Identifying informative signs for the recognition of non-stationary signals of information-measuring systems on the base of spectral analysis and filtration on the wavelet basis

Imanova Ulkar

***Abstract*— To identify the informative signs used in the procedure of recognition of non-stationary signals of information-measuring systems are suggested to use spectral analysis based on discrete wavelet-transformation and filtrationin the wavelet display area. Informative signs are formed via processing of useful signalsobtained by filtrationand recognition signalsare produced on the base of these signs.**

***Index Terms*— discrete wavelet transform, spectral analysis, wavelet filtration, discrimination threshold**

I. INTRODUCTION

Traditional practical interest for the recognition of signals in information-measuring systems (IMS) are those methods that the predetermined conditions provide the required level of reliability of the classification by them. Until recently, the dominant approach was to build a recognition device, which does not impose restrictions on the duration of the processed signal implementation, as required by the accuracy of recognition was achieved by static processing of the results, as well as increasing the dimension of the space features. However, it is difficult to ensure these conditions by recognizing transient signals of modern IMS. An additional requirement can benecessary for decision-making under time pressure. Therefore, it is necessary to use other methods whichcan provide the desired contrast of the signals generated in the attribute space according to the stated conditions.

For this purpose,using a spectral signal analysis on the base of wavelet (1) and the selection of the useful signal by filtering in the wavelet display area (wavelet filtration) [2,3] was suggested in the paper.

Recognition signs of signalsare formed as the sum of the squares of the WC (i.e, capacity) calculated in each decompositionband for purifying signal noise. This allows usto use the given values of their deterministic (non-random) features of known recognition algorithms for the recognition of objects (signals).

II. SPECTRAL ANALYSIS OF THE SIGNALS ON THE BASE OF DISCRETE WAVELET-TRANSFORMATION

Signal $s(t)$ can be represented as [4]:

$$s(t)=\sum_{k=1}^{\infty} c_{j_0}(k)\varphi_{j_0,k}(t)+\sum_{k=1}^{\infty}\sum_{j=j_0}^{\infty} d_j(k)\psi_{j,k}(t) \quad (2.1)$$

It means that:

$$s_{j_0}(t)=\sum_k c_{j_0}(k)\varphi_{j_0,k}(t),$$

$$s_j(t)=\sum_k d_j(k)\psi_{j,k}(t) \quad (2.2)$$

where coefficients are determined by the condition of basicorthogonal functions $\varphi_{j,k}(t)$ and $\psi_{j,k}(t)$:

$$c_{j_0}(k)=\int_{-\infty}^{\infty} s(t)\varphi_{j,k}(t)dt,$$

$$d_j(k)=\int_{-\infty}^{\infty} s(t)\psi_{j,k}(t)dt \quad (2.3)$$

The equations for the basic functions:

$$\varphi(t)=\sqrt{2}\sum_{n\in Z} h(n)\varphi(2t-n) \quad (2.4)$$

$$\psi(t)=\sqrt{2}\sum_{n\in Z} g(n)\ (2t-n) \quad (2.5)$$

Where $g(n)=(-1)^n h(L-n)$, L - is the length of the matter function $\varphi(t)$, allows to obtain fast algorithms for computing the wavelet-coefficients $c_{j0}(k)$ and $d_j(k)$ [1,4]:

$$c_{j_0}(k)=\sum_m h(m-2k)c_{j+1}(m),$$

$$d_j(k)=\sum_m g(m-2k)c_{j+1}(m) \quad (2.6)$$

For Daubechies wavelets [5] $L=2M-1$ ($M \geq 1$- integer) coefficients are calculated $h(n)$ as $M \leq 7$ [6].

It is convenient to change the scale resolution (or measurementscale) of the signal $s(t)$, determining the value $j=0$ for the best resolution level $j=j_{max}$. In this case, the iterative formula of fast wavelet transformation (FWT) takes the following form:

$$c_{j+1}(k)=\sum_m h(m)c_j(2k+m),$$

Imanova Ulkar Galib. born in Lankaran in 1969. In 1992 I graduated from the Azerbaijan State University of Oil and Industry, Baku, Azerbaijan on information measurement engineering specialty and I work as an assistant on my specialty

$$d_{j+1}(k) = \sum_m g(m) c_j(2k+m) \qquad (2.7)$$

For quick calculation of the valuesof $c_j(k)$ and $d_j(k)$ coefficientsformula (2.7) is a flow diagram of the division called a pyramid or algorithm Malate [6], which is interpreted as a sequence of two way input filtering using the cascade-connected filter units of low (H) and high (G) frequencies. Convolution (digital filtering) is calculated respectivelyby the followings formulas in blocks H and G.

$$y(i) = \sum_{t=0}^{N-1} x(t) h(i-t),$$

$$y(i) = \sum_{t=0}^{N-1} x(t) g(i-t) \qquad (2.8)$$

It was suggested to improve Malate algorithm by further processing of high-frequency components of the analyzed signal pyramid in the development of the theory of a brief scale analysis (BSA) byR.Koyfman and M.Vikerhauzer [7] . This algorithm is called complete decomposition of wavelet packet decomposition and consists of some steps appropriate to m - level wavelet decomposition of the signal, $m = 1,2,\ldots,K$, $K \leq log_2 N$.

Analytically, this algorithm can be represented by the following equations:

$$Z_{m,2n}(i) = \sum_{t=0}^{N-1} Z_{m-1,n}(t) h_{m,n}(i-t),$$

$$Z_{m,2n+1}(i) = \sum_{t=0}^{N-1} Z_{m-1,n}(t) g_{m,n}(i-t) \qquad (2.9)$$

For wavelet coefficients $Z_{m,n}(i)$, the index m indicates the number of decomposition levels, n isthe index number ofsub-bands at level m. Unlike Malate algorithm (2.7), application of wavelet packetdefined by formulas (2.9) makes it possible to take into account the fine structure of the analyzed process. Indeed, the absolute values of the coefficients in a wavelet packet decomposition have smaller values in comparison with the algorithm Malate. Therefore, it can be stated that the approximation for wavelet packet has a much smaller error.

III. WAVELET FILTERING SIGNAL ADAPTATION TO THE FORM OF DYNAMICS AND NOISE

Filtering in wavelet display signal (wavelet filtering) is discrimination wavelet coefficient (WC), which provides noise removal and the spectrum of the signal. During wavelet filtration with a hard threshold WC $d_j(k)$ j -space expansion is compared with the discrimination threshold. The coefficients whose values are below than the discrimination threshold, are nulled, the value of the remaining coefficients remains unchanged:

$$\tilde{d}_j(k) = \begin{cases} d_j(k) \; where \; |d_j(k)| \geq \theta, \\ 0 \quad where \; |d_j(k)| \leq \theta, \end{cases} \qquad (3.1)$$

where $\tilde{d}_j(k)$ and $d_j(k)$ are WC on j -mlevel before and after filtration, θ - is the value of discrimination threshold.

During wavelet filtration with flexible discrimination threshold of coefficients $d_j(k)$, whose values are below the discrimination threshold, are nulled, the value of the remaining coefficients is reduced according to the moduleonthe threshold value:

$$\tilde{d}_j(k) = \begin{cases} Sign(d_j(k)) \cdot (|d_j(k)| - \theta) \; where \; |d_j(k)| \geq \theta, \\ 0 \quad where \quad |d_j(k)| \leq \theta. \end{cases} \qquad (3.2)$$

For the normal distribution of wavelet coefficients of discrimination thresholds can be obtained by Formula Kramer. The following value of discrimination thresholds is given for Gaussian white or colored noiseon the base of this formula [9] .

$$\theta = \sigma \sqrt{2 \ln N} \qquad (3.3)$$

where σ - is the standard deviationof noise.

Formula (3,3) based on the assumption of distribution normality of the signal interference, has limited scope of practical application in connection with the following circumstances:

1) In most cases filtering the signal of noise spectrum differs from the spectrum of the white noise, so standard deviations σ of the noise presented in formula (3,3) really will have various values for spaces with different levels of resolution.

2) During filtrationof the signal,interferenceusuallyis nonstationary and in relation to equations (3.3) the value is changed in time for each j -space decomposition.

3) The assumption of a Gaussian nature of the noise is not satisfied for the coefficients of the wavelet decomposition of the noise, since manifesting itself in the selection of a higher order of wavelet basis,it contradicts the characteristic of wavelet transformationof signal compression property (i.e. an increasingpeaked distributions).

Selecting the discrimination levels of wavelet coefficients that exceeding the threshold level, determines the filter quality. Discrimination levels can be determined on the base of distribution function of WC noiseand distribution function of their maximum values. Therefore, a necessary step is the statistical analysis of the WC noise. A widespread approach to wavelet filtering based on the assumption of normal distribution of the WC of signal and noise are not always adequate to the filtrationtasksof non-stationary signals. So, a necessary condition for the development of the wavelet filter algorithmis a statistical study of empirical distributions of WC signal and noise.

To ensure the removal of noise from each j - space wavelet decomposition,it is necessary to know the distribution of the maximum values of WC noise contained in a sample of a given volume N . The distribution of the maximum values of the random variable obeying a normal distribution law, was investigated by Kramer [8]. In work [2] the results [8] were obtained on the base of the exponential distribution classwith density distribution.

$$f(x) = \frac{\alpha}{2\lambda\sigma\Gamma(1/\alpha)} exp\left(-\left|\frac{x}{\alpha\sigma}\right|^{\alpha}\right), \qquad (3.4)$$

where $\Gamma(x)$ - is gamma function, α and λ are the parameters of shape and scale. For $\alpha = 2$, $\lambda = \sqrt{2}$ obtained in [2] the formula for the discrimination threshold θ coincides with known formula of Kramer (with the average values of $m = 0$), and formula (2.3) is the augend of this formula for $m = 0$.

Though exponential-power distribution class (2.4) involves some known theoretical distributions (exponential and normal), it does not cover a majority of distributions used in practice, for example, Gibrat, logistic and gamma distributions.

A common approach to the estimation of the discrimination threshold was proposedin work [3] in the case of a continuous distribution $F(x)$ of noise interference. This approach is based on the asymptotical distribution of extreme order statistics [10]. While being applied to the highest value of the signal interferencewhich is the essence of the proposed approach is as follows.

$x_1, x_2, \ldots, x_n$ means independent and identically distributed (n.o.r) random variables with the same distribution function (f.r) $F(x) = Prob\{x_j < x\}$. Assume that $z_n = max\{x_1, x_2, \ldots, x_n\}$. It's obvious that,

$$H_n(x) = Prob\{z_n < x\} = F^n(x) \qquad (3.5)$$

For an arbitrary distribution of the random variable z_n, resulting in the highest value $(z_n - a_n)/b_m$, generally speaking, will not have a limit distribution (d.f)even after appropriate normalization (centering and regulation). However, if $F(x)$ is such a limit distribution, i.e. it means there is a limit.

$$\lim_{n\to\infty} H_n(a_n + b_n x) = H(x) \qquad (3.6)$$

This limit distribution should cover three types

$$H_{1,\gamma}(x) = \begin{cases} 0, & when\ x \le 0, \gamma > 0, \\ \exp(-x^{-\gamma}), & when\ x > 0; \end{cases}$$

$$H_{2,\gamma}(x) = \begin{cases} \exp[-(x^{-\gamma})], & when\ x \le 0, \gamma > 0, \\ 1, & when\ x > 0; \end{cases} \qquad (3.7)$$

$$H_{3,0}(x) = exp(-exp(-x)), \quad -\infty < x < \infty.$$

Assume that $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$ and $g(x) = (x-a)/b$ in (2.5) and (2.6) for d.f $H_n(x)$ random variable z_n implies limiting equality $\lim_{n\to\infty} H_n(x) = H(g(x))$. Thus z_n has a limit distribution. $P(x) = \lim_{n\to\infty} H_n(x)$, where in

$$P(x) = H(g(x)), \quad g(x) = (x-a)/b, \qquad (3.8)$$

where $H(x)$ - is one of the limiting distribution (2.7), and the parameters a and b are to be determined. Formula (3.8) is written as:

$$P(x) = H\left(\frac{x-\mu}{\sigma}\right)$$

While determining discrimination threshold θ, z_n-is assumedto be the highest value of the signal interferenceand $H(x) = H_{3,0}(x)$. Non-parametric (distribution-free) methods will be usedfor the construction of point estimate of the parameters a and b . One of these methods is the method [11], based on invariantstatistics.

Knowing d.f $P(x) = F(x)$ of the random variable z_n, quantile of $F^{-1}(P_g)$ that will be the estimation of the discrimination threshold θ can be foundat a given confidence level P_g .

Being calculated for each j –space wavelet decomposition of signal thresholds θ_j are used to ensurewavelet filtrationsignalby flexible threshold method (3.2). Thus, the use of different values of the thresholds in the wavelet decomposition areasallows to adapt to the spectral characteristics of the noise and detect (i.e., track and identify) the change in the properties of non-stationary processes. This is especially important for signal processing in real time, in which algorithmprocessing should adapt to the processstate.

IV. THE FORMATION OF THE ATTRIBUTE SPACE AND THE DECISION PROBLEM OF SIGNALRECOGNITION

An integral part of the signal recognition process is to determine the feature set $x_1, x_2, \ldots, x_p$, i.e. formation of attribute space so that it can make possible to provide a desired p recognition accuracyat the minimum possible dimension. Application of the approachesfor this purpose, based on the traditional spectral-time Fourier analysis, is associated with certain difficulties (see [12]). Firstly, great demands for the input stream on the signal/noise ratio (SNR);secondly, the lack of recognition for the accuracy of the multicomponent signals and thirdly, the need for significant capacity of the realization. These restrictions are unacceptable in recognizing transient signals IMS, especially if it is necessary to take decisions under time pressure.

In view of the above mentioned observations regarding the spectral-temporal Fourier analysis, the most effective means of formationof attribute space is an approach based on spectral analysis in the wavelet basis with regard to their frequency-local properties. This analysis is related to the calculation of wavelet coefficients signal via a fast wavelet transformation (FWT) using Malate algorithm (1.7) or the algorithm of wavelet packet decomposition (1.9).

As a sign of recognition the average power of the wavelet coefficients is used in each sub-bandin [12] , being calculated according to the following equation.

$$\overline{P}_{m,n} = \frac{\sum_{i=n\cdot N/2^m}^{((n+1)\cdot N/2^m)^{m-1}} (Z_{m,n}(i))^2}{N/2^m} \ . \qquad (4.1)$$

A set of signs (4.1) reflects the content of the sub-band of recognized signal similar to the frequency representation. Moreover, transition to the average power can be used for relatively recognitionof short input realization which is an important point in the expression system-analysis. Broadband frequency caught in each sub-band will be narrowed by increasing the number of decomposition levels, which is followed by analytical formulas ofwavelet -Package (1.9).

The quality of the formed signs was estimated in [12] on the value of the separableindex of classes J, which is a ratio of spurs of matrices S_1 and S_2: $J = trS_1 / trS_2$, where the matrix S_1 characterizes the average interclass distance, and the matrix S_2 characterizes the average interclass analysis calculated in accordance with [13]. The analysis of numerical experiments carried outin [12] shows that the contrast of the image of the selected test signals on average is 4 times higher in the attribute space formed by the formulas (4.1) than on the base of the Omelchenkomethod [14], using the spectral Fourier coefficients. In addition, the differenceof total value calculation from obtained featuresof their reference values according to [13], shows that during SNRsignsbelow 20 (4.1) are 1,5- 2 times more stable with respect to the signs obtained by conventional range [14].

The recognitionprocedure of signals on the base of signs (4.1) essentially uses the normal law of distribution, which can significantly simplify the selection and identification of decision rules. As stated in [12], generally this distribution pattern was installed after a large series of 500 experiments.

Since testingnormal law of distributionsigns (4.1) made in [12] only for certain values of the SNR, in the general case such an assumption is unfounded and needs to have rigorous statistical testing in each particular case.

In this connection, appropriately the noise is isolated from the test signal via wavelet filtration (see. Section 2) using a flexible discrimination threshold (2.2) of wavelet coefficients signal (WC), and then the wavelet coefficients of the filtered signal generate deterministic signs of vector $Y = \{y_r\}_R$ with components $y_r, r = 2^m - 1 + n$ equal to the values of the calculated formula (4.1) [12], or as a vector:

$$p_j = \frac{(\sigma_{j1}, \sigma_{j2}, \ldots \sigma_{j\rho})}{\sqrt{\sigma_{j1}^2 + \sigma_{j2}^2 + \sigma_{j\rho}^2}} \tag{4.2}$$

Wher σ_{jl}^2 Pis the sum of the squares of the WC (i.e.the capacity of WC) belonging to the l sub-band of j -space wavelet decomposition.

Various recognition procedures ofnon-stationary signals can be constructed on the base of signs (3.1) or (4.2) by the known algorithmsrecognition.

V. CONCLUSION

Development method of informative signs formationis used for the recognition of non-stationary signals of information - measuring systems. The method is based on the spectral analysis using wavelet packet decomposition of the non-stationary signal and wavelet filtering adaptation to the form of noise and dynamics with arbitrary distribution.

REFERENCES

[1] Novikov L.V. Spektralnyiy analiz signalov v bazise veyvletov // Nauchnoe priborostroenie, 2000, t.10, #3, s. 70-77.

[2] Merkusheva A.V. Filtratsiya nestatsionarnyih signalov (rechi) v veyvlet-oblasti s adaptatsiey k vidu i dinamike shuma // Nauchnoe priborostroenie, 2003, t. 13, #2, s. 73-87.

[3] Dyishin O.A., Musaev G.A., Musaeva S.A. Veyvlet-filtratsiya fonogramm s adaptatsiey k vidu i dinamike shuma. – V sb.: Aktualnyie problemyi sudebnoy ekspertizyi, kriminalistiki i kriminologii. Baku, 2005, vyip. 41, s. 118-130.

[4] Dremin I.M., Ivanov O.V., Nechitaylo V.A. Veyvletyi i ih ispolzovanie // Uspehi fizich. nauk, 2001, t. 171, #5, s. 465-501.

[5] Dobeshi I. Desyat lektsiy po veyvletam / Per. s angl. Moskva-Izhevsk, NITs «Regulyarnaya i haoticheskaya dinamika», 2004, 464 s.

[6] Mallat S. A theory for multistage signal decomposition: The wavelet representation // Trans. Pattern Analysis and Machine Intelligence, 1989, July, v.11, p. 674-693.

[7] Coifman R.R., Meyer Y., Wickerhauser V. Wavelet analysis and signal processing wavelets. Wavelets and their applications. Boston: Jones and Barlett, 1992, p. 153-178.

[8] Kramer G. Matematicheskie metodyi statistiki. / Per. s angl. – M.: Mir, 1975, 648 s.

[9] Donoho D.L., Johnstone I.M. Ideal Spatial Adaption by Wavelet Shrinkage // Biometrika, 1994, v. 81, p. 425-455.

[10] Galambosh Ya. Asimptoticheskaya teoriya ekstremalnyih poryadkovyih statistik / Per. s angl. – M.: Nauka, 1984.

[11] Mordvinov V.Yu. Opredelenie doveritelnyih intervalov pokazateley nadezhnosti mashin po tsentrirovannyim vyiborkam // Problemyi mashinostroeniya i nadezhnosti mashin. 1992, #2, s. 56-64.

[12] Dvornikov S.V., Saukov A.M. Metod raspoznavaniya radiosignalov na osnove veyvlet-paketov // Nauchnoe priborostroenie, 2004, t.14, #1, s. 85-93.

[13] Fukunaga K. Vvedenie v statisticheskuyu teoriyu raspoznavaniya obrazov. – M.: Nauka, 1974, 415 s.

[14] Omelchenko V.A., Matetskiy E.O., Balabanov V.V., Bazruk V.M. Raspoznavanie sluchaynyih signalov po spektru // Izv. vuzov. Radioelektronika, 1979, t.22, #12, s. 16-22.

Imanova Ulkar Galib. born in Lankaran in 1969. In 1992 I graduated from the Azerbaijan State University of Oil and Industry, Baku, Azerbaijan on information measurement engineering specialty and I work as an assistant on my specialty. The direction of her scientific work - on information measurement. Scientific direction area - on information measurement