

Neuro-Fuzzy Prediction for Brain-Computer Interface Applications

Wei-Yen Hsu

*Graduate Institute of Biomedical Informatics, Taipei Medical University
Taiwan*

1. Introduction

The brain-computer interface (BCI) work is to provide humans an alternative channel that allows direct transmission of messages from the brain by analyzing the brain's mental activities [1-7]. The brain activity is recorded by means of multi-electrode electroencephalographic (EEG) signals that are either invasive or noninvasive. Noninvasive recording is convenient and popular in BCI applications so it is commonly used. According to the definition suggested at the first international meeting for BCI technology, the term BCI is reserved for a system that must not depend on the brain's normal output pathways of peripheral nerves and muscles [2]. It has become popular for BCI systems on motor imagery (MI) EEG signals in the last decade [8]. It reveals that there are special characteristics of event-related desynchronization (ERD) and synchronization (ERS) in mu and beta rhythms over the sensorimotor cortex during MI tasks by discriminating EEG signals between left and right MIs [9, 10]. ERD/ERS is the task-related or event-related change in the amplitude of the oscillatory behavior of specific cortical areas within various frequency bands. An amplitude (or power) increase is defined as event-related synchronization while an amplitude (or power) decrease is defined as event-related desynchronization. As other event-related potentials, ERD/ERS patterns are associated with sensory processing and motor behavior [2]. The principal objective of this study is to propose a BCI system, which combines neuro-fuzzy prediction and multiresolution fractal feature vectors (MFFVs) with support vector machine, for MI classification.

A model is used for time series prediction to forecast future events based on known past events [11]. A variety of methods have been presented in time series prediction, such as linear regression, Kalman filtering [12], neural network (NN) [13], and fuzzy inference system (FIS) [14]. Linear regression is simple and common, but it has less adaptation. Kalman filtering is an adaptive method, but intrinsically linear. The NN can approximate any nonlinear functions, but it demands a great deal of training data and is hard to interpret. On contrary, FIS has good capability of interpretation, but its adaptability is relative low. FISs are fuzzy predictions that can learn fuzzy "if-then" rules to predict data. They are readable, extensible, and universally approximate [14]. Adaptive neuro-fuzzy inference system (ANFIS) [15] integrates the advantage of both NN and fuzzy system. That is, ANFIS not only has good learning capability, but can be also interpreted easily. In addition, the training of ANFIS is fast and it can usually converge only depending on a small data set.

These good properties are suitable for the prediction of non-stationary EEG signals. Therefore, ANFIS is used for time-series prediction in this study.

An effective feature extraction method can enhance the classification accuracy. An important component for most BCIs is to extract significant features from the event-related area during different MI tasks. A great deal of feature extraction methods has been proposed. Among them, the band power and AAR parameters are commonly used [16–19]. Feature extraction based on band power is usually obtained by computing the powers at the alpha and beta bands. The features are then extracted from band powers by calculating their logarithm values [16] or averaging over them [17]. AAR parameters are another popular feature in mental tasks [18, 19]. The all-pole AAR model lends itself well to modeling EEG signals as filtered white noise with certain preferred energy bands. The EEG time series is fitted with an AAR model.

Furthermore, fractal geometry [20] provides a proper mathematical model to describe complex and irregular shapes that exist in nature. Fractal dimension is a statistical quantity that effectively extracts fractal features. In the last decade, feature extraction characterized by fractal dimension has been widely applied in various kinds of biomedical image and signal analyses, such as texture extraction [21], seizure onset detection in epilepsy [22], routine detection of dementia [23], and EEG analyses of sleeping newborns [24]. In this study, discrete wavelet transform (DWT) together with modified fractal dimension is utilized for feature extraction. That is, MFFVs are extracted from wavelet data by modified fractal dimension. MFFVs contain not only multiple scale attributes, but important fractal information.

The support vector machine (SVM) [25] recognizing the patterns into two categories from a set of data is usually used for the analyses of classification and regression. For example, the SVM is used to classify attention deficit hyperactivity disorder (ADHD) and bipolar mood disorder (BMD) patients by proposing an adaptive mutation to improve performance [26]. The SVM is used for seizure detection in an animal model of chronic epilepsy [27]. Since it can balance accuracy and generalization simultaneously [25], it is used for classification in this study.

To evaluate the performance, several popular methods, including AAR-parameter approach and AAR time-series prediction, are implemented for comparison. This chapter is organized as follows: Section 2 presents the materials and methods. Section 3 describes experimental results. The discussion and conclusion are given in Sections 4 and 5, respectively.

2. Problem formulation

An analysis system is proposed for MI EEG classification, as illustrated in Fig. 1. The procedure is performed in several steps, including data configuration, neuron-fuzzy prediction, feature extraction, and classification. Raw EEG data are first filtered to the frequency range containing mu and beta rhythm components in data configuration. ANFIS time-series predictions are trained by the training data at offline. Information from ANFIS time-series predictions is directly applied to predict the test data. Modified fractal dimension combined with DWT is utilized for feature extraction. The extracted fractal features are used to train the parameters of SVM classifier at offline. Finally, the SVM together with trained parameters is utilized to discriminate the features.

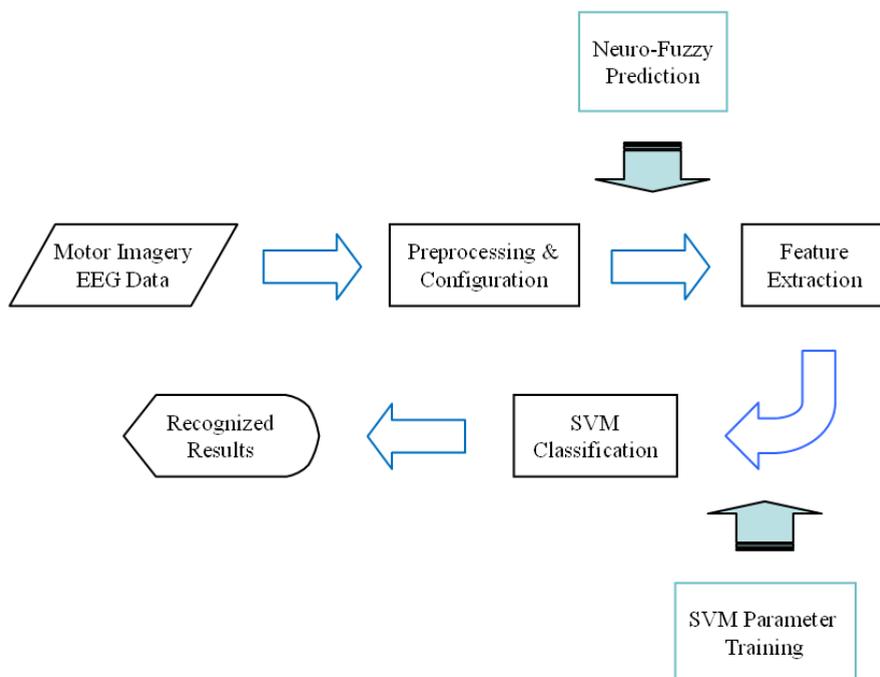


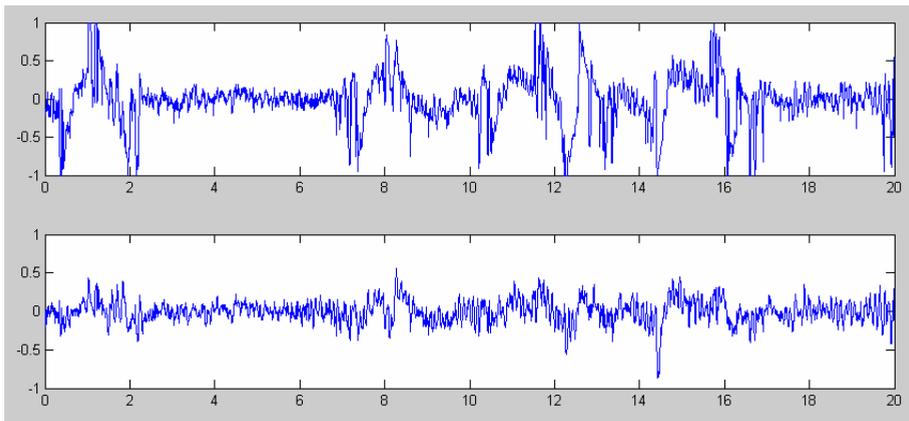
Fig. 1. Flowchart of proposed system.

3. Experimentation

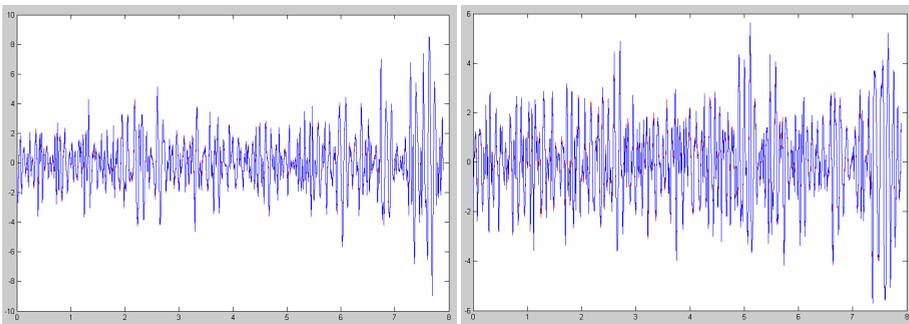
The EEG data was recorded by the Graz BCI group [19, 28–32]. Two data sets are used to evaluate the performance of all methods in the experiments. The first data sets were recorded from three subjects during a feedback experimental recording procedure. The task was to control a bar by means of imagery left or right hand movements [19, 28, 30, 31]. The order of left and right cues was random. The data was recorded on three subjects – the first subject S1 performs 280 trials, while the last two subjects, S2 and S3, hold 320 trials. The length of each trial was within 8–9s. The first 2s was quiet, an acoustic stimulus indicates the beginning of a trial at $t = 2s$, and a fixation cross + was displayed for 1s. Then at $t = 3s$, an arrow (left or right) was displayed as a cue (the data recorded between 3 and 8s are considered as event related). At the same time, each subject was asked to move a bar by imagining the left or right hand movements according to the direction of the cue. The recordings were made using a g.tec amplifier and Ag/AgCl electrodes. All signals were sampled at 128 Hz and filtered between 0.5 and 30 Hz. An example of a trial for C3 and C4 channels is given in Fig. 2(a).

The second data sets were recorded from three subjects by using a 64-channel Neuroscan EEG amplifier [29, 32]. The left and right mastoids served as a reference and ground, respectively. The EEG data was sampled at 250 Hz and filtered between 1 and 50 Hz. The subjects were asked to perform imagery movements prompted by a visual cue. Each trial started with an empty black screen; at $t = 2s$ a short beep tone was presented and a cross '+'

appeared on the screen to notify the subjects. Then at $t = 3\text{s}$ an arrow lasting for 1.25s pointed to either the left or right direction. Each direction indicates the subjects to imagine either a left or right hand movement. The imagery movements were performed until the cross disappeared at $t = 7\text{s}$. No feedback was performed in the experiments. The data set recorded from subject S4 was 180 trials, while the data sets for subjects S5 and S6 were 120 trials. For each subject, the first half of the trials were used as training data and the later half of the trials were used as test data in this study.



(a) An example of a trial



(b) Actual and predicted signals (C3) (c) Actual and predicted signals (C4)
(Actual filtered signals: Red; Predicted signals: Blue)

Fig. 2. Intermediate results.

4. Methodologies

4.1 Data configuration

The mu and beta rhythms of the EEG are those components with frequencies distributed between 8-30 Hz and located over the sensorimotor cortex. In addition, using a wider frequency range from the acquired EEG signals can generally achieve higher classification accuracy in comparison with a narrower one [33]. A wide frequency range containing all mu and beta rhythm components is adopted to include all the important signal spectra for MI

classification. In this study, the raw EEG data are filtered to the frequency range between 8 and 30 Hz with a Butterworth band-pass filter.

To make a prediction at sample t , the measured signals extracted from the recorded EEG time-series data are used from samples $t-Ld$ to $t-d$. The parameters L and d are the embedding dimension and time delay, respectively. Each training input data for ANFIS prediction consist of respective measured signals of length L on both the $C3$ and $C4$ channels, which are important for BCI works because they are located in the sensorimotor cortex [34]. The training input data are represented as follows:

$$\left[(C3_{t-Ld}, \dots, C3_{t-d}, C4_{t-Ld}, \dots, C4_{t-d})^t \mid (C3_t, C4_t)^t \right] \quad (1)$$

There are event related data of approximately 5s length in each trial. All parameter selection is performed from the training data. All training data are used to train the parameters of prediction models, which will be further used for feature extraction. The test data are finally tested to evaluate the performance of the system by using the trained parameters.

4.2 Neuro-fuzzy prediction

Time series prediction is the use of a model to forecast future events based on known past events. Although all kinds of methods in time series prediction have been presented, ANFIS time-series prediction is slightly modified and adopted in this study since it integrates the advantages of NN and fuzzy system.

The ANFIS network architecture applied for the time-series prediction of EEG data is introduced. A detailed description of ANFIS can be found in [15]. ANFIS enhances fuzzy parameter tuning with self-learning capability for achieving optimal prediction objectives. An ANFIS network is a multilayer feed-forward network where each node performs a particular node function on incoming signals. It is characterized with a set of parameters pertaining to that node. To reflect different adaptive capabilities, both square and circle node symbols are used. A square node (adaptive node) has parameters needed to trained, while a circle node (fixed node) has none. The parameters of the ANFIS network consist of the union of the parameter sets associated to each adaptive node. To achieve a desired input-output mapping, these parameters are updated according to given training data and a recursive least square (RLS) estimate.

In this study, the ANFIS network applied for time-series prediction contains L inputs and one output. There are 2^L fuzzy if-then rules of Takagi and Sugeno's type [35] in the representation of rule base. The output is a current sample, and the inputs are the past L samples in the time delay t . The output of the i th node in the l th layer is denoted by O_i^l . The node function for each layer is then described as follows.

Layer 1: Each node in this layer is a square node, where the degree of membership functions of input data is calculated. The output of each node in this layer is represented as

$$O_i^1 = \mu_{M_{jk}}(C_{t-(L-j+1)d}), \quad j = 1, 2, \dots, L; k = 1, 2; i = 1, 2, \dots, 2L \quad (2)$$

where $i = 2(j-1) + k$, C representing $C3$ or $C4$ is the input to node i , and M_{jk} is the linguistic label associated with this node function. The bell-shape Gaussian membership function $\mu_{M_{jk}}(C)$ is used

$$\mu_{M_{jk}}(C_{t-(L-j+1)d}) = \exp\left(-\left(\frac{C_{t-(L-j+1)d} - a_{jk}}{\sigma_{jk}}\right)^2\right) \quad (3)$$

where the parameter set $\{a_{jk}, \sigma_{jk}\}$ adjusts the shape of the Gaussian membership function. Parameters M_{jk} in this layer are referred to as premise parameters.

Layer 2: Each node in this layer is a circle node labeled Π multiplying the incoming signals together and sends out their product.

$$O_i^2 = w_i = \prod_{j=1}^L \mu_{M_{ji}}(C_{t-(L-j+1)d}), \quad i = 1, 2 \quad (4)$$

Each node output represents the firing strength of a rule.

Layer 3: Each node in this layer is a circle node labeled N . The firing strength of a rule for each node in this layer is normalized.

$$O_i^3 = \bar{w}_i = \frac{w_i}{\sum_j w_j}, \quad i = 1, 2 \quad (5)$$

Layer 4: Each node in this layer is a square node with its node function represented as

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i \left(\sum_{j=1}^L p_{ij} x_j + r_i \right), \quad i = 1, 2 \quad (6)$$

where the output f_i is a linear combination of the parameter set $\{p_{ij}, r_i\}$. Parameters f_i in this layer is referred to as consequent parameters.

Layer 5: The single node in this layer is a circle node labeled Σ computing the overall output y as the sum of all incoming signals.

$$O_1^5 = y = \sum_i \bar{w}_i f_i = \frac{\sum_j w_j f_j}{\sum_j w_j} \quad (7)$$

The architecture of neuron-fuzzy prediction in this chapter is shown in Fig. 3. The consequent parameters are updated by the RLS learning procedure in the forward pass for ANFIS network learning, while the antecedent parameters are adjusted by using the error between the predicted and actual signals. The parameter optimization for ANFIS training is adopted an approach that is mixed least squares and back-propagation method. Two

ANFISs are used to perform prediction. That is, they labeled IANFIS and rANFIS are used to predict left and right training MI EEG data, respectively. The actual filtered signals and their predicted results for C3 and C4 channels are shown in Fig. 2(b) and 2(c), respectively.

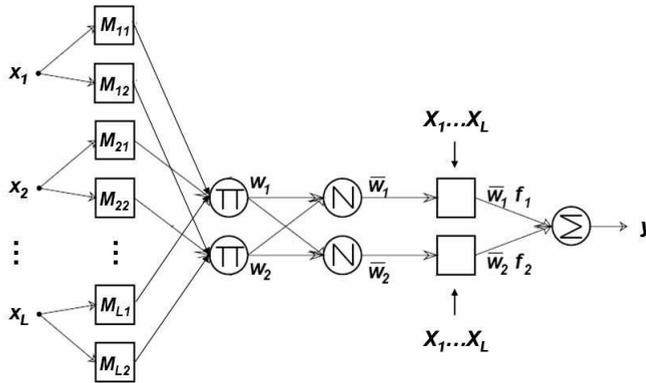


Fig. 3. Architecture of neuro-fuzzy prediction.

4.3 Feature extraction

After IANFIS and rANFIS are trained by using the left and right MI training data trial by trial respectively, they are used to perform one-step-ahead prediction. The test data are then input to these two ANFISs sample by sample, and features are extracted by continually calculating the difference of MFFVs between the predicted and actual signals as the length of predicted signals achieves 1-s window. The MFFV will be outlined in the next paragraph. In this study, feature extraction is performed on the 1-s window of predicted signals instead of directly classifying native predicted signals. A flowchart of feature extraction is shown in Fig. 4.

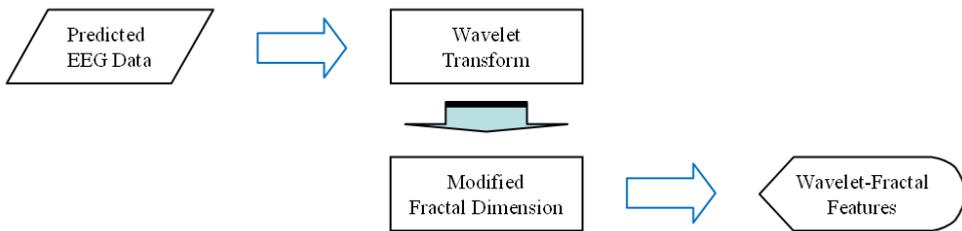


Fig. 4. Flowchart of feature extraction.

A signal is decomposed into numerous details in multiresolution analysis, where each scale represents a class of distinct physical characteristics within the signal. Wavelet transform is used to achieve multiresolutional representation in this study [21, 33, 36–39]. The 1-s segment is decomposed into numerous non-overlapping subbands by wavelet transform.

Fractal geometry provides a proper mathematical model to describe a complex shape that exists in nature with fractal features. Since fractal dimension is relatively insensitive to

signal scaling and shows a strong correlation with human judgment of surface roughness [20], it is chosen as the feature extraction method. A variety of approaches were proposed to estimate fractal dimension from signals or images [21–24]. A differential box counting (DBC) method covering a wide dynamic range with a low computational complexity is modified and used in this study [33]. A MFFV is extracted by modified fractal dimension from all the non-overlapping subbands of a 1-s segment.

The MFFV reflects the roughness and complexity of non-overlapping subbands of a signal. These MFFV calculations reduce prediction cost from a 1-s window to a feature vector for each signal. Features are extracted by continually calculating the difference of MFFVs between the predicted and actual signals as the length of predicted signals achieves 1-s window. In other words, two sets of MFFV features are first extracted from the predicted and actual signals respectively as the length of predicted signals achieves 1-s window. They are then subtracted for each respective subband. Finally, features are obtained by continually calculating their difference. The left and right test data are input to both the IANFIS and rANFIS, and each ANFIS provides two predictions from the C3 and C4 channels. Accordingly, four sets of MFFVs can be extracted after each new set of predictions is obtained. Each time a new set of predictions is produced, the oldest one is removed from the 1-s segment and a new MFFV is then extracted from the signals within the window. Since a large window is too redundant for the real time application, a 1-s window is short and selected for feature extraction. The length of a 1-s segment is a compromise between the computation cost and event-related potential (ERP) component applications. If the window length is selected properly, the extracted MFFVs will produce the maximum feature separability and obtain the highest classification accuracy.

4.4 Classification

It can be difficult to establish stable NNs since appropriate number of hidden layers and neurons usually need to carefully choose to approximate the function in question to the desired accuracy. The SVM first proposed by Vapnik [25] not only has a very steady theory in statistical learning, but guarantees to obtain the optimal decision function from a set of training data. The main idea of SVM is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. The SVM optimization problem is

$$\min_w \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \quad (8)$$

subject to $\xi_i \geq 0, \forall i$, and $d_i (w^T x_i + b) \geq 1 - \xi_i, \forall i = 1, 2, \dots, N$

where $g(x) = w^T x + b$ represents the hyperplane, w is the weighting vector, b is the bias term, x is the training vector with label d , C is the weighting constant, and ξ is the slack variable. It is then transformed into a convex quadratic dual problem. The discriminant function with optimal w and b , $g(x) = w_o^T x + b_o$, posterior to the optimization form becomes

$$g(x) = \sum_{i=1}^N \alpha_i d_i K(x, x_i) + b_0 \quad (9)$$

where α is a Lagrange multiplier and $K(x, x_i)$ is a kernel function. Generally, appropriate kernel functions are the polynomial kernel function $K(x_i, x_j) = (x_i^T x_j + 1)^p$ and the radial basis function (RBF) kernel function $K(x_i, x_j) = \exp\left(\left(-1/2\sigma^2\right)\|x_i - x_j\|^2\right)$. In this study, the latter is chosen for the SVM.

In the proposed system, classification is performed on MFFVs for recognizing the corresponding state at the sample rate. A different SVM classifier at each sample point is produced to classify each set of MFFVs for the training data. The classification sample point possessing maximal classification rate for training data is used as the standard classifier, which will be used for all classification performed on the test data. The best parameters selected from the training data are then applied to the test data to estimate the classification accuracy of test data.

5. Results

5.1 Performance of prediction methods

To assess the performance of proposed time-series prediction method, several prediction methods combined with power spectra features are implemented for comparison. They are AAR-parameter approach and AAR time-series prediction. The power spectra features are obtained by calculating the powers at the alpha and beta bands [16, 17]. The AAR-parameter method is an AAR signal modeling approach. The all-pole AAR model lends itself well to modeling the EEG as filtered white noise with certain preferred energy bands. The EEG time series is fitted with an AAR model. In the experiments, the order of AAR model is chosen as six and the AAR parameters are estimated with the RLS algorithm. To select the best value for the order of AAR model, an information theoretic approach is adopted [3]. The AAR parameters are used as features at each sample point for each trial. The AAR time-series prediction method is a time-series prediction approach, where left and rights ANFISs in the

Classification Accuracy [%]	AAR Parameters	AAR Prediction	Neuro-Fuzzy Prediction
S1	71.5	81.4	86.9
S2	66.3	76.6	84.2
S3	64.9	78.3	77.2
S4	72.6	79.6	88.6
S5	65.7	73.1	80.1
S6	61.0	77.0	79.8
Average	67.0	77.7	82.8

Table 1. Comparison of performance among different time-series prediction frameworks using power spectra features

ANFIS time-series prediction method are replaced by left and right AAR models. The lengths of windows for the AAR-parameter approach and AAR time-series prediction are all 1-s windows, which are the same as that for the ANFIS time-series prediction.

The comparison results of classification accuracy among different time-series prediction using power spectra features are listed in Table 1. The average classification accuracy of AAR-parameter approach is 67.0%, while AAR time-series prediction is 77.7% in the average classification accuracy. ANFIS time-series prediction obtains the best average classification accuracy (82.8%).

5.2 Performance of features

To further estimate the performance of proposed ANFIS time-series prediction method and MFFV features, ANFIS time-series prediction method combined with power spectra features is used for comparison in Table 2. The average classification accuracy for ANFIS time-series prediction method combined with power spectra features is 82.8%, while MFFV features under ANFIS time-series prediction method obtain 91.0 in the average classification accuracy.

Classification Accuracy [%]	Power Spectra	MFFV
S1	86.9	92.8
S2	84.2	88.5
S3	77.2	90.3
S4	88.6	93.9
S5	80.1	88.2
S6	79.8	92.0
Average	82.8	91.0

Table 2. Comparison of performance between power spectra and MFFV features under the use of ANFIS time-series prediction

5.3 Statistical analysis

Two-way analysis of variance (ANOVA) and multiple comparison tests [40] are performed in the experiments. The statistical analyses with two-way ANOVA are used to evaluate that the difference is significant or not for the two factors, methods and subjects. After analyzing with the two-way ANOVA, multiple comparison tests are used to estimate the p -values and significance of each pair of methods. The results of tests will be discussed in detail in the next section.

6. Discussion

6.1 Statistical evaluation of prediction methods

ANFIS combines the advantage of NN with that of FIS. Moreover, the training of ANFIS is fast and it can generally converge from small data sets. These attractive properties are

suitable for the prediction of non-stationary EEG signals. Table 1 lists the comparisons of performance among different prediction frameworks using power spectra features. In addition, two-way ANOVA and multiple comparison tests are performed to verify if the prediction methods are significantly different or not. The results indicate that AAR time-series prediction method is much better than AAR parameter approach in classification accuracy (p -value 0.0007) that is improved by 10.7% on average, while ANFIS time-series prediction method is slightly better than AAR prediction method (p -value 0.0195). The classification accuracy increases by 5.1%. Accordingly, ANFIS time-series prediction has the best performance in classification accuracy among these three methods. The results deduce that ANFIS time-series prediction is the best prediction framework in MI classification.

6.2 Statistical evaluation of features

Wavelet-fractal features are extracted from wavelet data by modified fractal dimension. MFFVs are utilized to describe the characteristic of fractal features in different wavelet scales, which are greatly beneficial for the analysis of EEG data. The comparison of performance between power spectra and MFFV features under the use of ANFIS time-series prediction is listed in Table 2. In addition, two-way ANOVA and multiple comparison tests are performed again to validate whether the two features are significantly different. The results indicate that MFFV features are significantly better than power spectra features in classification accuracy (p -value 0.0030), which is improved by 8.2% on average. The results indicate that MFFV features are better. These two results also suggest that ANFIS prediction framework together with MFFV features is a good combination in BCI applications.

6.3 Advantage of proposed method

The proposed ANFIS prediction framework combined with MFFV features provides a good potential for EEG-based MI classification. Furthermore, the proposed method has other potential advantages as follows: Firstly, the MFFV features really improve the separability of MI data, because the power spectra feature extracted from the predicted signals results in poorer performance. Secondly, the MFFV features can effectively reduce the degradation of noise. In other words, the MFFV features are extracted by DWT and modified fractal dimension. The former obtains multiscale information of EEG signals while the latter decreases the effect of noise. It is because the calculation of an improved DBC method is proposed and applied to modified fractal dimension.

7. Conclusion

We have proposed a BCI system embedding neuro-fuzzy prediction in feature extraction in this work. The results demonstrate the potential for the use of neuro-fuzzy prediction together with support vector machine in MI classification. It also shows that the proposed system is robust for the inter-subject use under careful parameter training, which is important for BCI applications. Compared with other well-known approaches, neuro-fuzzy prediction together with SVM achieves better results in BCI applications. In future works, more effective prediction/features and powerful classifiers will be used to further improve classification results.

8. Acknowledgements

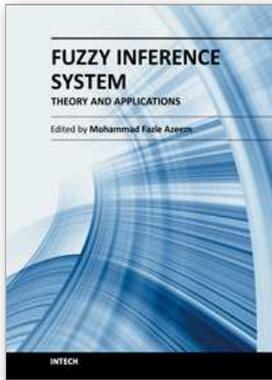
The author would like to express his sincerely appreciation for grant under shared facilities supported by the Program of Top 100 Universities Advancement, Ministry of Education, Taiwan.

9. References

- [1] R. Leeb, F. Lee , C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain-computer communication: Motivation, aim, and impact of exploring a virtual apartment," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 15, no. 4, pp. 473-482, 2007.
- [2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767-791, 2002.
- [3] W. Y. Hsu, "EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features," *Journal of Neuroscience Methods*, vol. 189, no. 2, pp. 295-302, June 2010.
- [4] W. Y. Hsu, C. Y. Lin, and W. F. Kuo, "Unsupervised Fuzzy C-Means Clustering for Motor Imagery EEG Recognition," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 8, pp. 4965-4976, Aug. 2011.
- [5] W. Y. Hsu, "EEG-based Motor Imagery Classification using Enhanced Active Segment Selection and Adaptive Classifier," *Computers in Biology and Medicine*, vol. 41, no. 8, pp. 633-639, Aug. 2011.
- [6] W. Y. Hsu, "Continuous EEG signal analysis for asynchronous BCI application," *International Journal of Neural Systems*, vol. 21, no. 4, pp. 335-350, Aug. 2011.
- [7] W. Y. Hsu, "Application of Multiscale Amplitude Modulation Features and FCM Clustering to Brain-Computer Interface," accepted for published in *Clinical EEG and Neuroscience*, June 2011, In Press.
- [8] L. Parra, C. Alvino, A. C. Tang, B. A. Pearlmutter, N. Yeung, A. Osman, and P. Sajda, "Linear spatial integration for single trial detection in encephalography," *NeuroImage*, vol. 7, no. 1, pp. 223-230, 2002.
- [9] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 12, no. 2, pp. 258-265, 2004.
- [10] G. Pfurtscheller, and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842-1857, 1999.
- [11] R. H. Shumway, and D. S. Stoffer, "Times Series Analysis and its application," New York: *Springer-Verlag*, 2000.
- [12] X. Hu, D. V. Prokhorov, I. I. Wunsch, and C. Donald, "Time series prediction with a weighted bidirectional multi-stream extended Kalman filter," *Neurocomputing*, vol. 70, pp. 2392-2399, 2007.
- [13] N. Stamatis, D. Parthimos, and T. M. Griffith, "Forecasting chaotic cardiovascular time series with an adaptive slope multilayer perceptron neural network," *IEEE Trans. Biomedical Engineering*, vol. 46, no. 12, pp. 1441-1453, 1999.

- [14] H. J. Rong, N. Sundararajan, G. B. Huang, and P. Saratchandran, "Sequential Adaptive Fuzzy Inference System (SAFIS) for nonlinear system identification and prediction," *Fuzzy Sets and Systems*, vol. 157, no. 9, pp. 1260-1275, 2006.
- [15] J. S. Jang Roger, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. SMC*, vol. 23, no. 3, pp. 665-685, 1993.
- [16] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller, "Information transfer rate in a five-classes brain-computer interface," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 9, pp. 283-288, 2001.
- [17] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 11, pp. 145-147, 2003.
- [18] D. P. Burke, S. P. Kelly, P. de Chazal, R. B. Reilly, and C. Finucane, "A parametric feature extraction and classification strategy for brain-computer interfacing," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 13, pp. 12-17, 2005.
- [19] C. Guger, A. Schlögl, C. Neuper, D. Walterspacher, T. Strein, and G. Pfurtscheller, "Rapid prototyping of an EEG-based brain-computer interface (BCI)," *IEEE Trans. Rehab. Eng.*, vol. 9, no. 1, pp. 49-58, 2001.
- [20] B. B. Mandelbrot, "Fractal geometry of nature," *Freeman Press*, San Francisco, 1982.
- [21] W. L. Lee, Y. C. Chen, and K. S. Hsieh, "Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform," *IEEE Trans. Medi. Imag.*, vol. 22, pp. 382-392, 2003.
- [22] P. Paramanathan, and R. Uthayakumar, "Application of fractal theory in analysis of human electroencephalographic signals," *Computers in Biology and Medicine*, vol. 38, no. 3, pp. 372-378, 2008.
- [23] G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram, S. Wimalaratna, C. Del Percio, and F. Vecchio, "Development and assessment of methods for detecting dementia using the human electroencephalogram," *IEEE Trans. Biomedical Engineering*, vol. 53, no. 8, pp. 1557-1568, 2006.
- [24] L. Rankine, N. Stevenson, M. Mesbah, and B. Boashash, "A nonstationary model of newborn EEG," *IEEE Trans. Biomedical Engineering*, vol. 54, no. 1, pp. 19-28, 2007.
- [25] B. E. Boser, I. M. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," *Proc. 5th Annu. Workshop on Computational Learning Theory*, pp. 144-152, 1992.
- [26] K. Sadatnezhad, R. Boostani, and A. Ghanizadeh, "Proposing an adaptive mutation to improve XCSF performance to classify ADHD and BMD patients," *J. Neural Eng.*, vol. 7, pp. 066006, 2010.
- [27] M. Nandan, S. S. Talathi, S. Myers, W. L. Ditto, P. P. Khargonekar, and P. R. Carney, "Support vector machines for seizure detection in an animal model of chronic epilepsy," *J. Neural Eng.*, vol. 7, pp. 036001, 2010.
- [28] A. Schlogl, C. Keinrath, R. Scherer, and G. Pfurtscheller, "Estimating the mutual information of an EEG-based brain computer interface," *Biomedizinische Technik*, vol. 47, pp. 3-8, 2002.
- [29] A. Schlogl, F. Lee, H. Bischof, and G. Pfurtscheller, "Characterization of four-class motor imagery EEG data for the BCI-competition 2005," *J. Neural Eng.*, vol. 2, pp. L14-L22, 2005.

- [30] E. Haselsteiner, and G. Pfurtscheller, "Using time-dependent NNs for EEG classification," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 457-462, 2000.
- [31] Graz Data Sets and description for the BCI 2003 competition. [Online]. Available: <http://ida.first.fraunhofer.de/projects/bci/competition/>
- [32] Graz Data Sets and description for the BCI 2005 competition. [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_iii/
- [33] W. Y. Hsu, C. C. Lin, M. S. Ju, and Y. N. Sun, "Wavelet-based fractal features with active segment selection: Application to single-trial EEG data," *Journal of Neuroscience Methods*, vol. 163, no. 1, pp. 145-160, 2007.
- [34] H. Jasper, "Report of committee on methods of clinical exam in EEG," *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, pp. 370-375, 1958.
- [35] K. Y. Lian, C. H. Su, and C. S. Huang, "Performance enhancement for T-S fuzzy control using neural networks," *IEEE Trans. Fuzzy Systems*, vol. 14, no. 5, pp. 619-627, 2006.
- [36] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.
- [37] W. Y. Hsu, P. W. F. Poon, and Y. N. Sun, "Automatic seamless mosaicing of microscopic images: enhancing appearance with colour degradation compensation and wavelet-based blending," *Journal of Microscopy*, vol. 231, no. 3, pp. 408-418, 2008.
- [38] W. Y. Hsu, and Y. N. Sun, "EEG-based motor imagery analysis using weighted wavelet transform features," *Journal of Neuroscience Methods*, vol. 176, no. 2, pp. 310-318, 2009.
- [39] W. Y. Hsu, "Wavelet-based Envelope Features with Automatic EOG Artifact Removal: Application to Single-Trial EEG Data," *Expert Systems with Applications*, Aug. 2011, DOI:10.1016/j.eswa.2011.08.132, In Press.
- [40] F. J. Anscombe, "The Validity of Comparative Experiments," *Journal of the Royal Statistical Society. Series A (General)*, vol. 111, no. 3, pp 181-211, 1948.



Fuzzy Inference System - Theory and Applications

Edited by Dr. Mohammad Fazle Azeem

ISBN 978-953-51-0525-1

Hard cover, 504 pages

Publisher InTech

Published online 09, May, 2012

Published in print edition May, 2012

This book is an attempt to accumulate the researches on diverse inter disciplinary field of engineering and management using Fuzzy Inference System (FIS). The book is organized in seven sections with twenty two chapters, covering a wide range of applications. Section I, caters theoretical aspects of FIS in chapter one. Section II, dealing with FIS applications to management related problems and consisting three chapters. Section III, accumulates six chapters to commemorate FIS application to mechanical and industrial engineering problems. Section IV, elaborates FIS application to image processing and cognition problems encompassing four chapters. Section V, describes FIS application to various power system engineering problem in three chapters. Section VI highlights the FIS application to system modeling and control problems and constitutes three chapters. Section VII accommodates two chapters and presents FIS application to civil engineering problem.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Wei-Yen Hsu (2012). Neuro-Fuzzy Prediction for Brain-Computer Interface Applications, Fuzzy Inference System - Theory and Applications, Dr. Mohammad Fazle Azeem (Ed.), ISBN: 978-953-51-0525-1, InTech, Available from: <http://www.intechopen.com/books/fuzzy-inference-system-theory-and-applications/neuro-fuzzy-prediction-for-brain-computer-interface-applications>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.