Available online at www.elixirpublishers.com (Elixir International Journal)

Electrical Engineering

Elixir Elec. Engg. 73 (2014) 26225-26229

ISSN: 2229-712X

Compare Wavelet statistical and power features in classification and separation of arousal signals from scalp

Naser Ziaei[1,*], Pezhman Aghaei[1] and Ali Rafiee[2]

[1]Department of Electrical and Computer Engineering, Islamic Azad University, Semirom Branch, Esfahan, Iran.

[2]Department of Electrical and Computer Engineering, Islamic Azad University, Kazerun Branch, Fars, Iran.

ARTICLE INFO

Article history:

Received: 25 June 2014;

Received in revised form:

25 July 2014;

Accepted: 6 August 2014;

Keywords

Classification, Wavelet, EEG,
DWT, Evoked potentials.

ABSTRACT

Aim of this article, classification statistical and power characteristics of the command brain motor with different structures of powerful perceptron neural network Combined with Levenberg Marquardt algorithm and evaluating the best structure to separate these signals. Detection of motor steering signals in the brain is an important classification issue. The discrete wavelet transform to extract features and investigate of the scale-frequency electroencephalogram signals are used. Results show perceptron network with two hidden layers and twelve neurons with linear output transfer function at best 92% and then the multilayer perceptron with one hidden layer and transfer function tangent sigmoid 86% have ability to separate.

Introduction

The human brain is a complex system and Shows spatial and temporal activity. Among a variety of non-invasive techniques to evaluate human brain activity, electroencephalogram systems provides a direct method of the external activity of the brain in a millisecond time resolution. Important factor in the accuracy and efficiency in classifiers, reducing the dimensions of feature space. Reduce the computational cost and classification accuracy, reducing the feature space are two major reasons [1]. DWT wavelet transfer method a single framework for the different techniques are used for various applications, it offers [9-6]. Because the wavelet transform method is suitable for the study of motion signals and this is an important advantage for studying the spectrum, of this method used for unstable events that may occur in the brain during imagined voluntary movement of the arm or leg [2]. via wavelet analysis recorded electroencephalogram, unstable characteristics are properly recorded and are focused on both time and frequency. In this article, we try to extract time - scale features of voluntary commands imagery of hands used for the best structure of Multilayer Perceptron network to do this. Multilayer Perceptron networks have the ability to learn and generalize, Require less training, quick and easy implementation we used for our classification and Levenberg–Marquardt algorithm learning classical application are used for training this specific networks. We want to detect the Artificial Neural Network which has the best resolution for the separation signal and Simultaneously, they also have separate left and right hand motor imagery. Here are, the brain signals to the separation of the two modes of related mental imagery of right hand and left hand use. Type of mental activities considered in this paper are in terms of functional neural prosthesis or the human-machine interface. Also we use a threshold also Inactivity mode features get added, That in machine computer interfaces applications it's utmost importance [12-2].

The Data Records

Data from a healthy subject (female, 25 years old) is collected in one feedback session. The purpose of this experiment is to control the screen cursor with hand movements left or right with subjective imagine. Orders left and right movement is random [12]. Test consists of seven stages, each consisting of 40 tested. All the steps are performed several times a day with a stop between each step . Each of the 280 test takes 9 seconds. the first 2 seconds are Generally and 2 seconds duration at the beginning of a sound stimulus is heard. Trigger channel moves from the bottom up and the sign (+) is shown for 1 second, Then in 3 seconds an arrow cursor (left or right) is shown. feedback based on a AAR channel parameter # 1 (C3) and # 5 (C4) are. Data recorded by an amplifier G.tec and Ag / Agcl electrodes done. Three channels of bipolar EEG (Former(+) or posterior (-)) on channels C3, Cz, C4 are measured. EEG sampling rate of 128 Hz and filtered between 0.5 and 30 Hz [12]. The format of the data. Mat have been saved. X_train variables including three EEG channels and 280 test is 9 seconds. The indicator is evident from the third to the ninth seconds. At the same time, feedback from the screen is shown in to the person tested [11].

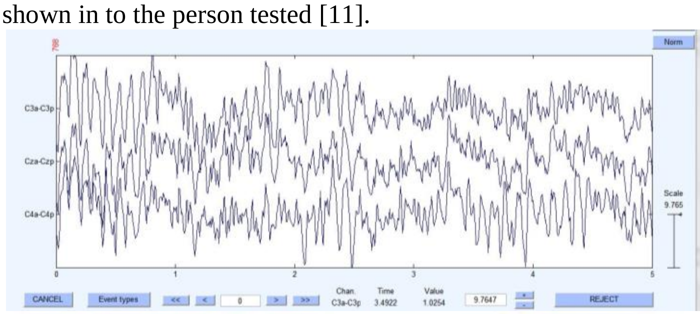


Figure 1: electrodes View the raw signals from the electrodes taken by EEGlab under MATLAB software, scale 9.765.

Since the data selected in the range between 0.5 and 30Hz is filtered in Programing don't need to filter the noise power (50 Hz) as well as an eye artifact removal filter because According to sampling these data team there's no effect of artifacts . Also, if there are visual artifacts in the area of the brain it have less motion in sensory motor [5]. The low number of electrodes used in this study is one of the other innovations this study.

Wavelet Features

There The problem Fixed resolution in Fourier time-short is root in Heisenberg's uncertainty principle. According to this

Tele:
E-mail addresses: Naserziaei@gmail.com

principle can't be described exact as time-frequency of a signal. And we can't know exactly what the signal frequency components when are there, It can be found only in what intervals, what the frequency band is[1]. Principles of discrete wavelet transform returns to sub-band coding method called. In the discrete case, filters with different cutoff frequencies for signal analysis in different scales were use. with High-pass and low-pass filters signals pass, different frequencies are analyzed. In discrete mode, signals resolution is controlled by the sample scale and scale variant from high or low sample changes. Typically, the process of changing the sample rate dyadic on a network with s0 = 2 and τ0 = 1 is done. The corresponding time scale and shift are respectively s = 2j and τ = k2j. The wavelet method is usually calculated on a scales in only 2 from (ie s2 ∈j and j ∈Z) The resulting wavelet transform and the wavelet transform is defined Dyadic. So Dyadic wavelet function f (x) in equation (1) would be [7]:

$$W_{\psi} f(j,k) = 2^{\frac{-j}{2}} \int_{-\infty}^{+\infty} f(x)\psi(2^{-j}x-k)dx \quad (1)$$
$$j = 0,1,2,\ldots$$
$$k = 0,\pm 1,\pm 2,\ldots$$

Dyadic continuous wavelet coefficients obtained both can be used as features [14]. Wavelet technique used to EEG signals, is detected related to the characteristics of the transient nature of the signal that Fourier don't shows [5]. All wavelet should be based on a low-pass filter g standard squaring circuit is doing, to be determined.

$$G(z)G(z^{-1}) + G(-z)G(-z^{-1}) = 1 \quad (2)$$

G (z) represents the transfer of the g of the z filter. It can be used as a complementary high-pass filter can be defined as:

$$H(z) = zG(-z^{-1}) \quad (3)$$

decomposing Signal into different frequency high-pass and low-pass filters of the time domain signal is obtained. The first filter, h [.] a single mother wavelet is the nature of the high-pass filter and the second g [.] is a filter with mirror image which have the nature of the low pass. The first outputs of high-pass and low-pass filters respectively provides D1 detailed and A1 approximation. First approximation A1 further decomposed and this process continues [6-16].

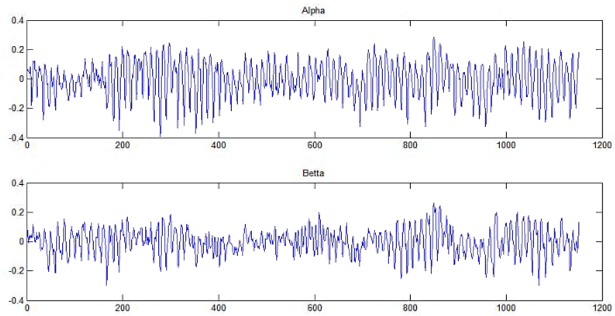
Feature Extruction

Select an appropriate wavelength and the number of levels decomposition of the signals that they use of DWT is very important. The number of levels of decomposition based on frequency elements are selected. Levels are chosen such that those parts of the signal are well correlated with the frequency remain of the wave signal classification coefficients [6]. In this paper, due to the frequency components EEG signals have not high 30 Hz, the number of decomposition levels is 4. Thus, EEG signals to detail D1 to D2 and a final approximation A4, have been analyzed. Soft wave of property daubechies (db2) it makes perfect waves for detecting changes in EEG signals. Thus the wave coefficients are calculated using db2 in this article. Frequency ranges, beta, alpha, delta and theta make brain band Beta and alpha bands in to the notion of involuntary body movements contribute and in imagined hand movement will have most motivation [5]. After considering these tips for reducing dual-band characteristics of alpha and beta will use statistical and power parameters our view derive these two frequency bands obtained by the discrete wavelet decomposition. Statistical characteristics are used to represent the time-signal frequency EEG distribution :

- Mean for two-band wavelet coefficients alpha and beta
- Standard deviation values in the two bands, beta and alpha
- Combine these features together

First the characteristics of a frequency distribution of the signal and the second, the change in the frequency distribution the third and fourth features the band shows the high and low frequency bands. The feature vector obtained for different frequency bands, A4 and D2-D4 can be used in classifying EEG signals. Now we have a matrix that is the X-Train. That to the left The standard deviation of alpha band electrodes C3, the mean alpha band electrodes C3, SD beta band electrodes C3, the average beta band electrodes C3, (Shown in Table 1), and likewise for the C4 electrode to the eighth column, and then the features of high and low throughput frequency bands for C3 and C4 to the eighteenth column. In Figure 1 of the two beta and alpha frequency bands after separation wavelet schema is given. As can be seen in the beta frequency band (30-13 Hz) than the alpha band frequency (13-8 Hz) is clearly evident. In Table 1, the features extracted from recorded in the right hand imagery is shown. Discrete wavelet transform and artificial neural network discriminant using MATLAB version 8.0.0.783_64bit took and the original waveform with the MATLAB software package BIOsig and EEGlab are taken.

Figure 2: Wavelet both alpha and beta frequency bands and ranges are shown with a symbol rate of 1152.



Classifier

Place Multilayer Perceptron MLP network, which has the ability to learn and generalize, requires less training, quick and easy implementation We used for our classification the Powerful learning programs and specific network Levenberg–Marquardt algorithm is used to train. To construct a neural network with more efficiency, Input feature vector is normalized. Thus, the data in the range [0,1] are reduced. Since the number of class is 2 numbers the output of our neural network output is 2. To learn the of weights a multi-layer BP method used. This method uses the gradient descent tries to minimize the objective function is the squared error between the network outputs [4]. The error is defined as follows:

$$E\left(\vec{W}\right) \equiv \frac{1}{2}\sum_{d\in D}\sum_{k\in outputs}\left(t_{kd}-o_{kd}\right)^2 \quad (6)$$

Levenberg Marquardt Optimizer

Propagation method, by the diffusion output error layer to the adjusting the weights on each rear side layer, the network weights are trained. Marquardt method, this method has a very low convergence rate, this compensates for the defect. The LMA interpolates between the Gauss–Newton algorithm (GNA) and the method of gradient descent. The LMA is more robust than the GNA, which means that in many cases it finds a solution even if it starts very far off the final minimum.

Table 1: Standard deviation of alpha band electrodes C3, the mean alpha band electrodes C3, SD beta band electrodes C3, the average beta band electrodes C3. (Typically the first 12 numbers of the 280 training data)

SD of Alpha band	Mean of Alpha	S D of Beta band	Mean of Beta band
0.0361138379614577	0.000108269966116095	0.0391532284416653	-2.316818393e-06
0.0383765667527218	-1.496921138e-05	0.0538205131741770	-2.722492425e-05
0.0386091785182868	3.39707654720731e-05	0.0438326075342212	-4.295146019e-06
0.0317475815082760	-7.674452918e-05	0.0480219997978874	-0.00014219605
0.0398016696771668	-2.749689140e-05	0.0531578082667693	6.75654901570762e-05
0.0476989241182419	-5.78357861e-05	0.0645483825058020	-9.253881207e-05
0.0366740631599504	3.48159658623174e-05	0.0419910427828653	-1.399277962e-05
0.0711928131888616	1.19458307620687e-05	0.0801417857137126	3.94012325410009e-05
0.0410763109513764	2.69005311664288e-05	0.0595416755711951	2.91975946891861e-05
0.0447050235507325	1.83535366777525e-05	0.0515914925065320	1.07037577613173e-05
0.0519397174772735	-8.702817367e-06	0.0630641784444973	2.06812790357358e-05
0.0445045672145851	-3.442263068e-06	0.0639015333146983	0.000115918682506822

Table 2: Performance of Separation MLP networks with different structures with statistical features wavelet.

two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 50% of data for training and testing the network			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	0.999	59.25%	$2.44e^{-05}$
8 neuron	0.99	56.4%	$4.4e^{-06}$
10 neuron	1	49.5%	$2.81e^{-08}$
12 neuron	0.98	43.2%	0.00714
14 neuron		43.4%	$2e^{-20}$
values obtained from the various iterations of the two-layer network and an output layer with two neurons and Linear transfer function and 50% of data for training and testing the network			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	1	22.3%	$6.17e^{-07}$
8 neuron	1	41.6%	$6.82e^{-17}$
10 neuron	0.97	36.5%	$5.93e^{-06}$
12 neuron	0.98	34.5%	$4.66e^{-05}$
14 neuron	0.99	39.9%	0.00079
values obtained from the various iterations of the two-layer network and an output layer with two neurons and Linear transfer function and 70% of data for training and 30% testing			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	0.972	57%	0.0143
8 neuron	0.999	50.2%	$4.15e^{-05}$
10 neuron	0.97	51.8%	0.0143
12 neuron		46.6%	$3.68e^{-06}$
14 neuron		41.6%	$2.49e^{-10}$

Table 3: two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 60% of data for training and 25% of data for testing validation check function in the network.

values obtained from the various iterations of the two-layer network and an output layer with one neurons and Linear transfer function and 60% of data for training and 25% testing and 15% validation			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	74.7%	69%	0.0343
8 neuron	80%	66%	0.0243
10 neuron	86.7%	55.4%	0.0099
12 neuron	84.6%	75.3%	0.0101
14 neuron	94.7%	62.4%	0.00134

Table 4: Performance of MLP networks separation with different wavelet features structures with validate function in the second and third sections

Parameters / network	train regression	Test regression	MSE network
the two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 60% of data for t raining and 40% of data for testing			
6 neuron	0.993	34.3%	0.00335
8 neuron	0.999	45.6%	$1.09e^{-05}$
10 neuron	1	43%	$1.85e^{-21}$
12 neuron	1	41.25%	$1.87e^{-19}$
14 neuron	1	42.6%	$1.65e^{-20}$
values obtained from the various iterations of the two-layer network and an output layer with two neurons and Linear transfer function and 60% of data for training and 25% testing and 15% validation in GUI			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	1	81%	$1.06e^{-08}$
8 neuron	0.957	65%	$1.34e^{-06}$
10 neuron	0.964	72.8%	$4.09e^{-05}$
12 neuron	0.978	61%	$4.97e^{-06}$
14 neuron	0.997	86%	$1.01e^{-08}$
values obtained from the various iterations of the two-layer network and an output layer with two neurons and Linear transfer function and 60% of data for training and 25% testing and 15% validation			
Parameters / network	**train regression**	**Test regression**	**MSE network**
6 neuron	0.782	56%	0.0511
8 neuron	0.811	73.2%	0.0308
10 neuron	0.832	67.2%	0.0159
12 neuron	0.825	76.4%	0.0179
14 neuron	0.893	76%	0.00622

For well-behaved functions and reasonable starting parameters, the LMA tends to be a bit slower than the GNA. LMA can also be viewed as Gauss–Newton using a trust region approach[15]. In this section, the square-error method is used to minimize the error. Whatever squared error are close to zero the difference between network output and actual output levels will be decreased[3]. During network training, the input data is repeatedly presented to the network. Therefore, when using the neural network must be assigned to the extent of about two sets of training data and test data to make decisions [4]. In this algorithm, the number of iterations to converge is less than other methods and LMBP algorithm is Fastest neural network learning algorithm in the average effective number of parameters. This method is the fastest method implemented in MATLAB and for a medium-sized networks with a few hundred variables is very high efficiency [7].

Simulation Results

In this section the statistical features mean and standard deviation of the data's using wavelet analysis of experimental data for alpha and beta bands were extracted, this data is then placed in a matrix and encoded the target matrix as data target of with feature matrix for Two mental moving right and left hands classification will giving to the neural network. Simulations in several stages Simulations in several stages for obtain correct values higher with varied number of hidden layer neurons and also with different amounts of data for testing, training and validation for network used. To evaluate and compare the simulation results are given in Tables 2 to 4.

Table 2, values obtained from the various iterations of the two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 50% of data for training and testing the network, best resolution with six input neurons is 59% and observed that the increasing the number of neurons regulation will be reduced. Values obtained from the various iterations of the two-layer network with two neurons output layer and a linear transfer function and using 50% of data for training and testing the network, can be seen that the resolution of output the linear function is less than the tangent sigmoid function transfer. values obtained from the various iterations of the two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 70% of data for training and 30% for testing the network, in this structure was also observed that increasing the number of neurons resolution reduced.

Table 3, values obtained from the various iterations of the two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 60% of data for training and 40% of data for testing the network. values obtained from the various iterations of the two-layer network and an output layer with two neurons and Tangent sigmoid transfer function and 60% of data for training and 25% of data for testing and 15% of data for validation the network and table 4, power of the MLP network classification with two layers and one neuron in output layer and linear transfer function and 60% of data for training and 25% of data for testing and 15% of data for validation with wavelet characteristics with validation function are showing. To evaluate and compare the simulation results are given in Tables 2 to 4:

Results

The as though a multi-layer BP network with suitable number of neurons can be implement any function, but it can't always find the correct weight to optimal solution [4]. One solution is that we created the network again and learning it several times to ensure that we have find the best answer, Similar to what was done in this study. also the LM number of iterations to converge is less than other methods And has high convergence rate. Two-layer neural network with an tangent sigmoid output layer than the network with a linear output layer have less percentage error and higher regression for Data testing and training and better classification has done, the number of hidden layer neurons, which are used Whatever increase in the number neurons more than 8 number causes a lower regression and speed of learn network will decrease due to increasing neurons in the discussing brain computer interface is not good. Therefore, the involvement of more than 14 neurons for the separation characteristics of the two electrodes was avoided up the in addition to review the network with a higher resolution for this purpose identify faster network with less neurons to do this. In table 4 Simulations were performed in the GUI ambience and of a neuron with a linear transfer function for the output of the Network is used, Validation data Network has a significant impact on accuracy, noted that MSE and regression for classification test data more than 68% was not obtained. Highest regression related to two layers network with 6 neurons in the first layer and 2 neurons in the output tangent sigmoid transfer function, and 60% for training data, 25% for testing and 15% for the validation data was obtained. The three-layer network with first hidden layer 8 tangent sigmoid neurons, second hidden layer 2 neuron tangent sigmoid and linear transfer function in the output 2 neuron makes classification up to 92% . Although At the entrance of the network, increased most neurons up to 16 numbers can slightly increased efficiency but the learning network time goes up, this deceleration in human prosthetic interface applications issue is not desirable. Therefore, involved more number of neurons in the application of neural networks for classification of this data wouldn't be necessary, because a total of 12 neurons in the network have reached the desired response and 82% regression to test data was obtained. It should be considered that the features used for comparison discussion and classification are mean and standard deviation of alpha and beta bands and the use of alpha-and beta-band power features in the results can sometimes be more efficient. given that LM is a classic network optimization and determine the coefficients and the combination of method of BP, AL, GD has arisen. Specific advantages of the method in reference [2] reasons has been use of this method. if the input data and parameters used in the process are very much it is suggested multi-layer perceptron networks with smart optimization techniques used in the classifier and compared.

References

[1] A. Bnakar, wavelet neural network and its application in fuzzy systems - nervous, Reza Bayat edition, numbered edition 149, Scientific Publishing Center of Tarbiat Modarres University,ISBN:978-600-5394-63-4, 2013.

[2] Naser Ziaei, Ali Rafiee, M. Maesoumi, diagnosis and control of motor cortical processing of EEG signals in order to use the wheelchair using neural networks, Master's thesis, Islamic Azad University, Kazerun-branche, Department of Electrical, Kazerun, Iran, 138 pages, 2013.

[3] Martin T. Hagan, Howard B. Demuth and Mark H. Beale, Neural Network Design, translated by Seyed Mustafa Kia, Second Edition, ISBN 9786005237559, place of publication kian rayaneh sabz, Tehran, 2011.

[4] Mohammad Bagher Menhaj, Fundamentals of Neural Networks, Volume I, Eighth Edition, ISBN 9789644630873, published at Amirkabir University of Technology, Tehran, 2012.

[5] EEG SIGNAL PROCESSING Saeid Sanei and J.A. Chambers . Centre of Digital Signal Processing .Cardiff University, UK John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.

[6] Wavelet Toolbox User's Guide R2011b Michel Misiti, Yves Misiti, Georges Oppenheim, Jean-Michel Poggi info@mathworks.com

[7] Digital Signal Processing Using MATLABr and Wavelets Michael Weeks Georgia State University . Infinity Science Press LLC Hingham, Massachusetts. ISBN: 0-9778582-0-0

[8] *Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery*. *Journal of clinical neurophysiology*, Vol. 16, No. 4. (July 1999), pp.IEEE. 373-382.

[9] *Automatic Sleep Stage Classification Based on EEG Signals by Using Neural Networks and Wavelet Packet Coefficients* .Farideh Ebrahimi, Mohammad Mikaeili, Edson Estrada, Homer Nazeran, Senior Member, Vancouver, British Columbia, Canada, August 20-24, 2008 IEEE

[10] *EEG Eye Blink Classification Using Neural Network*. Brijil Chambayil, Rajesh Singla, R. Jha .Proceedings of the World Congress on Engineering 2010 Vol I WCE 2010, June 30 - July 2, 2010, London, U.K.

[11] University of Technology *Graz*. (Gert Pfurtscheller) http://www.bbci.de/competition/ii/

[12] *Implementation of Epileptic EEG using Recurrent Neural Network* M. Gayatri, Arun Kumar, Manish Janghu, Mandeep Kaur. and Dr. T.V. Prasad. IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.3,IEEE March 2010.

[13] Subasi, A. *Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients.* Expert Systems with Applications,IEEE 28, 701–711,2005.

[14] G. G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Lugger, "*Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters*," IEEE.

[15] http://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt_algorithm

[16] Combined neural network model employing wavelet coefficients for EEG signals classification. Elif Derya Ubeyli. Digital Signal Processing 19 (2009) 297–308 2009 Elsevier.

[17] *ANN-based classification of EEG signals using the average power based on rectangle approximation window*. Mahmut HEKIM PRZEGLĄD ELEKTROTECHNICZNY (Electrical Review), ISSN 0033-2097, R. 88 NR 8/IEEE 2012.