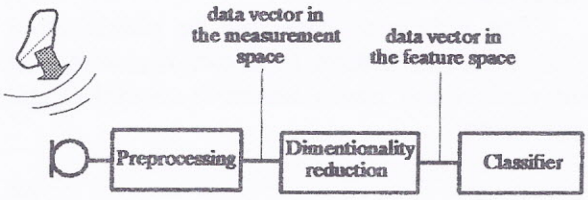

Fig.1. Scheme underwater noise processing

The underwater noise signal from sensor (hydrophone) was preprocessed in programmable signal analyzer GC-89. Next, the records of data from signal analyzer were processed in computer system to achieve pattern of disturbance noise and classify shipping noise data.

The hydrophones were submerged to 20 meters below the sea surface and were mounted on a stationary measurement system. The ambient noise from output of hydrophone was recorded by means of analog tape recorder. Each sample was 60 sec long.

The underwater noise were investigated by analysis of data in the frequency domain using Fourier transform. Each recorded underwater signal was processed by the digital signal analyzer. Records of digital data were analyzed in frequency bands covering range 2 - 128 Hz.

The third octave spectrum was applied to parametrize the measured data.

The central problem in implementing the methods stated above is grouping objects into classes according to their "similarity". For this aim the data were collected in the south part of the Baltic Sea in the coastal range during one year period time. Noise observations were made in variety of hydrological conditions.

From full measurements set, the data consisting of coastal disturbances were preselected using perception method. In multidimensional analysis terminology this set of data is called set of norm space for disturbances underwater noise. The set of data of norm space consist of 30 measures underwater noise disturbances and it is assume that it is representative for coastal region.

One can see that the spectrum characteristic is like wideband noise. The tonal frequen-

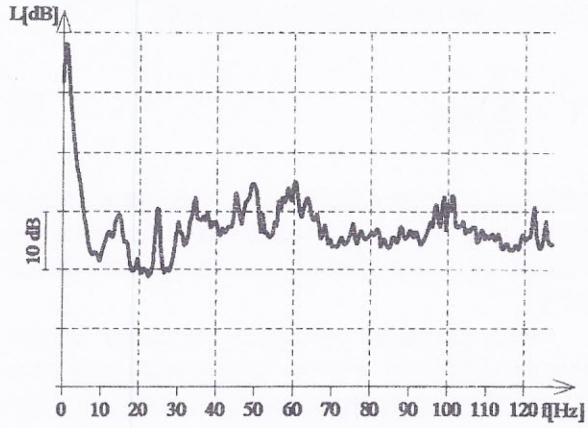

Fig. 2. RMS spectrum disturbance noise

cies to appear randomly and their amplitude is random too.

For identification of underwater shipping noise nature the measures of five representative ships were done. The measures of shipping noise were carried out for typical working regimes of mechanisms of ships. In the analyzed band of frequency the dominant sources of shipping noise are propeller and propulsion machinery. The radiated acoustic spectrum consist of broad, continues spectral components, as well as narrow-band components (also refereed to as discrete lines or tonals). This behavior allows to deduce that shipping noise is a wide-band process in the selected range; anywhere spectrum in low frequency range (less than about 100 Hz) is higher than elsewhere. This perception was employed in the discrimination procedure of underwater noise.

The distinctive futures for disturbance and shipping noise were extracted from measuring training set of data.

For underwater disturbances the training set of data consisted of 40 results of measures. From this set of data 10 measures were extracted as outliers. Finally, the training set of data in measurement space consisted of 30 data vectors in form RMS spectrum of disturbance noise in band 2-128 Hz. The data from measurement space were processed to feature space in form of third octave spectrum of disturbance noise.

The measurement space for shipping noise was covered the set of measures for five representative ships. This set of data called testing data consisted of 25 RMS spectrum measures. Like before, the data were processed to form of vectors of third octave spectrum.

Future extraction procedures may be based on intuition or physical consideration of the problem or they may be purely mathematical techniques that seek to reducing the dimensionality of the observation space in a prescribed way. For future extraction of disturbance and shipping noise, discriminant procedures were used.

The main results of discriminant procedures applying for futures extraction are shown in tab. 1.

The variables V1...V24 represent the centroid frequency of 1/3 octave spectrum bands covering 2 - 128 Hz.

The standard Wilk's λ statistic is used to denote the statistical significance of the discriminatory power of the current model. Its value can range from 1.0 (no discriminatory power) to 0.0 (perfect discriminatory power). Statistics for overall discrimination are computed as the ratio of the determinant of within-groups covariance matrix over the determinant of the total covariance matrix. So called partial lambda (λ_p) is used to determine the unique contribution of the respective variable to the discrimination between groups. It is computed as multiplicative increment in lambda that results from adding the respective variable. The standard F statistic is used for a variable to indicate its statistical significance in the discrimination between groups, that is, it is a measure of the extent to which a variable makes a unique contribution to the prediction of the group membership. Significant level - p is associated with the respective F. Generally the smaller the Wilk's lambda value, the greater is contribution to overall discrimination.

The variables from tab. 1 perfectly discriminate between disturbance and shipping noise and were used as distinctive discriminatory terms for shipping noise recognition. Thus, we may conclude at this point that the third octave frequencies represented by V2, V3, V4, V5, V6, V11, V12, V14, V17 and V18 are the major variables that

Tab. 1 Discriminant function analysis summary

N=16	Wilks	λ_p	F	p
Ship A				
V3	0.121	0.805	6.80	0.00
V4	0.280	0.347	52.77	0.00
V5	0.217	0.447	34.63	0.00
V6	0.325	0.298	65.92	0.00
V11	0.343	0.283	71.04	0.00
V14	0.104	0.934	1.98	0.10
Ship B				
V2	0.363	0.273	80.04	0.00
V3	0.128	0.772	8.87	0.06
V4	0.331	0.299	70.23	0.00
V5	0.158	0.626	17.96	0.00
V11	0.223	0.444	37.58	0.00
V12	0.264	0.375	50.00	0.00
V14	0.206	0.481	32.37	0.00
Ship C				
V2	0.98	0.614	16.96	0.00
V4	0.15	0.379	44.23	0.00
V6	0.06	0.896	3.15	0.08
V14	0.21	0.287	67.22	0.00
V18	0.20	0.288	66.64	0.00
Ship D				
V3	0.05	0.931	2.06	0.10
V11	0.36	0.133	182.97	0.00
V12	0.32	0.152	156.52	0.00
V14	0.33	0.149	159.66	0.00
V18	0.26	0.183	124.98	0.00
Ship E				
V6	0.01	0.597	18.87	0.10
V11	0.09	0.094	269.37	0.00
V12	0.07	0.119	205.41	0.00
V14	0.24	0.036	740.35	0.00
V17	0.23	0.037	736.37	0.00

allow to discriminate between the groups of measurement data.

Using discriminatory terms from tab. 1 the norm space in field of discriminatory variables for disturbance noise was constituted. The basic statistics for norm space are shown in tab. 2.

Tab. 2 The basic statistics for norm space

Mahalanobis distance			
Min	Max	Mean	Std. dev.
0.98	27.72	8.69	9.02

The shipping noise recognition was tested on set of 10 results of measure underwater noise. For all cases the Mahalanobis distance was computed. The results of tests are shown in tab. 3.

Tab. 3 Mahalanobis distance for shipping noise

Mahalanobis distance			
Min	Max	Mean	Std. dev.
128.6	325.8	224.3	84.5

One can see that for all cases the Mahalanobis distances for shipping noise data are extremely greater than the Mahalanobis distance for data set from norm space. They are statistically significant on significant level $p=0.05$.

SUMMARY

The method for shipping noise recognition was described. The problem is extremely important in the coastal range where observed data consist of ambient and coastal disturbances together with shipping noise. The main emphasis is put upon multidimensional analysis of observed data for identification of underwater noise nature. By determining distinctive features of shipping and coastal noise the classification of underwater noise was done.

The Mahalanobis distance was applied as measure of distance between the center of classes corresponding to ambient or coastal noise.

To demonstrate the effectiveness of the proposed methodology, some experimental results were enclosed. As may be seen the multidimensional approach to the underwater noise investigation proved to be powerful and succesfull one.

REFERENCES

1. Bouvet M., Schwartz S.C., (1988), Underwater Noises: Statistical Modeling Detection, and Normalization JASA, Vol 83, (pp. 1023-1033).
2. Duda R. O., Hart P. E., (1986), Patern classification and scene analysis. John Wiley & Sons, New York.
3. Hand D.J., (1981), Discrimination and classification. Wiley & Sons,. New York.
4. Kiciński W., Kozaczka E., (1996), Classification of Underwater Noise by linear Discriminant Functions. 3rd European Conference on Underwater Acoustics, Heraklion, Crete, Greece, (pp.1135-1139).
5. Serviere C., Baudois D., (1995), Source separation with noisy observation: A noise cancelling application. Signal Processing, Vol 42, (pp.45-57).
6. Tacconi G., Astengo F., Marras P., Oddone M., (1994), Experimental results of a MASAT project on shipping noise evaluation. Undersea Defence Technology London., (pp. 413-417).
7. Zhang Q.T., (1996), A CFAR processor for the detection of unknown random signals in nonstationary correlated noise. Signal Processing Vol. 48, (pp. 17-26).