18. Liying M., Khorasani K. Constructive Feedforward Neural Network Using Hermite Polinomial Activation Functions. IEEE Trans. On Neural Networks, 16, No. 4, 2005, p.821–833.

19. Baumann M. Nutzung neuronale Netze zur Prognose von Aktionkursen. – Report Nr. 2/96, TU Ilmenau, 1996.

20. Fueser K. Neuronale Neteze in der Finanzwirtschaft. – Wiesbanden: Gabler, 1995.

## Authors' Information

**Yevgeniy Bodyanskiy** - Dr. Sc., Prof., Head of Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Lenin Av., 14, Kharkiv, 61166, Ukraine, e-mail: bodya@kture.kharkov.ua

*Iryna Pliss* - Ph.D., Senior research scientist, Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Lenin Av., 14, Kharkiv, 61166, Ukraine, e-mail: pliss@kture.kharkov.ua

*Oleksandr Slipchenko* - Ph.D., Senior research scientist, Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Lenin Av., 14, Kharkiv, 61166, Ukraine, e-mail: aslipchenko@kture.kharkov.ua

# DOUBLE-WAVELET NEURON BASED ON ANALYTICAL ACTIVATION FUNCTIONS

Yevgeniy Bodyanskiy, Nataliya Lamonova, Olena Vynokurova

*Abstract: In this paper a new double-wavelet neuron architecture obtained by modification of standard wavelet neuron, and its learning algorithm are proposed. The offered architecture allows to improve the approximation properties of wavelet neuron. Double-wavelet neuron and its learning algorithm are examined for forecasting non-stationary chaotic time series.*

*Keywords: wavelet, double-wavelet neuron, recurrent learning algorithm, forecasting, emulation, analytical activation function.*

*ACM Classification Keywords: I.2.6 Learning – Connectionism and neural nets*

## Introduction

Recently, in the analysis tasks and the non-stationary series processing under the uncertainty conditions computational intelligence techniques particularly hybrid neural networks are widely used. The most important tasks related to signal processing are forecasting and emulation of dynamic non-stationary states of systems in the future.

For solving such kind of forecasting problems a variety of neural network architectures including hybrid architectures are used. However they are either bulky because of their architecture (for instance multilayer perceptron) or poorly adjusted to learning process in real time. In most cases the activation functions for these neural networks are sigmoidal functions, splines, polynomials and radial basis functions.

In addition wavelet theory is widespread [1-3] and allows to recognize the local characteristics of the non-stationary signals with high accuracy. At the confluence of the two approaches, hybrid neural networks and wavelet theory, have evolved the so-called wavelet neural networks [4-18] that have good approximating properties and sensitivity to the characteristics changes of the analyzed processes.

Previous studies have proposed and described [19-21] attractive features of wavelet neuron such as technical realization, ensured accuracy and learning simplicity. At the same time the wavelet functions are incarnated either at the level of synaptic weights or the neuron output, and as a learning algorithm the gradient learning algorithm with constant step is used. For the improvement of approximation abilities and the acceleration of the learning process the present work introduces a new structure called double-wavelet neuron and learning algorithm with smoothing and approximation properties.

## Wavelet Analytical Activation Functions

Various kinds of analytical wavelets can be used as the activation functions of double-wavelet neuron. Among them we propose to use the triangular wavelet [8] and analytic wavelets generator [9], that have large spectrum of properties.

Fig. 1 shows the proposed triangular wavelet. To fulfil the standard condition $\int_{-\infty}^{\infty}\varphi(x)dx = 0$ (the main condition of wavelet existence), it is necessary to perform the condition $h_2 = \dfrac{(b-d)+h_1(c-a)}{(c-e)}$. After that we can write the mathematical expression for this function in the following form

$$\varphi(x,[a,b,c,d,e,h_1]) = \begin{cases} 0, \ if \ \ x < a \ \& \ x > e; \\ -h_1\dfrac{(x-a)}{(b-a)}, \ if \ \ a \le x \le b; \\ \dfrac{(h_1+1)(x-b)}{(c-b)} - h_1, \quad if \ \ b \le x \le c; \\ -\dfrac{(h_2+1)(x-c)}{(d-c)} + 1, \quad if \ \ c \le x \le d; \\ h_2\dfrac{(x-d)}{(e-d)} - h_2, \quad if \ \ d \le x \le e. \end{cases} \tag{1}$$

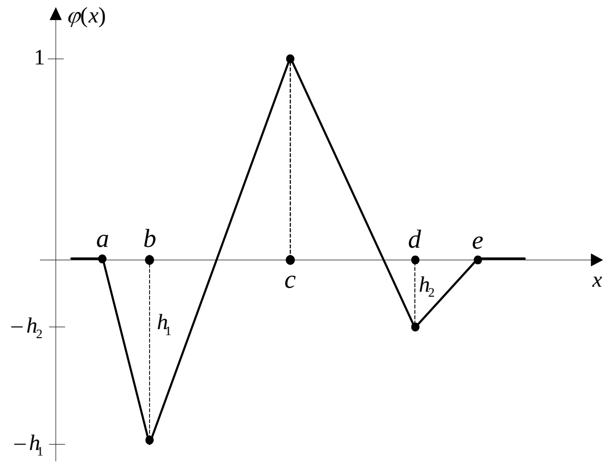



Fig. 1 – Triangular wavelet

Distinctive feature of the proposed triangular wavelet is that this function can be both even and odd, according to the values of function parameters. Such function can be implemented in chip as an activation function of wavelet neuron.

In some situation instead of triangular wavelet using a universal wavelet activation function based on analytic wavelet generator [9] is useful. As is known most of wavelets can be divided into the even and odd functions. Therefore because of the processing signal type it is necessary to choose even analytic wavelets generator

$$y_{even}(x(k)) = \sum_{i=1}^{n} a_i \cos(i\,x(k)) = a^T \varphi_{even}(x(k)) \tag{2}$$

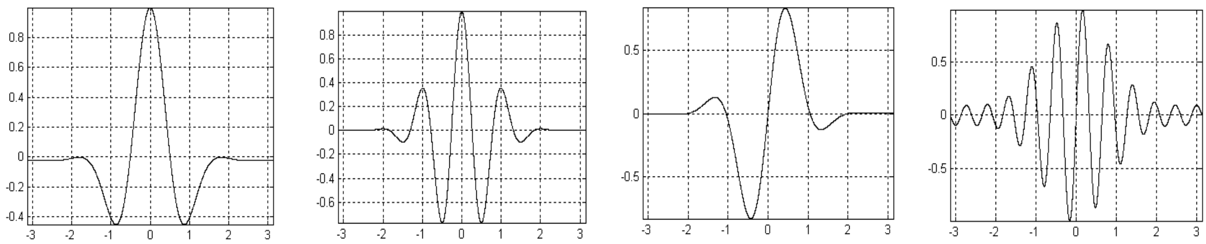
(here $\varphi_{even}(x(k)) = (\cos x(k), \cos 2x(k), \ldots, \cos nx(k))$; $a_i$ are the spectral decomposition coefficients of even wavelet) or odd analytic wavelets generator

$$y_{odd}(x(k)) = \sum_{i=1}^{n} b_i \sin(i\, x(k)) = b^T \varphi_{odd}(x(k)), \tag{3}$$

where $\varphi_{odd}(x(k)) = (\sin x(k), \sin 2x(k), \ldots, \sin nx(k))$; $b_i$ are the spectral decomposition coefficients of even wavelet.

These analytic wavelets generators allow to get the large number of wavelet functions and to tune their parameters during the wavelet neuron training.

Fig. 2 a, b show the representatives of wavelets obtained using even analytical wavelets generator (2), and fig. 2 c, d show the representatives of wavelets obtained using odd analytical wavelets generator (3).



a) b) c) d)

Fig. 2 – Representatives of wavelets that obtained using analytical wavelets generator

It is easy to see, that the situation shown on fig. 2 a is the most similar to the Mexican hat wavelet, on fig. 2 b – the Morlet wavelet, on fig. 2 c – the POLYWOG 2 wavelet, on fig. 2 d – the RASP 2 wavelet [4].

## Structure of Double-Wavelet Neuron

Fig. 3 introduces the structure of double-wavelet neuron that consists of nonlinear wavelet synapses which use analytical activation functions.

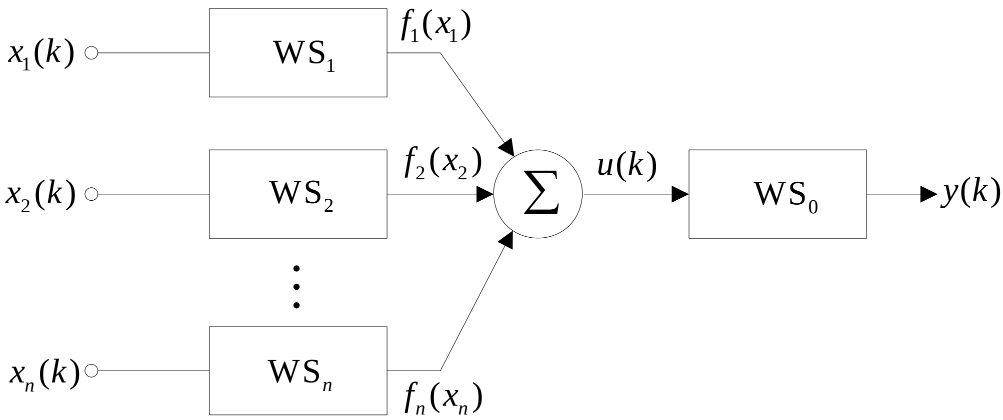



Fig. 3 – Generalized structure of double-wavelet neuron

If a vector signal $x(k) = (x_1(k),\ x_2(k),\ \ldots,\ x_n(k))^T$ (here $k = 0,\ 1,\ 2,\ \ldots$ is the number of sample in the training set or current discrete time) is fed to the input of the double-wavelet neuron shown in Fig. 4 then the output is described by the expression

$$y(k) = f_0\left(\sum_{i=1}^{n} f_i(x_i(k))\right) = f_0(u(k)) =$$
$$= \sum_{l=0}^{h_2} \varphi_{l0}\left(\sum_{i=1}^{n}\sum_{j=0}^{h_1} \varphi_{ji}(x_i(k))w_{ji}(k)\right)w_{j0} = \sum_{l=0}^{h_2} \varphi_{l0}(u(k))w_{l0}(k), \quad (4)$$

and depends from the synaptic weights $w_{ji}(k)$, $w_{l0}$ as well as by the values of the used wavelet functions $\varphi_{ji}(x_i(k))$, $\varphi_{l0}(u(k))$, assuming that $\varphi_{00}(\bullet) = \varphi_{0i}(\bullet) \equiv 1$.

The double-wavelet neuron is composed of two sublayers: hidden layer that contains *n*-wavelet synapses with $h_1$ wavelet-functions in each and output layer that contains one wavelet-synapse with $h_2$ wavelet-functions.

In each wavelet-synapse, the wavelets that differ between each other by dilation, translation and bias factors are realized.

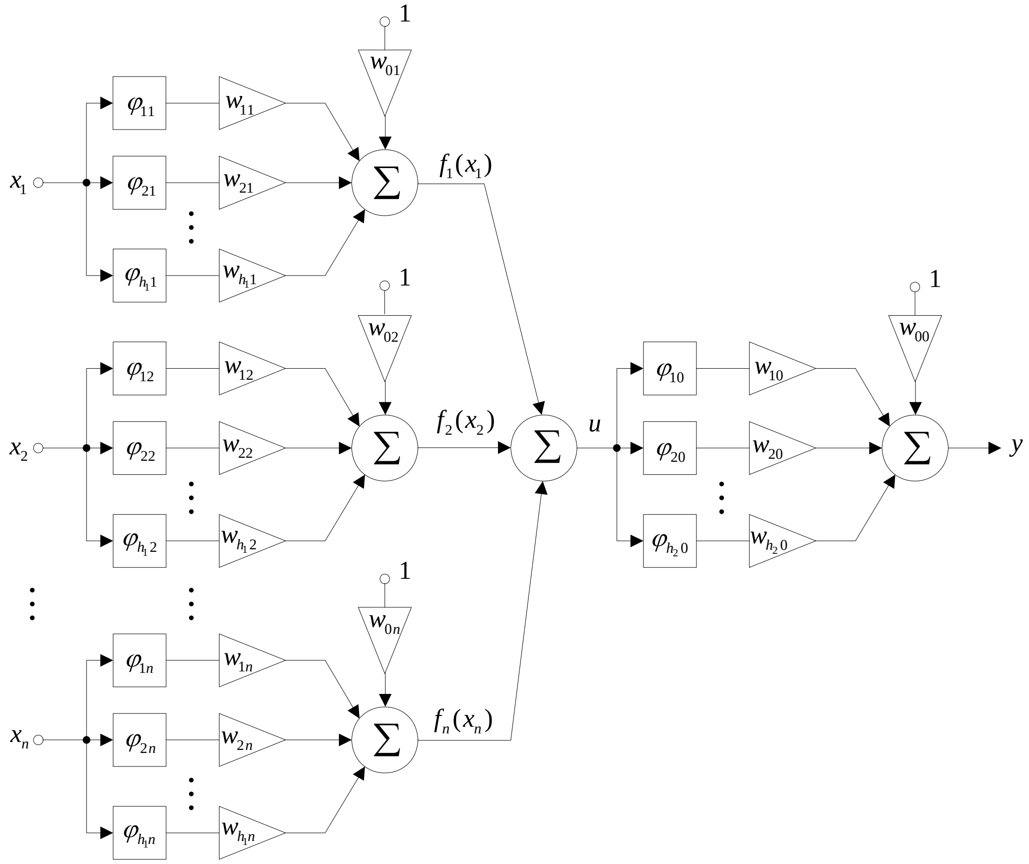



Fig. 4 - Architecture of double-wavelet neuron with nonlinear wavelet-synapses

## Synthesis of Double-Wavelet Neuron Learning Algorithm

For the tuning of the output layer of double-wavelet neuron we shall use the one-step criterion

$$E(k) = \frac{1}{2}(d(k) - y(k))^2 = \frac{1}{2}e^2(k), \quad (5)$$

where $d(k)$ is the external training signal.

The learning algorithm for the output layer of double-wavelet neuron on the basis of gradient approach can be written as

$$w_{j0}(k+1)=w_{j0}(k)+\eta_0(k)e(k)\varphi_{j0}(u(k)), \quad (6)$$

or in the vector form

$$w_0(k+1)=w_0(k)+\eta_0(k)e(k)\varphi_0(u(k)), \quad (7)$$

where $w_0(k)=(w_{00}(k),w_{10}(k),w_{20}(k),\dots,w_{h_2 0}(k))^T$ is $(h_2+1)\times 1$ vector of synaptic weights, $\varphi_0(u(k))=(1,\varphi_{10}(k),\varphi_{20}(k),\dots,\varphi_{h_2 0}(k))^T$ is a vector of activation functions, $e(k)$ is a learning error, $\eta_0(k)$ is a learning rate parameter which is subject to determination.

To increase the rate of convergence of the training process it is necessary to turn from gradient procedures to the second-order procedures.

We propose to use the following learning algorithm

$$\begin{cases} w_0(k+1)=w_0(k)+\dfrac{e(k)\varphi_0(u(k))}{\gamma_i^{w_0}(k)}, \\ \gamma_i^{w_0}(k+1)=\alpha\gamma_i^{w_0}(k)+\left\|\varphi_0(u(k+1))\right\|^2, \end{cases} \quad (8)$$

where $\alpha$ is the forgetting factor of out-dated information ($0\le\alpha\le 1$).

This algorithm has both the smoothing and approximating properties. The tuning of hidden layer is carried out in the same way on the basis of error backpropagation approach by using the same criterion written in the form

$$E(k)=\frac{1}{2}(d(k)-f_0(u(k)))^2=\frac{1}{2}\left(d(k)-f_0\left(\sum_{i=1}^{n}\sum_{j=0}^{h_1}\varphi_{ji}(x_i(k))w_{ji}(k)\right)\right)^2. \quad (9)$$

The learning algorithm for the hidden layer of double-wavelet neuron on the basis of gradient optimization has the form

$$w_{ji}(k+1)=w_{ji}(k)+\eta(k)e(k)f_0'(u(k))\varphi_{ji}(x_i(k)), \quad (10)$$

or in the vector form

$$w_i(k+1)=w_i(k)+\eta(k)e(k)f_0'(u(k))\varphi_i(x_i(k)), \quad (11)$$

where $w_i(k)=(w_{0i}(k),w_{1i}(k),w_{2i}(k),\dots,w_{h_1 i}(k))^T$ is the vector of synaptic weights, $\varphi_i(x_i(k))=(1,\varphi_{1i}(k),\varphi_{2i}(k),\dots,\varphi_{h_1 i}(k))^T$ is the vector of wavelet-activation functions, $e(k)$ is the learning error, $\eta(k)$ is the learning rate.

By analogy with (8) one can introduce the procedure

$$\begin{cases} w_i(k+1)=w_i(k)+\dfrac{e(k)\ f_0'(u(k))\,\varphi_i(x_i(k))}{\gamma_i^{w_1}(k)}, \\ \gamma_i^{w_1}(k+1)=\alpha\gamma_i^{w_1}(k)+\left\|\varphi_i(x_i(k+1))\right\|^2, \end{cases} \quad (12)$$

where $0\le\alpha\le 1$.

Besides the quadratic goal function (5) for the learning algorithm synthesis the goal function based on Trefferquote criterion, Wegstrecke criterion and hybrid criterions can be used [23-25].

## Simulation Results

Effectiveness of performance of the proposed double-wavelet neuron and its learning algorithm (8), (12) were investigated in the process of solving forecasting problem and chaotic behaviour emulation of nonlinear dynamic system described by the equation

$$x_{n+1} = \frac{5x_n}{1+x_n^2} - 0.5x_n - 0.5x_{n-1} + 0.5x_{n-2} \tag{13}$$

with initial values $x_0 = 0.2,\ x_1 = 0.3,\ x_2 = 1.0$.

The training set contained 10000 samples, and checking set – 500 samples. Double-wavelet neuron had 5 synapses in the hidden layer corresponding to 5 inputs $x(k-4),\ x(k-3), x(k-2), x(k-1),\ x(k),\ (n=5)$ with 20 wavelets in each synapse ($h_i = 20,\quad i = 1\ldots5$). Output layer consists of 5 wavelets in synapse $WS_0$. Initial values of synaptic weights were generated in a random way from $-0.1$ to $+0.1$.

Several criteria were used for the quality rating of forecast:

- mean-square error (RMSE)

$$RMSE = \frac{1}{N}\sum_{k=1}^{N}(x(k)-\hat{x}(k))^2\ ; \tag{14}$$

- mean absolute percentage error (MAPE)

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|x(k)-\hat{x}(k)\right|}{x(k)}100\%\,; \tag{15}$$

- Trefferquote [24, 25] representing percentage ratio of correctly predicted directions to actual direction of the signal

$$Trefferquote = \frac{N - \frac{1}{2}\sum_{k=1}^{N}\left|sign(\hat{x}(k)-x(k-1)) - sign(x(k)-x(k-1))\right|}{N}\cdot 100\%\,; \tag{16}$$

- Wegstrecke [24, 25], representing quality rating of the predicted model (value +1 corresponds to the optimal predictive model, and -1 – to the incorrect forecast) and described by the equation

$$Wegstreke = \frac{\sum_{k=1}^{N} signal(k)(x(k)-x(k-1))}{\sum_{k=1}^{N}\left|x(k)-x(k-1)\right|}\ , \tag{17}$$

where $signal(k)$ is a sign-function in the form

$$signal(k) = \begin{cases} 1, & if \quad \hat{x}(k)-x(k) > 0, \\ -1, & if \quad \hat{x}(k)-x(k) < 0, \\ 0, & in \quad other\ cases, \end{cases}$$

where $x(k)$ is the actual value of forecasting process, $\hat{x}(k)$ is the forecast, $N$ is the length of training set.

Fig. 5 shows the results of forecasting process on the basis of data from test set after 10 training epoch with the parameter $\alpha = 0.99$.

Table 1 shows the results of forecasting process on the basis of the double-wavelet neuron compared the results of forecasting process on the basis of standard wavelet-neuron with the gradient learning algorithm, radial basis neural network and multilayer perceptron.

Thus as it can be seen from experimental results the proposed double-wavelet neuron with the learning algorithm (8), (12) having the same number of adjustable parameters ensures the best quality of forecast and high learning rate in comparison with conventional architectures. The experimental results of multilayer perceptron and radial basis function network are worse because the multilayer perceptron had not time to be trained at such a small number of epochs and radial basis function network suffers from the curse of dimensionality.

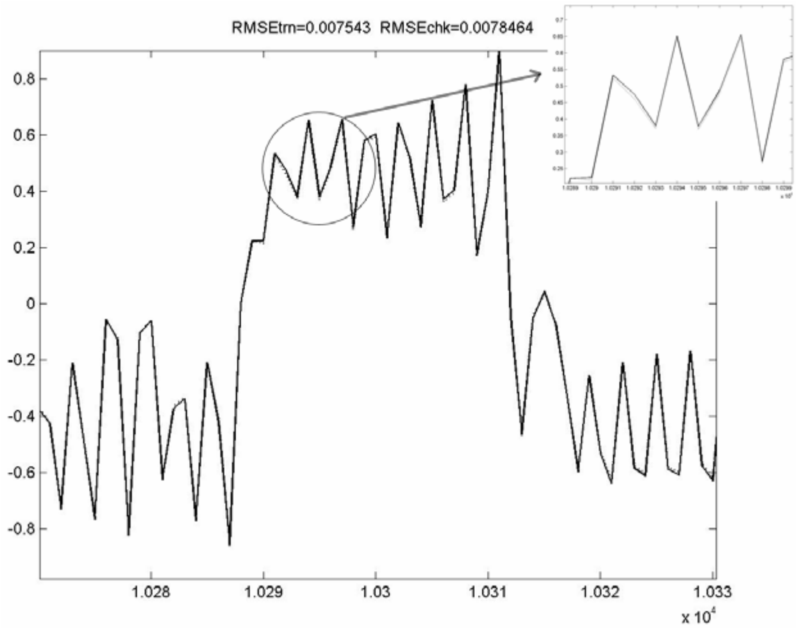



Fig. 5 – Forecasting of chaotic dynamic system behaviour using the double-wavelet neuron

Table 1 - The results of time series forecasting

| Neural network/ Learning algorithm | Number of adjustable parameters | Criteria | | | |
|---|---|---|---|---|---|
| | | RMSE | MAPE | Wegstrecke | Trefferquote |
| Double-wavelet neuron/ Proposed learning algorithm of wavelet-synapses parameters (8), (12) | 105 | 0.0076 | 1.9% | 1 | 99.8% |
| Wavelet-neuron/ Gradient learning algorithm of parameters of wavelet-synapses with constant step | 105 | 0.0101 | 3.1% | 0.98 | 98.8% |
| Radial basis function network / RLSE | 105 | 0.3774 | 7% | 0.4883 | 55,2% |
| Multilayer perceptron / Gradient learning algorithm | 115 | 0.5132 | 9% | 0.5882 | 75,5 % |

## Conclusions

The double-wavelet neuron architecture and learning algorithm which allows to adjust its parameters are proposed. This algorithm is very simple in the way of its numerical implementation, possesses high rate of convergence and additional smoothing and approximation properties.

## Bibliography

[1]. Chui C. K. An Introduction to Wavelets. New York: Academic. 1992, 264 p.

[2]. Daubechies I. Ten Lectures on Wavelets. Philadelphia, PA: SIAM. 1992, 228 p.

[3]. Meyer Y. Wavelets: Algorithms and Applications. Philadelphia, PA: SIAM. 1993, 133 p.

[4]. Lekutai G., van Landingham H.F. Self-tuning control of nonlinear systems using neural network adaptive frame wavelets. Proc. IEEE Int. Conf. on Systems, Man and Cybernetics. Piscataway, N.J. 2, 1997, P. 1017-1022.

[5]. Bodyanskiy Ye., Lamonova N., Pliss I., Vynokurova O. An adaptive learning algorithm for a wavelet neural network. Blackwell Synergy: Expert Systems. 22 (5), 2005, P. 235-240.

[6]. Bodyanskiy Ye., Kolodyazhniy V. Pliss I., Vynokurova O. Learning wavelet neuron based on the RASP-function. Radio Electronics. Computer Science. Control. 1., 2004, P. 118-122.

[7]. Vynokurova O., Bodyanskiy Ye., Lamonova N. Adaptive hybrid wavelet neural network for the solving problem of forecasting and emulation // Proc. of 12th International Conference on Automatic Control "Automatics-2005", Kharkiv: NTU "KPI". - 2005. – Vol.3. – P.40-41. (in Russian)

[8]. Bodyanskiy Ye., Vynokurova O. Triangular wavelet and the formal neuron based on it // Proc. of 3-th Int. Scient. Conf. «Mathematical support and software for artificial intelligence» (MPZIS-2005) – Dnepropetrovsk: DNU. - 2005. – P. 14-15 (in Russian)

[9]. Vynokurova O. Lamonova N., Pliss I. Analytic wavelets generator // Problemy Bioniki. – 2004. - N 60. - P. 104 – 109 (in Russian)

[10]. Billings S. A., Wei H.-L. A new class of wavelet networks for nonlinear system identification. IEEE Trans. on Neural Networks. 16 (4), 2005, P. 862-874.

[11]. Szu H. H., Telfer B., Kadambe S. Neural network adaptive wavelets for signal representation and classification. Opt. Eng. 31, 1992, P. 1907–1916.

[12]. Zhang Q. H., Benveniste A. Wavelet networks. IEEE Trans. on Neural Networks. 3 (6), 1992, P. 889–898.

[13]. Dickhaus H., Heinrich H. Classifying biosignals with wavelet networks. IEEE Eng. Med. Biol. Mag. 15(5), 1996, P. 103–111.

[14]. Cao L. Y., Hong Y. G., Fang H. P., He G. W. Predicting chaotic time series with wavelet networks. Phys. D. 85, 1995, P. 225–238.

[15]. Oussar Y., Dreyfus G. Initialization by selection for wavelet network training. Neurocomputing. 34, 2000, P. 131–143.

[16]. Zhang J., Walter G. G., Miao Y., Lee W. N. W. Wavelet neural networks for function learning. IEEE Trans. on Signal Process. 43(6), 1995, P. 1485–1497.

[17]. Zhang Q. H. Using wavelet network in nonparametric estimation. IEEE Trans. on Neural Networks. 8(2), 1997, P. 227–236.

[18]. Casdagli M. Nonlinear prediction of chaotic time series. Phys. D. 35, 1989, P. 335–356.

[19]. Soltani S. On the use of wavelet decomposition for time series prediction. Neurocomputing. 48, 2002, P. 267–277.

[20]. Yamakawa T., Uchino E., Samatu T. Wavelet neural networks employing over-complete number of compactly supported non-orthogonal wavelets and their applications. IEEE Int. Conf. on Neural Networks, Orlando, USA., 1994, P. 1391-1396.

[21]. Yamakawa T., Uchino E., Samatu T. The wavelet network using convex wavelets and its application to modelling dynamical systems. The Trans. on the IEICE. J79-A. 12, 1996, P. 2046-2053.

[22]. Yamakawa T. A novel nonlinear synapse neuron model guaranteeing a global minimum – Wavelet neuron. Proc. 28th IEEE Int. Symp. On Multiple-Valued Logic. Fukuoka, Japan: IEEE Comp. Soc., 1998, P.335-336

[23]. Bodyanskiy Ye., Lamonova N., Pliss I., Vynokurova O. Learning algorithm wavelet neuron based on combined criterion Radio Electronics. Computer Science. Control. 2., 2005, P. 83-89.

[24]. Baumann M. Nutzung neuronaler Netze zur Prognose von Aktienkursen. Report Nr. 2/96, TU Ilmenau., 1996, 113 S.

[25]. Fueser K. Neuronale Netze in der Finanzwirtschaft. – Wiesbaden: Gabler, 1995. – 437 S.

## Authors' Information

Bodyanskiy Yevgeniy – Doctor of Technical Sciences, Professor of Artificial Intelligence Department and Scientific Head of the Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Lenina av. 14, Kharkiv, Ukraine 61166, Tel +380577021890, e-mail: bodya@kture.kharkov.ua

Lamonova Nataliya – Candidate of Technical Sciences (equivalent Ph.D.), Senior Research Scientist of Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Lenina av. 14, Kharkiv, Ukraine 61166, Tel +380577021890, e-mail: webmaster@natashka.de

Vynokurova Olena – Candidate of Technical Sciences (equivalent Ph.D.), Senior Research Scientist of Control Systems Research Laboratory Kharkiv National University of Radio Electronics, Lenina av. 14, Kharkiv, Ukraine, 61166, Tel +380577021890, e-mail: vinokurova@kture.kharkov.ua