10th International Conference
″Research and Development in Mechanical Industry″
RaDMI 2010
16 - 19. September 2010, Donji Milanovac, Serbia

COMPUTER SYSTEM FOR IDENTIFICATION OF WATER-ETHANOL BINARY MIXTURES

Nicolay Shopov[1], Raycho Ilarionov [2], Ivan Simeonov [2], Hristo Kilifarev[2], Hristo Ibrishimov[2]
[1] University of Food Technologies, Plovdiv, BULGARIA, nikshop@abv.bg
[2] Technical University of Gabrovo, Gabrovo, BULGARIA, ilar@tugab.bg

Summary: *In the current paper is proposed a method for reading of the concentration of alcohol solutions, based on the echolocation method for obtaining of initial information and the spectral methods for image recognition. A classical classifier is used with the rule of "k-nearest neighbors" with wavelet forming of symptoms space. The base advantage of the proposed method is that it is non-contact, fast and with easy technical realization.*

Keywords: *Ultrasonic, Wavelets, Patern recognition, Ethanol-water mixtures.*

1. INTRODUCTION

The quality and harmless of the foods is a problem with constant increasing importance for the European foods market. This determines the necessity from the developing of contemporary technical means for analysis of the foods in their production. The analysis is essential part from the quality control. The control protects the user from the consumption of a product with lower quality or a product which is hazardous for the health; protects the manufacturer from issues and expenses for confiscation of the product from the market network; helps for finding of the changes in the quality and influences in the manufacturing process.

In many cases in the food industry is necessary to be made analysis of the content of different liquids. In this field are known from long ago the potentiality of the ultrasonic waves for determining of the content of different liquids.

The ethanol (ethyl alcohol - C_2H_5OH) is known in the practice as usual or vinegar alcohol (spirit of wine). It is liquid without color with specific smell and hot taste with relative density 0.79, boiling temperature 78.3°C and freezing temperature -117°C. Its lower boiling temperature according to this of the water is used at the processes distillation and rectification for its separation. In nature occurs as a product of fermentation processes of carbohydrates in fruits and cereals saccharified by the action of yeast. In the industry the ethanol can be produced also by catalytic hydratation of ethylene. The ethanol is produced mainly by two methods. Ethanol by fermentation have wide use in the food industry, perfumery, pharmacy, medicine. The produced by synthetic method ethanol is widely used the industry as resolvent and for producing of many other products. The synthetic ethanol is used for technical purposes because it contains undesired admixtures. Different methods for identification of alcohol solutions are known – aerometric, ebulliometry, spectroscopy in the near infrared field and etc.

In [7] is proposed a method for obtaining of the concentration of ethanol in sugar into water solutions with appliance at the process of fermentation.

At applied ultrasonic methods is used the purpose of the ultrasonic waves at their spreading along the continuous liquid system to interact with it and as result to change their parameters – amplitude and phase. The change of these parameters of the ultrasound depends from the properties of the medium [1, 2], i.e. carry information for the properties of the system (researched medium).

In [2] are presented results from experiments showing dependence of the speed and the attenuation of the ultrasonic pulses from the concentration of ethanol in water solutions. The results shows dependence on the two parameters from the ethanol concentration, but the speed gives higher sensitivity at low concentrations.

The authors researches, based on the above results, are directed to the opportunity for determining of the content of liquid foods on the base of analysis of the shape of the reflected ultrasonic signal. The analysis of the shape is made with the help of the contemporary direction in spectral analysis – the wavelet transforms.

2. TECHNICAL REQUIREMENTS

The ultrasonic diagnostics of food products have many advantages over traditional analytical methods - they are fast, nondestructive, accurate, enabling the automation and others. One of the methods for testing of substances is based on the relationship between parameters of ultrasonic wave and the composition of a given product in its spreading along or its reflection from it. The main feature of ultrasonic waves is their high frequency, allowing to be broadcast in the form of a narrow beam of radiation. The nature of distribution and absorption of ultrasonic waves in gases, liquids and solids can be obtained valuable information about their properties and structure.
There are two main functional schemes for the construction of an ultrasonic sensor system for non-contact measurement:
- Sensors for transmission and reception of acoustic waves are on the same side of the medium for recognition - echolocation method;
- Sensors are mounted at a certain distance, pointing at each other and the medium for recognition is on the middle between them.
The echolocation ultrasound method is based on the transmission of high frequency package of pulses from the *transmitter*, which is spread along the medium to the material for identification, reflects from it and returns as an echo to the second sensor - *receiver*. Fluctuations of pulse are scattered into the material and the resulting reflected signal is measured, transformed and analyzed. The spectrum of the reflected signal depends on the density and homogeneity of the material. This is due to factors as scattering, absorption, reflection, refraction and ray divergence. The absorption of ultrasonic energy from the medium (the tested product) in which the ultrasound is spread depends on the characteristics of that medium (density and elasticity) and also from the frequency of the used ultrasound. Consequently, the ultrasonic signal is modified and carries information about the characteristics of the medium between the transmitter and receiver - Figure 1. This gives an opportunity for identification of the composition of substances in qualitative aspect.
The computer system for processing of the information from the sensors has to extract and analyze the useful information for the identification purposes.

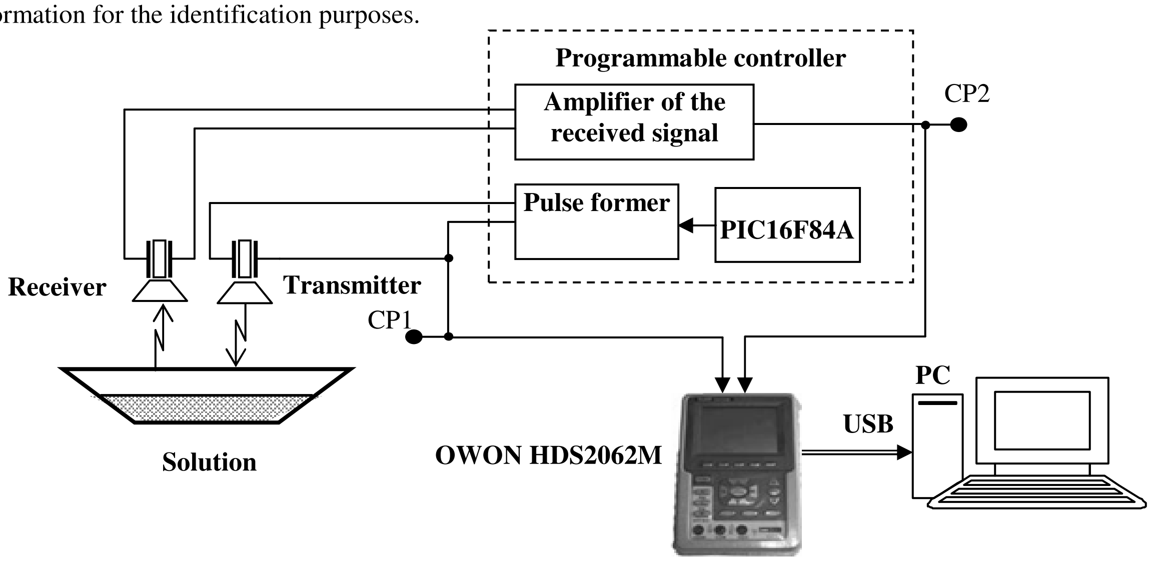

Figure 1: Scheme of the experimental setup

The experiments to obtain information were made with ultrasonic sensors type UST40T/UST40R with 40kHz frequency from *Nippon Ceramic* company.
in Figure 1. It consists from the following ***Hardware:***

- PC configuration with Microsoft Windows XP;
- microcontroller PIC16F84A - to form the ultrasonic pulses;
- oscilloscope OWON HDS2062M - used for recording and entering information into the PC;

and ***Software:***

- program for generating a pulse modulated signal (40 kHz) - loaded into PIC16F84A for initiating of the transmitter UST40T;
- OWON Wave analyzing software Version 6.6.03 - software to transfer measurement data from the oscilloscope to the PC;

- Developed software modules for processing of the results in the programming environment Matlab R2007b.

The first stage of transformation of input information is the process of analogue-digital conversion of information from the receiver. In the researches this stage is realized by using the built-in ADC into the digital oscilloscope OWON HDS2062M. In the same time are realized also functions for recording of the information and conversion according the standard USB for entering of the information into the computer system. These transformations are necessary but insufficient to extract potentially useful information for identifying the product. This requires the implementation of transformation related to change in the description area of the image. The choice of the symptoms space suitable for the concrete task is driven mostly by the technical device, which is a consumer of the information - the classifier.

The wavelet transform is one of the most promising and emerging methods for analyzing of signals. It appears to be intermediate between the full spectral and temporal presentation.

The wavelets [5] are functions with soliton like form located on the axis of the independent variable (*t* or *n*), with an option to moving along it and zooming (compression/expansion) and having a shape of short wave packets with zero mean of the wavelet function $\psi(n)$, i.e.:

$$\int_{-\infty}^{\infty} \psi(n)dn = 0 \quad \mathbf{(1)}$$

The bases of the wavelet transformation are two continuous and integrable functions on the axis of the independent variable:

- $\psi(n)$ - wavelet function giving the details of the signal and forming the detailizing coefficients;
- $\varphi(n)$ - scaling function obtaining the approximation of the signal and forming the approximating coefficients. φ-functions are the domain of the orthogonal wavelets - of Haar, Daubechies, Coiflet and others. [3, 5, 6].

The basis function $\psi_0(n)$ have to satisfy (1) and the implementation of the operations:

- moving along the axis of the independent variable:

$$\psi_0(n-b), \quad \mathbf{(2)}$$

where b is real number ($b \in R$) defining the location of the wavelet packet.

- scaling:

$$a^{-1/2}\psi_0\left(\frac{n}{a}\right), \quad \mathbf{(3)}$$

where $a \in R, a > 0$ - defining the thickness of the wavelet packet.

With respect to the requirements (2) and (3):

$$\psi(n) = a^{-1/2}\psi_0\left(\frac{n-b}{a}\right). \quad \mathbf{(4)}$$

At discrete values: $a = 2^m$ and $b = k\,2^m$, where *k* and *m* – integer numbers, dependence (5) adopt the form:

$$\psi_{m,k}(n) = 2^{-m/2}\psi_0\left(2^{-m}n - k\right). \quad \mathbf{(5)}$$

In general form the output signal on a level *m* is represented by the expression:

$$U(n) = \sum_{k=-\infty}^{\infty} A_{m,k}\varphi_{m,k}(n) + \sum_{j=1}^{m}\sum_{k=-\infty}^{\infty} D_{j,k}\psi_k(n)\,. \quad \mathbf{(6)}$$

At present are known great number of orthogonal basis wavelet functions (Figure 2) - the Haar, Daubechies, Coiflet, Symlet, etc., which has the ability to fast transformation [5]. It is realized on the base of an iterative algorithm according to the scheme shown in Figure 3.

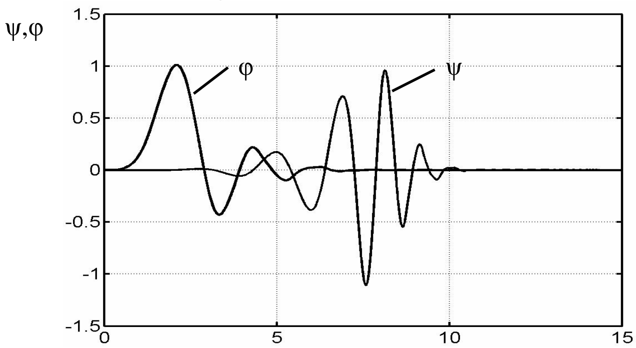

Figure 2: Wavelet (ψ) and scaling (φ) basis functions of the wavelet Daubechies 8

The initial signal passes through two filters - high frequency filter G and low frequency filter H. It was found that every second discrete value can be eliminated without loss of information. The operation is called reduction of discrete values (binary decimation). There are obtained two discrete sequences - the coefficients of approximation at m=1 $a_{1,k}$ from the filter G and detailizing coefficients $d_{1,k}$ from the filter H. - Figure 3. With decomposition at a higher level the approximating coefficients at level m=1 ($a_{1,k}$) are subjected to similar operations according the scheme of Figure 3.
The detailizing coefficients $d_{1,k}$ reflect mainly the high frequency noise, while the approximation coefficients $a_{1,k}$ reflect the characteristics of the original signal. The process of decomposition can be performed sequentially (pyramidal algorithm of Mallat [3, 6]), where the signal is decomposed of many low frequency and high frequency components. The obtained as a result wavelet coefficients are applied and researched as a symptoms for recognition. Used is one of the most frequently used parametric classifiers - kNN [4].

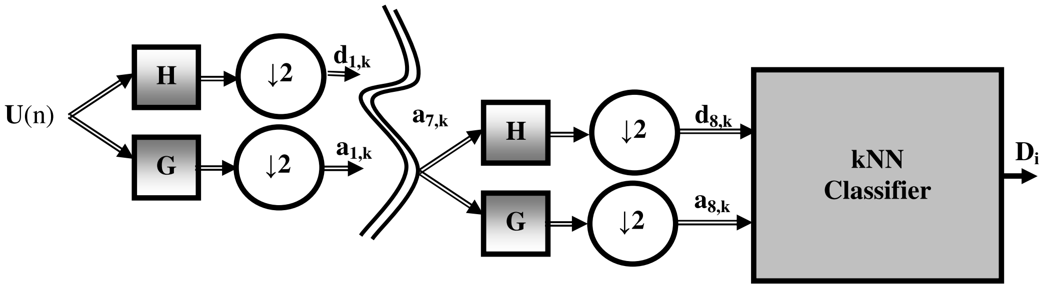

Figure 3: Block diagram of a kNN classifier with wavelet forming of symptoms space

Researched are clean water-alcoholic solutions containing ethanol, respectively 20%, 40%, 60% and 80% with total of 436 measurements with the distance between the transmitter and the medium for research - 25cm at ambient temperature of 20°C. After discretization of the reflected ultrasonic wave are obtained 2500 discrete values for the amplitude. From the discrete data obtained for the researched alcoholic solutions are formed learning and control excerpts. Training sample size is 250 numbers and is formed by the method of self-random excerpts with re-selection through random selection. This sample was used for synthesis of classifiers. Efficiency of the classifiers was tested using an independent control sample (246 numbers) which includes all conversions outside the learning sample. The workflow is by the structural diagram shown in Figure 4.

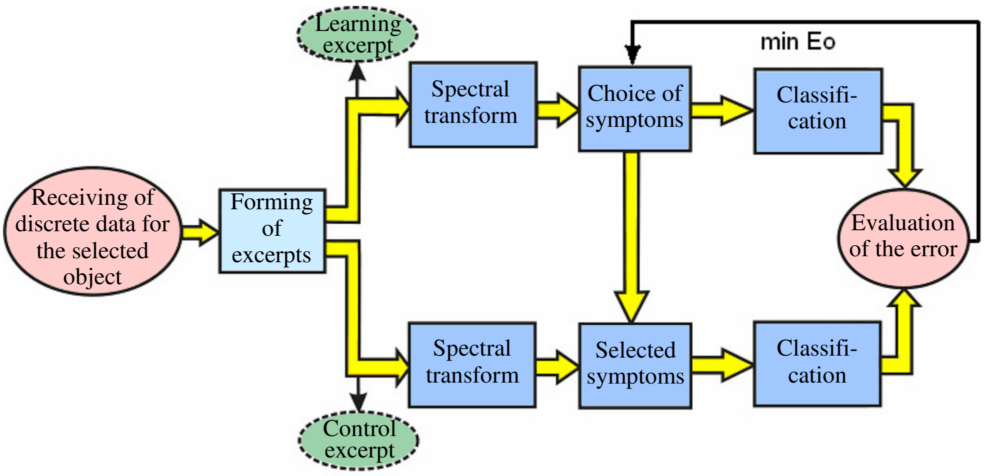

Figure 4: General structure of the process "synthesis of a classifier" with application of spectral methods

With the help of the software Matlab are obtained the wavelet approximating and detailizing coefficients of the all samples on the levels from m=1 to m=8 with the application of DWT with the orthogonal wavelets of Haar, Daubechies, Coiflet and Symlet. The wavelet spectrums of the approximating and detailizing coefficients obtained by applying of the DWT with wavelet of Haar are presented respectively in Figures 5 and 6.
Besides transformation of the image in Discrete Wavelet transformation takes place and reduction of the volume of input information, while retaining its value for identification purposes. For example, after reaching the level of m=8 as a result is obtained symptoms space of 10 approximating and 10 detailizing coefficients (Figures 5 and 6).
The obtained in this way wavelet coefficients are used as symptoms for synthesis of automatic classifiers for classification of the researched solutions.

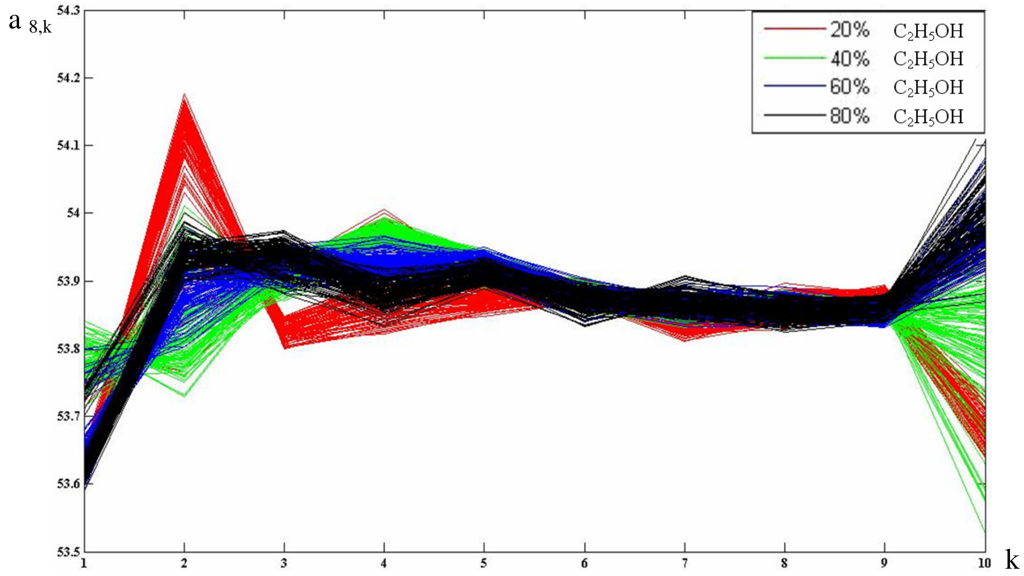

Figure 5: Spectrum from approximating coefficients obtained with wavelet of Haar at level m=8

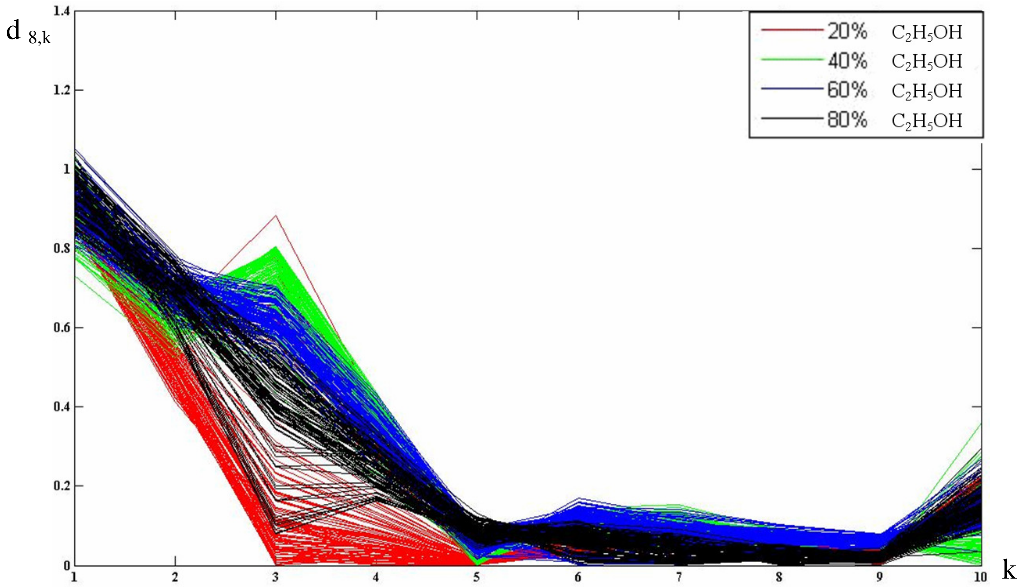

Figure 6: Spectrum from detailizing coefficients obtained with wavelet of Haar at level m=8

Selected are automatic classifiers with crucial functions by the method of "k-nearest neighbor" (kNN), taking into account the Euclidean distance to the three nearest neighbors. With the selected classifiers is performed further reduction using the method of consequent rejection where initially are selected all *k* symptoms and then removed one by one, while with the left multitude is performed classification. At each step is eliminated the symptom which leads to higher overall error from classification E_0. The condition for the end of the procedure is reaching of the lowest value of E_0 (min E_0). In the Figures 7 and 8 are shown the clusters of the four solutions into symptoms spaces of three from the obtained five "the best" symptoms with wavelet of Haar at level m=8.
The evaluation of the accuracy of the synthesized automatic classifiers is made after the classification of the control excerpt with the formed symptoms, as the lowest total error is obtained with Haar wavelet at level m=8. The results are summarized in Table 1.
The errors in classification have the following meaning:
– Highway (major) error e_i shows the relative share of objects from class *i*, referred incorrectly by the classifier into other classes.
– The marginal (actual) error g_i indicates the relative share of the objects from other classes incorrectly referred by the classifier into a given class *i*:
- The general error E_0 - shows incorrectly classified objects relative to all objects from the excerpt.

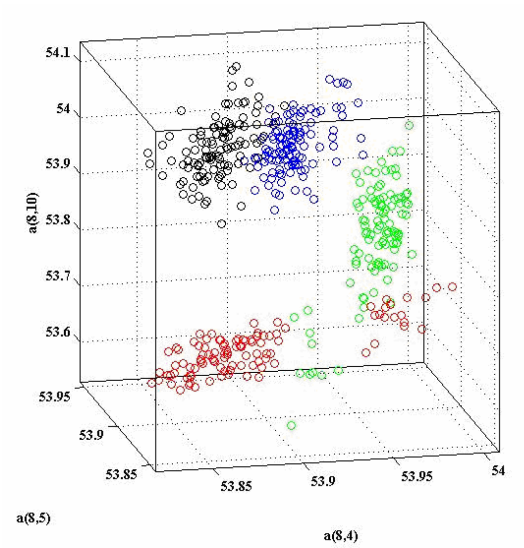

Figure 7: Clusters from approximating coefficients, obtained after DWT with wavelet of Haar at level m=8

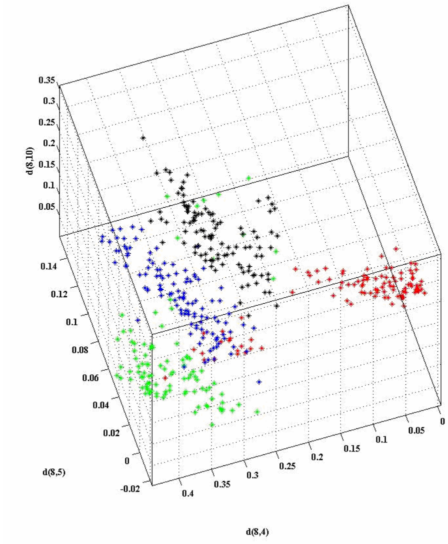

Figure 8: Clusters from detailizing coefficients, obtained after DWT with wavelet of Haar at level m=8

Table 1: Results from classification of the control excerpt (246 numbers) with kNN classifier and symptoms $a_{8,k}$, κ= 4, 5, 7, 8, and 10, obtained after DWT with wavelet of Haar at level m=8

Material		**Classified by the classifier, numbers**					**Errors**	
		20% C_2H_5OH	40% C_2H_5OH	60% C_2H_5OH	80% C_2H_5OH	Total	Marginal	Highway
Real	m_{ik}	m_{i1}	m_{i2}	m_{i3}	m_{i4}	numbers	g_i,%	e_i,%
20% C_2H_5OH	m_{1k}	53	2	0	0	55	1,85	3,64
40% C_2H_5OH	m_{2k}	1	57	2	0	60	3,39	5,00
60% C_2H_5OH	m_{3k}	0	0	69	3	72	8,00	4,17
80% C_2H_5OH	m_{4k}	0	0	4	55	59	5,17	6,78
Total	numbers	54	59	75	58	246	General error E_0=4,88%	

3. CONCLUSION

Proposed is a non-contact ultrasonic method for identification of four water-alcoholic solutions. From the results summarized in Table 1 can be seen that with the used method to identifying of water-alcoholic solutions is obtained an error 4.88%. This error is relatively small and allows for the practical use of the method. To reduce this error it is necessary to make new measurements of a large number of water-alcoholic solutions which are with less difference in the percentage concentration from 20. The applied wavelet approach for obtaining of symptoms for classification is with strictly defined mathematical procedure, avoiding the subjective factor in the formation of heuristic symptoms. Thus it is possible to automate this whole process of measurement, mathematical data processing and classification by selected criteria, without the subjective human intervention.

REFERENCES

[1] Dukhin, A.S., Goetz, P.J., "Ultrasound for characterizing colloids. Particle sizing, Zeta potential, Rheology". Elsevier, 2002

[2] Dukhin, A.S., P.J. Goetz, B.Travers, Ultrasound for characterizing liquid based food products. 1. Acoustic Spectroscopy., Dispersion Technology Inc., USA, 2003

[3] Mallat S., ''Multiresolution approximations and wavelet orthogonormal bases of L2(R),'' Trans. Amer, Math. Soc., 315:69--87, 1989

[4] Marques de Sa J.P. Pattern recognition. Concepts, methods, and applications., Springer, 2001

[5] Smolencev N. N., Base wavelet theory. Wavelets in MATLAB, DMK Press, Moscow, 2005

[6] Vetterli M., Wavelets, Approximation, and Compression, IEEE Signal Processing Magazine, September 2001, 59-73

[7] Schöck T., T. Becker, Sensor array for the combined analysis of water–sugar–ethanol mixtures in yeast fermentations by ultrasound, Elsevier, Food Control 21 (2010) 362–369