Scientific Bulletin of the „Petru Maior” University of Tîrgu Mureş
Vol. 10 (XXVII) no. 2, 2013
ISSN-L 1841-9267 (Print), ISSN 2285-438X (Online), ISSN 2286-3184 (CD-ROM)

WAVELET TRANSFORM BASED NONLINEAR PREDICTION OF SIGNALS

Zoltan GERMAN-SALLO[1], Horatiu-Stefan GRIF[2]

[1,2] "Petru Maior" University of Tîrgu Mureş
Nicolae Iorga Street, no.1, Tîrgu Mureş, Romania

[1]zoltan.german-sallo@ing.upm.ro
[2]horatiu.grif@ing.upm.ro

Abstract

Estimating signals from time series is a common task in many domains of science and has been addressed for a long time by specialists. Predicting a signal from recorded time series remains however a very specific task, a great challenge. The wavelet transform provides multi-resolution analysis and allows accurate time-frequency localization of different signal properties. This paper presents a nonlinear prediction method implemented on artificial neural network based learning structure. From a discrete wavelet transform provided tree structure, specific coefficients are obtained and predicted with the already mentioned method, the reconstruction of signal is carried out using these new coefficients. The predicted signal is compared with the original one through parameters as the absolute mean error, using different analyzing functions and different learning structures. To evaluate the prediction, noise of different levels is added and the absolute mean error is recomputed and compared after every prediction.

Keywords: prediction, discrete wavelet transform, artificial neural networks, nonlinear approximation, prediction error

1. Introduction

It is possible to use various approximations, for example regression of the dependency of the predicted variable on other events that is then extrapolated to the future [1], [15]. Finding such approximation can be also difficult [2]. This approach generally means creating the model of the predicted event. For non-stationary signals we are interested in the frequencies that are dominant at any given time. We want to localize the frequencies in time which cannot be achieved by the Fourier transform. Artificial Neural Networks (ANNs) are being increasingly used in electrophysiological applications such as ECG signal detection, classification and automated diagnosis[12]. In this work the authors have investigated potential applications of artificial neural networks in electrocardiographic (ECG) signals prediction carried out by a Wavelet Packet Transform based decomposition of this signal [9].

Usually, prediction methods should have to emphasize the most important long-term characteristics of the predicted signal [7].

2. Artificial Neural Networks (ANNs)

Neural networks can be used for prediction with various levels of success. The advantage of then includes automatic learning of dependencies only from measured data without any need to add further information (such as type of dependency like with the regression). The neural network is trained from the previous recorded data with the hope that it will discover hidden relations and that it will be able to use them for predicting into future [7], [11]. In other words, neural network is not represented by an explicitly given model. It is more an input-output system that is able to learn something. The advantage of the usage of neural networks for prediction is that they are able to learn from examples only and that after their learning is finished, they are able to find non-linear dependencies, even when there is a significant noise in the training set.

3. The Discrete Wavelet Transform

The continuous wavelet transform is calculated by the convolution between the signal and analysis function, where the trigonometric analysis functions are replaced by a wavelet function [3], [4], [5]. A wavelet is a short oscillating function which contains both the analysis function and the window [6]. The frequencies are changed by contraction and dilatation of the wavelet function. The discrete wavelet transform (DWT) uses filter banks to perform the wavelet analysis. Concerning „*the multiresolution theory of orthogonal wavelets proves that any conjugate mirror filter characterizes a wavelet ψ that generates an orthonormal basis of $L^2(R)$. Moreover, a fast discrete wavelet transform is implemented by cascading these conjugate filters*" (Mallat 2001, p. 8), we consider that A filter bank consists of filters which separate a signal into frequency bands. The discrete wavelet transform decomposes the signal into wavelet coefficients from which the original signal can be reconstructed again. The DWT decomposition of the signal into different frequency bands can be obtained by successive high-pass and low-pass filtering (digital FIR structures) of the time domain followed by downsamplig to eliminate the redundancy, as shown in fig. 1. [9]. The obtained frequency bands are presented on Fig. 2.

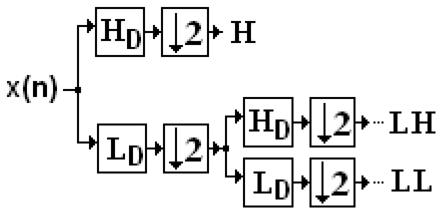

Fig. 1: Filter bank structure of DWT for a second level decomposition

In its most common form, the DWT employs a dyadic grid (integer power of two scaling in *s* and *l*) and orthonormal wavelet basis functions and exhibits zero redundancy[20].

$$\psi_{s,l}(x) = 2^{-\frac{s}{2}} \psi\left(2^{-s} x - l\right) \tag{1}$$

The variables *s* and *l* are integers that scale and dilate the mother function ψ to generate wavelets (analyzing functions). The scale index *s* indicates the wavelet's width, and the location index *l* gives its position. The mother wavelets are rescaled, or "dilated" by powers of two, and translated by integers, (dyadic decomposition structure[8], [10]).

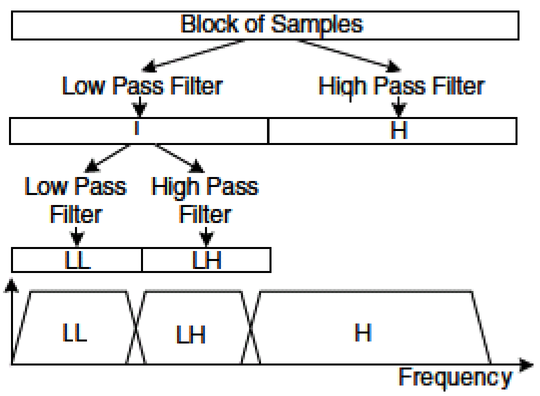

Fig. 2: Frequency bands of DWT for a second level decomposition

The decomposition procedure starts with passing the signal (sequence) through a half band digital low/pass and high-pass filters [4]. As we can see this dyadic division of the bandwidth could be a solution for subband processing techniques which allow independent processing in these frequency bands. A half band low-pass filter removes all frequencies that are above half of the highest frequency in the signal. The highest frequency component that exists in a signal is π radians, if the signal is sampled at Nyquist's rate (which is twice the maximum frequency that exists in the signal) [13]. That is, the Nyquist's rate corresponds to 2π rad/s in the discrete frequency domain [14]. After passing the signal through a half band low-pass filter, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency.

4. The proposed procedure

The main idea is to estimate a noiseless second average component from a second order DWT decomposition.

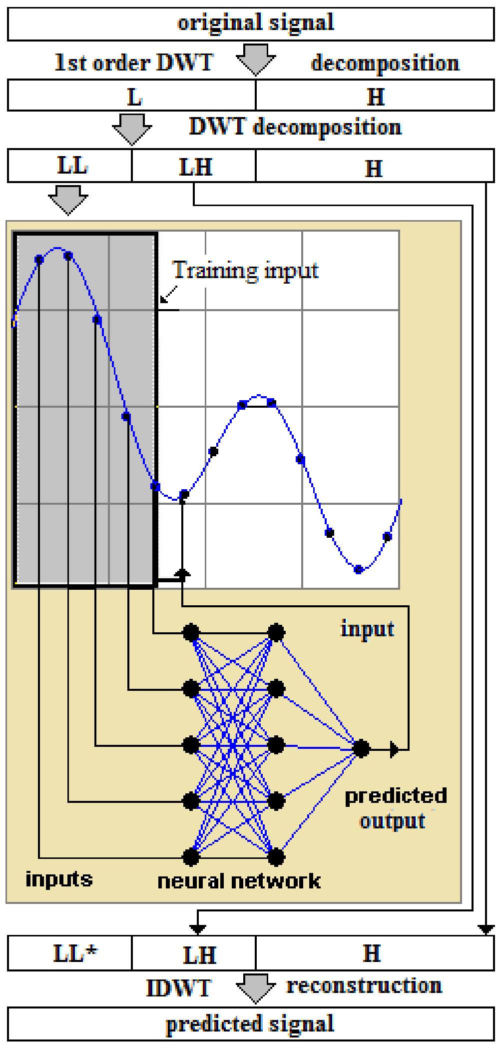

Fig. 3: The proposed procedure

The estimation task consists in the estimation of a sample from a certain number of previous samples using an artificial neural network. Before the estimation (prediction) task, the ANN is trained in off-line manner using a certain number (N_{tr}) of samples from the original signal (LL signal from fig. 3) with N samples ($N_{tr} < N$). The inputs of the ANN consist of the current sample and a fixed number of past samples. For simplification, the signal LL will be referred as x. The ANN is trained to reproduce the sample x_{k+1} when at the inputs of ANN are applied the samples $x_k, x_{k-1}, \ldots, x_{k-n+1}$, where n is the number of the inputs of the ANN and $k = n \ldots N_{tr}-1$. The samples $x_{k+1}, x_{k+2}, \ldots, x_{Ntr}$ from the original signal are used to construct the training target signal for the output of the ANN. The samples $x_1, x_2, \ldots, x_{Ntr-1}$ from the original signal are used to construct the n-dimensional training vectors for the input of the ANN. After the training of the ANN is done the prediction task follows. During the prediction task the ANN will compute the value $\tilde{x}_{k+1}$ using at its inputs the samples $x_k, x_{k-1}, \ldots, x_{k-n+1}$ where $k = n \ldots N$. The value $\tilde{x}_{k+1}$ represents the estimated (predicted) value for the sample x_{k+1}.

5. Results

For estimation, a feed forward artificial neural network with 2 hidden layers was used. The input layer has 30 inputs which represent the samples $x_k, \ldots, x_{k-29}$. The first hidden layer has 30 artificial neurons, the second hidden layer has 10 artificial neurons and the output layer has one neuron. The artificial neurons from the hidden layers have the hyperbolic tangent sigmoid transfer function type. The artificial neuron from the output layer has the linear transfer function. In fig. 4 are represented the training and the learned signals. The learned signal represents the signal - obtained at the output of the ANN – that approximates the training signal with a desired global training error.

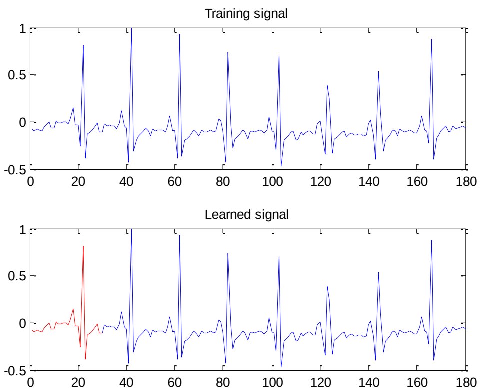

Fig. 4: The training and the learning signals

In the case of the signals from fig. 4, the ANN learned – using the Levenberg-Marquardt algorithm - the training signal in 7 epochs with a global training error equal to 10^{-8}.

In fig. 5 is depicted the training error which represent the error between the samples of training and the learned signals. The global error represents the average squared error between the ANN outputs and the samples of the training signal.

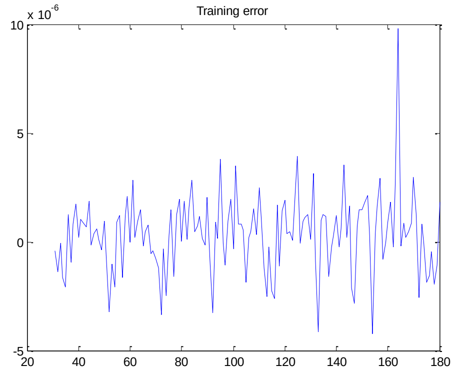

Fig. 5: Training error

In fig. 6 are presented the original signal and the predicted signal. The first part of the predicted signal is the same with the part which is learned, therefore is not changed.

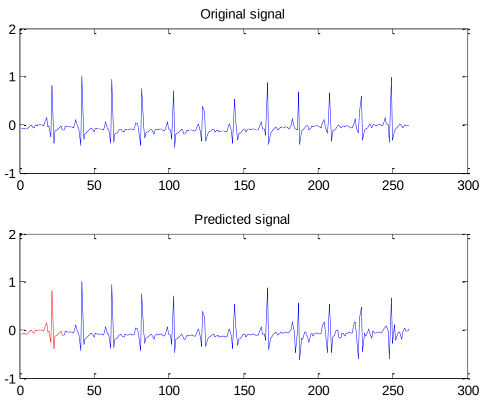

Fig. 6: The original and the predicted signals

The red part of the learned signal (fig. 4) and of the predicted signal (fig. 6) represents the first 30 samples of the original signal which were used as the initial input values for the ANN.

The different mean prediction errors obtained with different noise levels are presented on fig. 7.

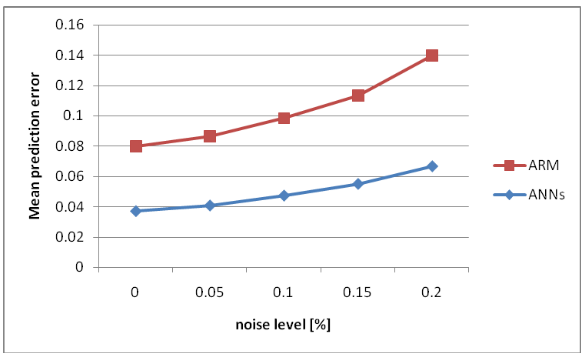

Fig. 7: The mean prediction error for different noise levels

6. Conclusion

The artificial neural network based method gave promising results, these results usually are better than the others obtained with autoregressive models, in the same perturbing conditions. It can be seen that the error obtained with autoregressive model based method increases faster than the other at higher level of noise. The average components obtained from a discrete wavelet transform are already filtered sequences, therefore an artificial neural network can learn with better results. The chosen analyzing function (wavelet) has a great influence in approximation because can represent the same sequence with a lower number of coefficients. The advantage of the usage of neural networks for prediction is that they are able to learn from examples only and that after their learning is finished, they are able to catch hidden and strongly non-linear dependencies, even when there is a significant noise in the training set. Also the training procedure strongly depends on the number of inputs and on the used learning method. As further work, a wavelet packet decomposition based prediction can be performed in order to have specific subband-components in which present interest from an analysis point of view.

References

[1] Abramovich, F., Bailey, T. C., and Sapatinas, T. (2000), Wavelet analysis and its statistical applications, *Journal of the Royal Statistical Society*, Series D , vol. 48, pp. 1–30.

[2] Barber, S. and Nason, G.P. (2004), Real nonparametric regression using complex wavelets. *Journal of the Royal Statistical Society*, Series B, vol. 66, pp. 927–939.

[3] Barber, S., Nason, G. P., and Silverman, B.W. (2002), Posterior probability intervals for wavelet thresholding, *Journal of the Royal Statisitcal Society*, Series B, vol. 64, pp. 189–205.

[4] Burrus, C.S., Gopinath, R.A. and Guo, H. (1998), *Introduction to Wavelets and Wavelet Transforms*, Prentice Hall.

[5] Keinert, F. (2004), *Wavelets and Multiwavelets*, Chapman and Hall / CRC Press, Boca Raton, Fl.

[6] Lamár, K. and Morva, G. (2013), Hardware and Software Functions of Standalone Field Data Acquisition Devices for the Low Voltage Power Distribution Grid, *Carpathian Journal of Electronic and Computer Engineering* , vol. 6/1, pp. 22-27.

[7] Lotri, U., (2004), Wavelet based denoising integrated into multilayered perceptron, *Neurocomputing*, vol. 62, pp. 179–196.

[8] Weeks, M. (2007), *Digital Signal Processing using Matlab and Wavelets*, Infinite Science Press LLC.

[9] Mallat, S. (2001), *A Wavelet Tour of Signal Processing*, Academic Press.

[10] Manly, B.F.G. (2005), *Multivariate Statistical Methods: A Primer*, Chapman & Hall/CRC, Boca Raton.

[11] Demuth, H. and Beale, M. (2002), Neural Network Toolbox User's Guide, Math Works Inc.

[12] Percival, D.B. and Walden, A.T. (2000), *Wavelet methods for time series analysis*, Cambridge University Press, Cambridge.

[13] Renaud, O., Starck, J.-L., and Murtagh, F. (2003), Prediction based on a multiscale decomposition, *International Journal of Wavelets, Multiresolution and Information Processing*, 1(2), pp. 217–232.

[14] Soltani, S., Boichu, D., Simard, P., and Canu, S. (2000), The long-term memory prediction by multiscale decomposition, *Signal Processing*, vol. 80, pp. 2195–2205.

[15] Van der Ville, D., Blu, T., and Unser, M. (2006), Surfing the brain — an overview of wavelet-based techniques for MRI data analysis, *IEEE Transactions on Signal Processing*, vol. 54, pp. 1256–1267.