XVII IMEKO World Congress
Metrology in the 3rd Millennium
June 22-27, 2003, Dubrovnik, Croatia

PROCESSING OF MEASURING SIGNALS FOR MONITORING OF TRANSIENTS IN UNDERWATER ENVIRONMENT

Wieslaw Kicinski

Naval University of Gdynia, Poland

Abstract – The problem of processing of the class of measuring signals, so called transient signals, caused by shocks of the oxygen tank of diver has been discussed. The method of transient signals processing based on wavelet transform and multivariate data analysis has been described.

Keywords: transient signals, wavelet transform, underwater noises.

1. INTRODUCTION

Detection and identification noise sources are among the most important and challenging goals in the passive monitoring of underwater environment.

This paper is devoted to the processing of the class of measuring signals, so called transient signals, caused by shocks of the oxygen tank of diver. In other words an approach to blind source separation in underwater environment has been presented. This problem is very difficult because of the lack of information on the signal of interest:

- the temporal signal shape is not exactly known,
- the duration is not known and generally needs to be estimated,
- the underwater environment brings reduced information for measuring signals characterization.

The transient signals caused by shocks of the oxygen tank of diver are distinguished by a short cyclic duration in time. They also have an oscillatory shape yielding high peaks in their spectrum.

The measuring signals of disturbances caused by shocks of the oxygen tank of diver naturally need to be processed with time-frequency/scale tools in order to localize them in the time domain.

Modern techniques in signal processing tend to greatly improve the quality of detection: higher order statistics (HOS), wavelet transform etc. The problem of constructing a robust feature space for automatic classification of transient signals has been considered in [7]. An approach to acoustic signal monitoring based on the generalized probabilistic descent method has been described in [9]. A model-based scheme for feature extraction and signal identification which uses likelihood criteria for transient detection has been discussed in [5]. The multiresolusion signal decomposition based on wavelet transform has been reported in [2,3,6,8].

In this paper a novel techniques of measuring signal analysis combining of wavelet decomposition and multivariate data analysis has been proposed.

2. MULTIRESOLUTION WAVELET DECOMPOSITION

The wavelet transform is a time-scale representation based on multiresolution signal decomposition. It was chosen for the work reported here since it has the ability to localize the frequency changes of a non-stationary signal and it is free of any assumptions regarding the statistical characteristics of the signal.

The theory of multiresolusion wavelet decomposition has been described in [4]. A discrete signal x_n (n=1,2,…,N) with $N=2^J$) is decomposed into its coarse $A_{2^j}x$ and detail $D_{2^j}x$ components in an iterative fashion. $A_{2^j}x$ is the discrete approximation of x_n and $D_{2^j}x$ represents the signal detail at resolution 2^j. It is worth noticing, that $D_{2^j}x$ is the difference of the information between the approximation of the x_n at the resolution 2^{j+1} and 2^j, i.e. between $A_{2^{j+1}}x$ and $A_{2^j}x$. The decomposition is equivalent to a series of filter bank decomposing a complex signal x_n into multiple sub-band components. Denoting the decomposition of x_n at the scale 2^j (j=1,2,...,J) by column vector $\mathbf{d}_j$ the wavelet representation of x_n vector **W**x consists of all concatenation of these decomposition

$$\boldsymbol{W}x = \left[A_J^T,\, D_J^T, D_{J-1}^T, \ldots, D_1^T,\right]^T \tag{2}$$

with $dim(\boldsymbol{A}_J)=1\text{x}N/2^J$ and $dim(\boldsymbol{D}_j)=1\text{x}N/2^j$ for $j=1,2,...,J$.

Since the subsampling by 2 has been used in wavelet transform, it is useful to consider the length of segment of measuring signal as $N=2^J$. In this case, there are total of J+1 levels of resolution, containing respectively: 2^N points, 2^{N-1},..., 2^0=1 points.

In this paper, by using the feature sets associated with transient signals as input, the process of classification segments of wavelet coefficients has been presented.

3. SIGNAL SEGMENTATION

The idea of measuring signal segmentation consists in finding the partitions of wavelet decomposition with respect to a given criterion. This time partition is adapted to the characteristics of the signal we want to highlight (transients, which may be interpreted as disturbances in underwater environment caused by shock of the oxygen tank of diver).

A number criteria for identifying discontinuities in a signal have been described in the literature. The criterion that has been formulated in this paper based on measure of statistical similarity two vectors in the feature space of measuring signal.

Let $\mathbf{w}=[w_1, w_2,...w_{J+1}]$ (where J – number of levels of decomposition of signal) to be a vector in a feature space of signal to be analyzed and j to be an index associated with $\mathbf{w}$. The clustering criterion has been expressed by definition.

Definition

Two vectors $\mathbf{w}_1$, and $\mathbf{w}_2$ belongs to the same cluster of the homogeneous segments in the measuring signal, if

$$D^2 = \left(\boldsymbol{w}_1 - \boldsymbol{w}_2\right)^T C^{-1}\left(\boldsymbol{w}_1 - \boldsymbol{w}_2\right) \le \chi^2_{df,1-\alpha} \qquad (2)$$

where:

D^2 - Mahalanobis distance,

$\mathbf{C}$ - covariance matrix,

$\chi^2_{df,1-\alpha}$ - quartile of χ^2 distribution with df degrees of freedom on 1-α confidence level.

They are three general motivations behind the use of formula (2) in signal segmentation. These include procedures to: (1) extract the valid data set in feature space for ambient noise data, (2) extract the valid data segments due to transients caused by shock of the oxygen tank of diver, (3) extract so called outliers in data segments of measuring signal.

The first procedure is based on the property, that data vectors belongs to norm space, for which $D^2 \le \chi^2_{df,1-\alpha}$. The square Mahalanobis distance iteratively has been computed and data vectors not satisfying condition above, have been rejected.

The second procedure deals with extraction of the data vectors associated with transients. Classification results have been obtained by testing statistics

$$D^2 = \left(\boldsymbol{w}_i - \overline{\boldsymbol{w}}_{an}\right)^T C^{-1}\left(\boldsymbol{w}_i - \overline{\boldsymbol{w}}_{an}\right) > \chi^2_{df,1-\alpha} \qquad (3)$$

where:

$\overline{\boldsymbol{w}}_{an}$ - the centroid of the norm space for ambient noise data vectors,

$\boldsymbol{w}_i$ - data vector associated with transients in measuring signal.

Due to deviation $\boldsymbol{w}_i$ from the centroid of the norm space $\overline{\boldsymbol{w}}_{an}$ the data vectors have been classified as associated with transients caused by shock of oxygen tank of diver. The third procedure deals with vector data validation. It is based on assumption, that valid data vectors have been concentrated in feature space restricted by the square of Mahalanobis distance from the centroid of the space on given significance level. That means, that all data vectors for which

$$D^2 = \left(\boldsymbol{w}_i - \overline{\boldsymbol{w}}_{dv}\right)^T C^{-1}\left(\boldsymbol{w}_i - \overline{\boldsymbol{w}}_{dv}\right) > \chi^2_{df,1-\alpha} \qquad (4)$$

where

$\overline{\boldsymbol{w}}_{dv}$ - the centroid of the norm space for transients data vectors,

have been rejected as outliers caused by confusers influence on measuring environment.

4. FEATURE EXTRACTION

The results of the wavelet decomposition of the measuring signals have proved, that the information on transients is contained in the variance of the wavelet coefficients $\boldsymbol{D}_j$. Once determined, these parameters computed for wavelet coefficients have been used to classify the segments of measuring signal.

The three steps of feature extraction in the measuring signal are as follows:

Step 1. Computation of $\boldsymbol{D}_j$ by multilevel decomposition of the measuring signal.

Step 2. Computation of the variance of wavelet coefficients on each level decomposition.

Step 3. Production of the feature vector defined as

$$\boldsymbol{w} = \left[var(D_1), var(D_2), \ldots, var(D_J)\right] \qquad (5)$$

where

$var(D_j)$ - variance of the wavelet coefficients at the j-th level of the signal decomposition.

Because of the subsampling by 2 of the wavelet coefficients, the sample sizes of the variance estimator are proportional to 2^j.

5. EXPERIMENTAL RESULTS

The experiment have been carried out using a real measuring signal recorded by single omnidirectional hydrophone in coastal zone. They concern practical situation of transient emission caused by oxygen tank of diver. The recordings were done with sampling frequencies equal to 5,12 kHz, corresponding to less than one second recording for 4096 samples. The measuring session consisted of registration of ambient noise signal (one record) and transient signal caused by shock of oxygen tank of diver (four records). The length of the records were equal to 13x4096 samples (one record consists of 13 segments of samples of measuring signal).

The 7-level decomposition of signal was done for each record. The db3 mother wavelet was used in decomposition process. An example of signal decomposition for record of transients are shown on fig.1.

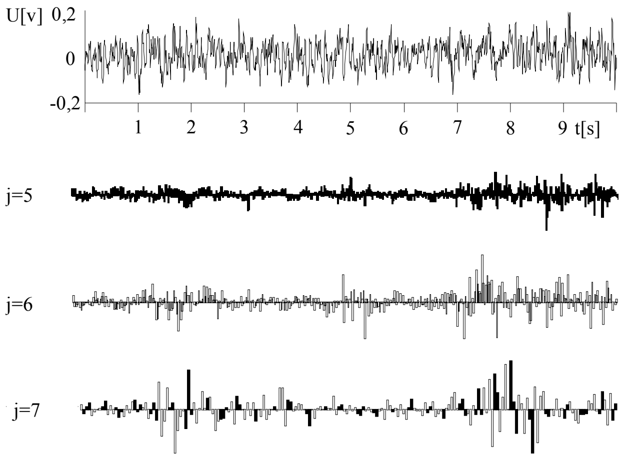

Fig. 1. Signal and wavelet coefficients on levels 5-7

As it was mentioned above, the consequence of the subsampling of wavelet coefficients are different sample sizes of the variance estimators (32,64,..., 2048).

Because of the discriminating power of the variables in feature space, the variances of the wavelet coefficients computed for second and seventh decomposition level has been used as 4-dimentional data vector in feature space of the measuring signal.

5.1. Norm space for ambient noise data

In the literature the notion of ambient noise has been defined as the residual noise background in the absence of any individual identifiable sources or that ambient noise is the natural noise of a measurement site [1]. The norm space for ambient noise data has been created from 13 data vectors using formula (3). All data vectors for which the condition (3) is satisfied belongs to the norm space on given significance level.

TABLE 1. The results of classification of the data vectors for ambient noise signal

Data vector	D^2
$\mathbf{w}_1$	5,27
$\mathbf{w}_2$	11,07
$\mathbf{w}_3$	7,58
$\mathbf{w}_4$	0,88
$\mathbf{w}_5$	7,38
$\mathbf{w}_6$	5,28
$\mathbf{w}_7$	18,62
$\mathbf{w}_8$	7,69
$\mathbf{w}_9$	1,05
$\mathbf{w}_{10}$	13,84
$\mathbf{w}_{11}$	2,16
$\mathbf{w}_{12}$	0,86
$\mathbf{w}_{13}$	6,19

The results of classification of the data vectors for ambient noise signal are shown in the table 1.

The set of data vectors indexed in table 1 by 1,3,4,5,6,8,9,11,12,13 creates norm space for ambient noise data (p=0,1, df=4, χ^2=7,78) and have been used in classification of the data vectors associated with presence of diver in underwater environment.

5.2. Classification of data vectors

The procedure of classification concerns of data vectors, which can be recognized as transients caused by diver activity in underwater environment. The problem can be formulated as a binary hypothesis testing

$\mathbf{H}_0 : D^2 \leq \chi^2_{df,1-\alpha}$ ambient noise is observed

$\mathbf{H}_1 : D^2 > \chi^2_{df,1-\alpha}$ appearance of a transient in measuring signal.

An example of data vector classification for samples record of measuring signal consisting of transients has been shown in table 2.

TABLE 2. The results of classification of the data vectors of second record of measuring signal

Data vector	D^2
$\mathbf{w}_{14}$	6,23
$\mathbf{w}_{15}$	8,94
$\mathbf{w}_{16}$	10,15
$\mathbf{w}_{17}$	7,85
$\mathbf{w}_{18}$	1,50
$\mathbf{w}_{19}$	3,49
$\mathbf{w}_{20}$	3,34
$\mathbf{w}_{21}$	2,47
$\mathbf{w}_{22}$	2,55
$\mathbf{w}_{23}$	9,94
$\mathbf{w}_{24}$	28,91
$\mathbf{w}_{25}$	9,71
$\mathbf{w}_{26}$	6,31

The vector data labeled $\mathbf{w}_{15}$, $\mathbf{w}_{16}$, $\mathbf{w}_{17}$, $\mathbf{w}_{23}$, $\mathbf{w}_{24}$, $\mathbf{w}_{25}$, has been interpreted as transients in underwater environment due to their deviation from norm space for ambient noise data. The procedure described above has been applied for four data records of measuring signal. Finally, the data vectors associated with transients in four data records of measuring signal have been extracted.

5.3. Outliers in transients

All data vectors which are the result of classification procedure have been validated for outliers rejection. In general, the outliers in observed transients may be caused by influence of different ingredients in measuring environment under registration of signal (for example man made activity in coastal zone). Therefore, after classification, it is necessary to check if the set of data vectors, derives from the same source.

The problem can be formulated as the following binary hypothesis in feature space of the measuring signal

$\mathbf{H}_0 : D^2 \leq \chi^2_{df,1-\alpha}$ diver transient is observed,
$\mathbf{H}_1 : D^2 > \chi^2_{df,1-\alpha}$ transient caused by confuser is observed.

4. CONCLUSIONS

In this paper, a signal processing method dedicated to the monitoring of the class of measuring signals, so called transient signals has been presented. This method is based on a complete process of measuring signal analysis and interpretation, including: signal analysis, features extraction and transients classification. It combines the wavelet transform of measuring signal and multivariate data analysis in feature space of signal.

The wavelet transform has been applied to extract some information in the raw measuring signals, and the resulting wavelet coefficients has been used to creation of data vectors in feature space of the signal.

The multivariate data analysis has been applied to: extraction of the valid data set in feature space for ambient noise data, extraction of the valid data segments due to transients caused by shock of the oxygen tank of diver and extraction so called outliers in data segments of measuring signal.

The described method of signal processing has been verified on the real measuring signal registered in underwater environment.

The results of signal analysis are promising and prove the efficiency of proposed method.

REFERENCES

[1] L. Bjorn, Sources of Ambient Noise in Shallow Waters. XLIV Open Acoustic Seminary, Jastrzebia Gora, Poland, pp. 35-50, September 1997.

[2] N. Hazarika, J.Z. Chen, C. Tsoi, A. Sergejew, Classification of EEG signals using the wavelet transform, Signal processing, vol. 59, no 1, pp. 61-72, 1997.

[3] S. Legendre, D. Massicotte, J. Goyette, T.K. Bose, "Neural Classification of Lamb Wave Ultrasonic Weld Testing Signals Using Wavelet Coefficients", IEEE Trans. on Instr. and Measurement, vol. 50, no 3, pp. 672-678, June 2001.

[4] G.S. Mallat, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, IEEE Transactions on Tattern Analysis and Machine Intelligence, vol. 11, no 1, pp. 674-693, 1989.

[5] S. Pycock, S. Pammu, A. Goode, S. Harman, Robust model-based signal analysis and identification, Pattern Recognition, vol. 34, no 5, pp. 2181-2199, 2001.

[6] P. Ravier, P. Amblard, "Combining an adapted wavelet analysis with fourth-order statistics for transient detection", Signal processing, vol. 70, no 2, pp. 115-128, 1998.

[7] A. Sabatini, A digital-signal-processing technique for ultrasonic signal modeling and classification, IEEE Trans. on Instr. and Measurement, vol. 50, no 1, pp.15-21, February 2001.

[8] W. Sandham, D. Hamilton, A. Fisher, W. Xu, M. Conway, „Multiresolution Wavelet Decomposition os the Seismocardiogram", IEEE Trans. on Signal Proc., vol. 4, no 9, pp. 2541-2543, September 1998.

[9] H. Watanabe, Y Matsumoto, S. Tanaka, S. Katagiri, A new approach to acoustic signal monitoring based on the generalized probabilistic descent method, IEEE Trans. on Signal Proc., vol. 47, no 9, pp. 2615-2618, September 1999.

Author: associate professor, Wieslaw Kicinski, Naval University of Gdynia, 81-103 Gdynia, Smidowicza 69, Poland, phone: 00-48 58 6262903, e_mail: wkicinski@amw.gdynia.pl

This work has been supported by Committee for Scientific Research under grant 0 T00A 016 23.