VOL. 11, NO. 1, JANUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DEVELOPMENT AND APPLICATION OF COMPOSITE LOGISTICS FUNCTIONS TO IMPROVE THE SPEED OF TRAINING OF WAVELET NEURAL NETWORKS IN SPEECH RECOGNITION SYSTEMS

Dmitry Ivanovich Popov and Artem Vladimirovich Gapochkin
Moscow State University of Printing Arts, Pryanishnikova 2A, Moscow, Russian Federation
E-Mail: damitry@mail.ru

ABSTRACT

This article suggests a new approach to handling voice communications, based on the joint application of wavelet analysis and neural networks. One of the most effective methods of speech signals currently performing wavelet analysis. The use of large-scale analysis allows estimating the speech signal as from the point of view of the spectrum and temporal variation. The advantages of wavelet neural network for the multi-layer network.

Keywords: artificial neural network, wavelet analysis, speech recognition, training neural networks.

1. INTRODUCTION

Artificial neural networks have been widely used recently to solve a wide range of problems concerning processing speech signals, since they guarantee high quality of approximation, anticipation, filtration, and other stochastic and chaotic signals. A wavelet analysis approach, which is very effective for local representation of speech signals in both time and frequency domains, was developed, along with neural networks [1]. These two approaches are used in neuro-wavelet networks that combine flexibility and ability of neural networks to train with compact description of different signals peculiar to wavelet analysis [2, 10].

2. METHODOLOGY

Multilayer neural network

Multilayer neural network (MNN) (Figure-1) includes input layer and output layer. Despite input and output layers, there are also hidden layers [8]. They consist of neurons, which are located between input layer and output layer. Neural network processes input signals and compares with output. In the case if the deviation is greater than a predetermined value, weighs of neural connections and threshold values of the neurons change as well. Again, it is necessary to calculate the output value and compare it with predetermined reference. If the deviation is less than predetermined error, the training process finishes [3].

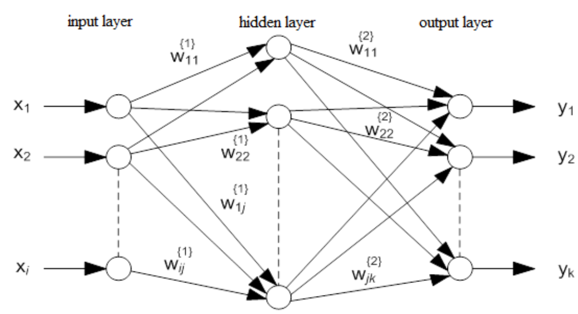


Figure-1. Example of Multilayer neural network.

Wavelet-neuron network

Wavelet-neuron network (WNN) (Figure-2) consists of one hidden layer of full value and one defective layer - linear combiner. Figure-2 shows the architecture of WNN, which allows approximating any signal and forms linear combination of secondary wavelets,

where

$$h_{a,b}(t) = h\left(\frac{t-b}{a}\right) \tag{1}$$

with expansion factor [9].

ARPN Journal of Engineering and Applied Sciences

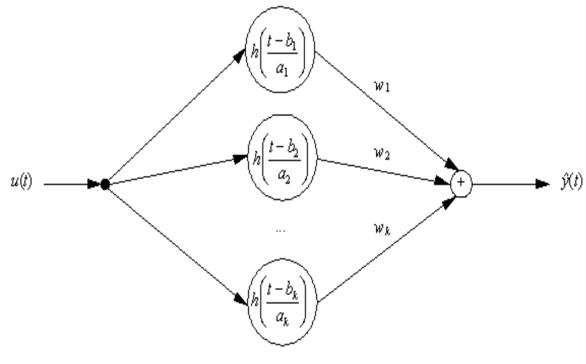


Figure-2. Example of Wavelet-neuron network.

Output signal is represented as

$$\widehat{y}(t)=u(t)\sum_{k=1}^{K}w_k h_{a_k,b_k}(t) \tag{2}$$

Where K - number of time intervals of wavelets, w_k - weigh coefficients. The task of training networks is to set w_k, a_k, b_k in order to minimize function of the energy of the error E in time *t*. The function of the error in time *t* will take the form:

$$e(t)=y(t)-\widehat{y}(t)$$

Where y(t) – the objective function. The function of energy is calculated by the formula:

$$E=\frac{1}{2}\sum_{t=1}^{T}e^2(t) \tag{3}$$

The minimization of the error is carried out by method of steepest descent that uses gradients $\frac{\partial E}{\partial w_k}$, $\frac{\partial E}{\partial a_k}$ and $\frac{\partial E}{\partial b_k}$ to renew every network parameter w_k, a_k *and* b_k respectfully.

For any primary wavelet, gradients are equal:

$$\begin{aligned}\frac{\partial E}{\partial w_k}&=-\sum_{t=1}^{T}e(t)h(\tau)u(t)\\ \frac{\partial E}{\partial b_k}&=-\sum_{t=1}^{T}e(t)u(t)w_k\frac{\partial h(\tau)}{\partial b_k}\\ \frac{\partial E}{\partial a_k}&=-\sum_{t=1}^{T}e(t)u(t)w_k\tau\frac{\partial h(\tau)}{\partial b_k}=\tau\frac{\partial E}{\partial b_k}\end{aligned} \tag{4}$$

where $\tau=\frac{t-b_k}{a_k}$.

Further calculation of every coefficient is equal to a negative meaning of all its gradients:

$$\Delta w=-\frac{\partial E}{\partial w},\quad \Delta b=-\frac{\partial E}{\partial b},\quad \Delta a=-\frac{\partial E}{\partial a} \tag{5}$$

This leads us to the conclusion that coefficient w, a, b of neural network are calculated as follows:

$$\underline{w}(n+1)=\underline{w}(n)+\mu_w\Delta\underline{w} \tag{6}$$

$$\underline{b}(n+1)=\underline{b}(n)+\mu_b\Delta\underline{b} \tag{7}$$

$$\underline{a}(n+1)=\underline{a}(n)+\mu_a\Delta\underline{a} \tag{8}$$

where *n* – constant parameter [6].

Compound logistic functions (CLOG)

Sigmoid function is a monotonically increasing, smooth function, which presents itself as the activation function in Neural Network Models (Figure-3) [3, 13]:

$$f(t)=\frac{1}{1+e^{-t}} \tag{9}$$

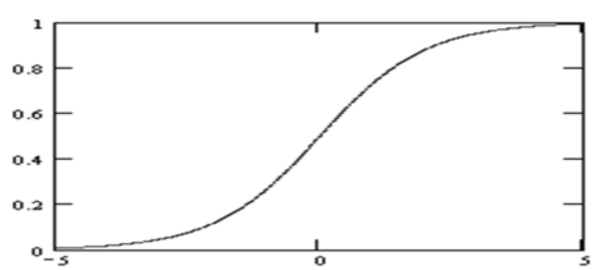
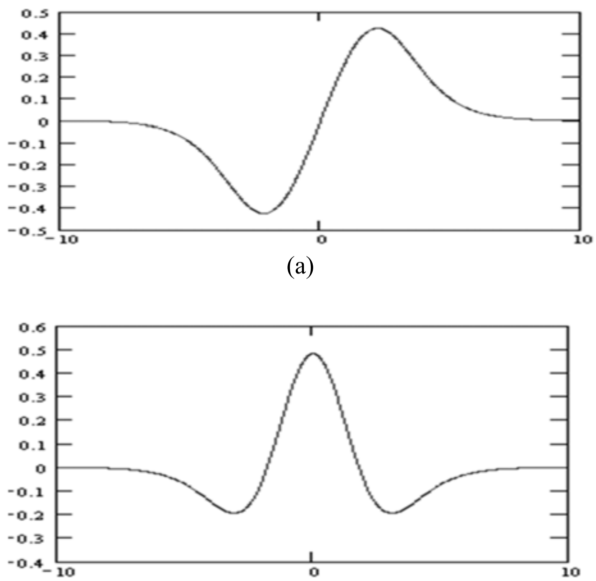


Figure-3. Logistic function.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com



(a)

(b)

Figure-4. Types of logistic functions of primary wavelets.

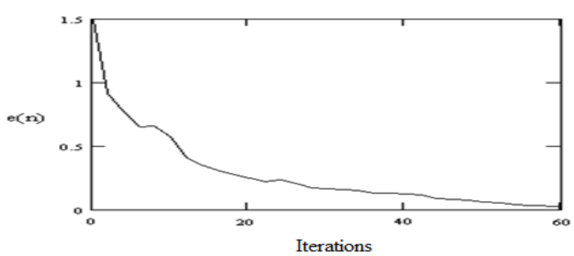


Figure-5. The graph of the error in MNN training.

CLOG is a result of the final sum of weighted logistic functions with a delay. CLOG is identical to functions of mother wavelets. Therefore, neural network can have the same approximating characteristics as wavelets [4, 6].

$$h_{\log 1}(t) = \frac{1}{1+e^{-t+1}} - \frac{1}{1+e^{-t+3}} - \frac{1}{1+e^{-t-3}} + \frac{1}{1+e^{-t-1}} \quad (10)$$

$$h_{\log 2}(t) = \frac{3}{1+e^{-t-1}} - \frac{3}{1+e^{-t+1}} - \frac{1}{1+e^{-t-3}} + \frac{1}{1+e^{-t+3}} \quad (11)$$

Figure-4 (a) displays wavelet transformation (WT) showing the sum of logistic sigmoids (2). This function of a primary WT represents the minimum fluctuation magnitude (“down-up-down”). Figure-4 (b) demonstrates another WT (3) that has complementary semi-fluctuations (“down-up-down-up”); thus, wavelets with a greater number of semi-fluctuations can also be generated [14].

Let us now examine the functions of approximation $u(t) = \cos(t) \cdot e^{-t}$ by means of MNN. Let us use MNN, which includes one output neuron and a hidden layer containing four neurons with sigmoid activation functions. Let us predetermine network training error $e = 0{,}03$. In Figure-4 we have got the graph of the training error for above-mentioned network. Multilayer network training is fulfilled during 51 iterations, and the training error is $e = 0{,}0287$. Consider approximation of the sane function by means of WNN. Let us apply neuron network containing four wavelet neurons, activation function of which is CLOG (11). Predetermined training error - $e = 0{,}03$. Table-1 represents the initial values of network parameters.

During the training, the approximation of the given function $u(t) = \cos(t) \cdot e^{-t}$ is achieved by 38 iterations, and the training error is $e = 0{,}0293$. Figures 6-8 show the graphs of changing of the parameters of WNN and the graph of the error in training process [12, 15].

Table 1. The initial parameters of WNN.

No. of neuron	*w*	*a*	*b*
1	0	10	1
2	0	10	6
3	0	10	13
4	0	10	17

Table-2 shows the figures of the parameters of WNN after the training.

Table-2. Parameters of WNN after the training.

No. of neuron	*w*	*a*	*b*
1	0	10	1
2	0	10	6
3	0	10	13
4	0	10	17

 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

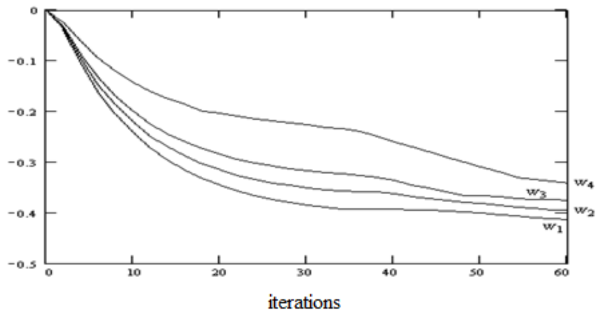


Figure-6. Values of the weighting coefficients during WNN training.

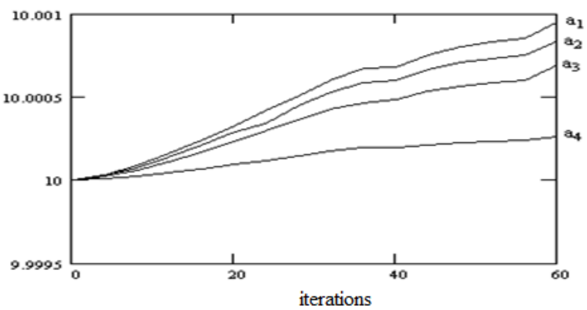


Figure-7. Values of the scale parameter during WNN training.

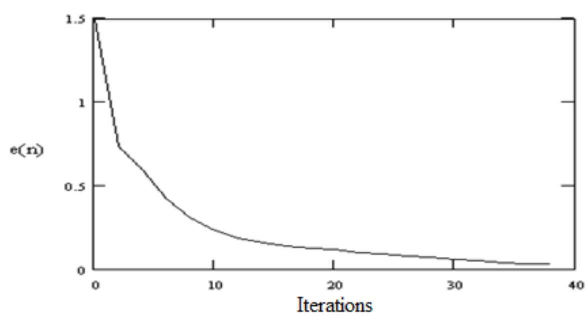


Figure-8. The graph of the error of WNN training.

3. RESULTS AND DISCUSSIONS

Results of the experiment can be stored in database [7] or knowledge base [16] and its show that using WNN for approximation of the function of one variable is more preferable rather than using traditional MNN. The advantage of WNN in the quickness of training can be explained as follows [11].

Learning algorithms for WNN and MNN are gradient methods [5]. At the same time changing of weighting coefficients both for WNN and MNN depend on the gradient $\partial E / \partial w$. Values of the energy of the error E for WNN depend on activation function, which is wavelet function. The energy of the error E for MNN is determined directly by sigmoid activation function.

In Figure-9 we can easily see that at equal change of function arguments, change in value of function is greater for wavelet function ($\Delta h_B > \Delta h_c$). The outcome of the present analysis is therefore that wavelet function has a steeper descent, compared with sigmoid function, and, at gradient method of training gradient value will be greater than gradient value $\partial E / \partial w$ for a standard multilayer network. Thereafter, WNN training process will run faster than MNN training process [6]. As is well known, wavelet function has a more flexible structure. Since the slope of a sigmoid graph is determined only by coefficient b, we can claim that wavelet function has two parameters: a, responsible for changes in the slope of the graph of wavelet function, and shift parameter b.

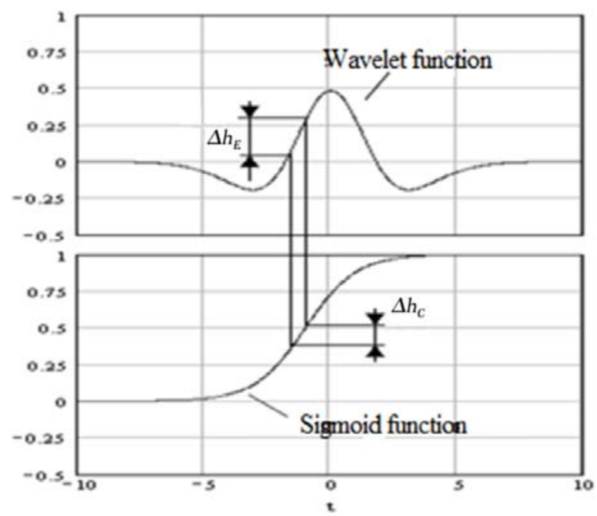


Figure-9. Activation function: wavelet function and sigmoid function (case $\Delta h_B > \Delta h_c$).

4. CONCLUSIONS AND FUTURE WORK

We have developed the model of a new type of wavelets and neural networks hybrid, intended for improvement of quality of neural networks, used in speech detection system. We have drawn a comparison between MNN and WNN in the task of approximation of functions of one variable. Advantages of WNN in contrast with multilayer network during its training, lead us to the conclusion that the quality of speech signals detection can be improved. It should be noted that further research will involve methods of using wavelet transformation for the protection and marking digital images [17].

REFERENCES

[1] Vorob'ev V. I., Gribenin V. G. 1999. Theory and practice of wavelet transform. -SPb.: Publishing in the Military University of Communication.

[2] Gapochkin A. V., Ternovoy N. V., Chervyakov N. I. 2014. The use of neural networks as classifiers in speech recognition systems. Neurocomputers: development and application. (9): 20-24.

[3] Krylov V. V., DLI M. I., Golunov R. J. 1996. Fuzzy logic and artificial neural networks. - M. Solon.

[4] Chervyakov N. I., KondrashovY.V. 2009. The use of compound logic functions to improve the speed of training the wavelet neural network. - Belgorod. pp. 113-118.

[5] Szu H.H., Telfer B.A., Kadambe S. 1992. Neural network adaptive wavelets for signal representation and classification. Optical Engineering. (9): 1907-1916.

[6] Koldaev A. I. 2012. Development of fuzzy neurocontroller uses the methods of wavelet transforms. abstract - Stavropol.

[7] Popov D. I., Popova E. D. 2009. Information technology. Database. Textbook for University students; Ministry of education and science of the Russian Federation, Federal Agency on education, Moscow state University of printing. Moscow. pp. 45-51.

[8] Popov D. I., Popova E. D. 2009. Computer systems networking and telecommunications. Lecture notes: for students of universities. Ministry of education and science of the Russian Federation, Federal Agency on education, Moscow state University of printing. Moscow. pp. 15-23.

[9] Maruti Limkar, Rama Rao and Vidya Sagvekar. 2012. Isolated Digit Recognition Using MFCC and DTW. International Journal on Advanced Electrical and Electronics Engineering, (IJAEEE), ISSN 2278-8948, 1(1).

[10] John Sahaya R. A., Thalkari S. S., Venkatesan N. 2015. Adapted DTW joint with wavelet transform for isolated digit recognition. ARPN Journal of Engineering and Applied Sciences. 10(1): 1-7.

[11] A. Jose Albin., N.M. Nandhitha, S. Emalda Roslin. 2014. ART Network Based Text Independent Speaker Recognition System for Dynamically Growing Speech Database. International Conference on Frontiers of Intell. Comput. (FICTA). Advances in Intelligent Systems and Computing. pp. 473-480.

[12] A. Jose Albin., N.M. Nandhitha, S. Emalda Roslin. 2015. Text independent human voice ranking system for audio search engines using wavelet features. ARPN Journal of Engineering and Applied Sciences. 10(4): 1650-1653.

[13] S. R. Krothapalli and S. G. Koolagudi. 2013. Emotion Recognition Using Vocal Tract Information. In: Emotion Recognition using Speech Features. Springer. pp. 67-78.

[14] Geeta Kaushik, H.P. Sinha and Lillie Dewan. 2014. Biomedical Signals Analysis by DWT Signal De-Noising with Neural Networks. Research Journal of Applied Sciences. 9: 244-256.

[15] M. Hariharan, *et al*. 2013. Objective evaluation of speech dysfluencies using wavelet packet transform with sample entropy. Digital Signal Processing: A Review Journal. 23, pp. 952-959.

[16] Ostroukh A.V., Popov D.I, Demidov D.G. 2015. Development of the rules base for an expert system choice adaptive learning strategy. ARPN Journal of Engineering and Applied Sciences. 10(10): 4430-4435.

[17] Popov D.I., Belubokova Yu.A. 2015. Bitmaps Protective Marking Model to Verify Their Integrity. International Journal of Applied Engineering Research. 10(16): 37384-37388.