New Biometric Approach Based on Motor Imagery EEG signals

Jian-feng HU
Institute of Information Technology
Jiangxi Bluesky University,
Nanchang, Jiangxi 330098, China
e-mail: huguess@21cn.com

***Abstract* — A research on biometry based on motor imagery EEG signals was described. In this study, I select EEG signals related to motor imagery, and a model was built. Estimated model parameters as feature vector were extracted, and then to classified by an artificial neural network. Two different classify cases, including authentication and identification, were investigated. Four types of motor imagery EEG signals and three subjects were compared. Experiment results show that EEG carrying individual-specific information can be successfully exploited for purpose of person authentication and identification.**

Keywords- Biometric, Electroencephalogram (EEG), Nonlinear analysis, ARMA model.

I. INTRODUCTION

The need for a biometric to identify trusted people is irrefutable. Biometric system measures unique physical or behavioral characteristics of individuals as a means to recognize or authenticate their identity. Common physical biometrics includes fingerprint, hand or palm geometry, and retina, iris, or facial characteristics. Behavioral characteristics include signature, voice, keystroke pattern and gait. The first shortcoming of above biometric security system is the discrimination of groups of people whose biometrics cannot be recorded well. For example, people whose fingerprints do not print well on, they even miss the required feature. The second shortcoming is such as: face can be copied using a photograph, and a voice can be printed using a voice recording. Another problem is the verification of the subject's aliveness. I propose an EEG signature (EEG password) as a new biometric. EEG signature analysis requires detection of large dissimilarity between the known EEG signature and test data from other people, as well as recognition of insignificant change between the known EEG signature and test data from the same person.

The usage of the EEG signals for person identification is still the least explored possibility of all [1-8]. Poulos et al. [1-2] adopted two classification algorithms and obtained the accuracies of around 80% and 95% respectively. Paranjape et al. [3] analyzed a data set of 40 subjects and 349 EEG trials and got a classification accuracy of about 80%. Palaniappan and Mandic [4] carried out a personal identification experiment with 102 subjects based on visual evoked potentials and the accuracies were around 95-98%. With a data set of nine subjects performing mental imaginary tasks of left hand movements, right hand movements and word generation beginning with a same letter, Marcel and Millán [5] got a highest accuracy rate for personal verification of 93.4%. The above early work has played an important role in studying the feasibility of EEG signals for usage in biometrics. Motor imagery EEG signals have been popularly applied in the research of Brain-Computer Interface (BCI) [9-13]. In this paper, I make an attempt to analysis EEG signatures from different motor imagery movements based on an ARMA (Auto-Regressive and Moving Average) model.

II. MATERIALS

The dataset of BCI competition 2003 was used in this investigation. This dataset was provided by Graz University of Technology, which included three subjects – K3b, K6b and L1b.

The recording was made with a 64-channel EEG amplifier from Neuroscan, using the left mastoid for reference and right mastoid as ground. The EEG was sampled with 250 Hz, and it was filtered between 1 and 50 Hz with Notchfilter on.

The subject sat in a relaxing chair with armrests. The task was to perform imagery left hand, right hand, foot or tongue movements according to a cue. The order of cues was random. The experiment consists of several runs with 40 trials each after each; after trial begin, the first 2s were quite, at t=2s an acoustic stimuli indicated the beginning of the trial, and a cross symbol is displayed; then from t=3s an arrow to the left, right, up or down was displayed for 1s; at the same time the subject was asked to imagine a left hand, right hand, foot or tongue movements, respectively, until the cross symbol disappeared t=7s (as shown in Fig. 1). Each of the 4 cues was displayed 10 times within each run in a randomized order.

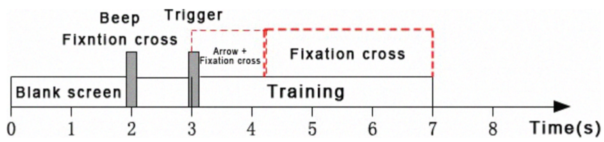

Figure 1. Timing of the paradigm

III. METHODS

In this study the EEG data from the electrodes C3, C4, P3, P4, O1 and O2 only are analyzed. These channels are converted to Hjort derivation in order to get the local activity [14]. Specifically, let x_i be the Hjort derivation corresponding to $x_i^{'}$. The Hjort derivation x_i is calculated as

$$x_i = x_i^{'} - \frac{1}{8}\sum_{j \in S_i} x_j^{'} \quad (1)$$

Where S_i is the reading of the center $x_i^{'}$ and S_i is the set of indices corresponding to the four electrodes surrounding $x_i^{'}$.

The EEG data are finally bandpass filtered and only information between 2 and 40 Hz was retained. I used all 90 trials available for each task to test the classification approaches discussed below.

A. *ARMA model* [15]

Both linear and non-linear components co-exist in the EEG signals. In order to model the linear component of an EEG signal, a linear model of Auto-Regression and Moving Average type, ARMA(p,q) is fitted to the EEG signal x_t. This model can be written as:

$$x_t = \sum_{i=0}^{q} c_i e_{t-i} + \sum_{i=1}^{p} a_i x_{t-i} \quad (2)$$

Where c_0=1, e_t is an independent, identically distributed driving noise process with zero mean and unknown variance σ_e^2, and other parameters {a_i, i=1,2,…,p; c_i, i=1,2,…,q } are unknown constants with respect to time.

In order to explain nonlinear component of the EEG signal, the above model can be modified as:

$$x_t = \sum_{i=0}^{q} c_i e_{t-i} + \sum_{i=1}^{p} a_i x_{t-i} + \sum_{i=1}^{k}\sum_{j=1}^{m} b_{ij} x_{t-i} e_{t-j} \quad (3)$$

Where c_0=1, e_t is an independent, identically distributed driving noise process with zero mean and unknown variance σ_e^2, and other parameters {a_i, i=1,2,…,p; c_i, i=1,2,…,q; b_{ij}, i=1,2,…,k; j=1,2,…,m } are unknown constants with respect to time.

It can be seen that equation (3) is produced from equation (3) with the addition of extra component $\sum_{i=1}^{k}\sum_{j=1}^{m} b_{ij} x_{t-i} e_{t-j}$. In this way, the model is considered which contain both the non-linear and linear components.

In this work, the elimination of the MA part of the model is compromise in the model type in order to facilitate the parameter estimation step. Thus q=0 is assumed. The model can be simplified into:

$$x_t = \alpha + \sum_{i=1}^{p} a_i x_{t-i} + \sum_{i=1}^{k}\sum_{j=1}^{m} b_{ij} x_{t-i} e_{t-j} + \mathrm{e}_t \quad (4)$$

Parameter α is allowed fitting of the model to non-zero mean data. In our case, EEG data are zero mean before further processing. So α can omitted from the model. Parameters {a_i, i=1,2,…,p; b_{ij}, i=1,2,…,k; j=1,2,…,m} are unknown constants with respect to time. It's difficult to estimate the values of these parameters due to unknown p, k, and m.

The choice of the order of the model is usually based on information theory criteria such as the Akaike Information Criterion (AIC) which is given by:

$$\mathrm{AIC}(r) = (N - M)\log\sigma_e^2 + 2r \quad (5)$$

$$\sigma_e^2 = \frac{1}{N-M}\sum_{t=M=1}^{N} e_t^2 \quad (6)$$

N is the length of the data record, M is the maximal order employed in the model, (N-M) is the number of data samples used for calculating the likehood function, r is the number of independent parameter presented in the model, and the optimal r is the minimum of AIC(r).

I have used the AIC to determine the order of the linear part of the model in equation (4). It can be seen that AIC (p) takes on its minimum values for model orders p ranging between 8 and 13, record–independent. For simplicity, I have set the model order p=8. Similarly, the other orders of the model m and k can be set experimentally to k=2 and m=3. Higher values of m and k have not been considered due to practical consideration (too complex).

For given (p,k,m), the parameters of model {a_i, i=1,2,…,p; b_{ij}, i=1,2,…,k; j=1,2,…,m } are estimated by linear minimization using standard least squares techniques.

B. *Feature extraction*

The parameters are estimated from the equation (3), where the model is fitted to real EEG signals. A set of such parameters {a_i, i=1,2,…,p; b_{ij}, i=1,2,…,k; j=1,2,…,m} obtained from a single EEG signal. These 8+2*3=14 parameters are used as feature vectors. In the data set, 36 EEG recordings (k3b), 20 EEG recordings (k6b and 11b) were analyzed. Take account of six electrodes, thus 36*6=216 data segments for k3b, 20*6=120 data segments for k6b and l1b, respectively for each type of motor imagery, were investigated.

C. *Classfication*

Multilayer back-propagation neural networks were trained to classify different subjects. The input layer had 14 units, the only output unit was valued between 0 and 1.

The traindx learning algorithm has been used to train the network, which uses gradient descent with momentum and variable learning rate in batch learning mode. For each subject, the neural networks were trained using 50% of the trials (training set). Subsequently, the trained network was used to classify the remaining 50% of the trials (testing set).

Two different types of classification are employed, as described in the following:

(1) Authentication procedure, aims to differentiate between a target subject and a pool of other subjects. Three different subjects of target, k3b, k6b and l1b, participate in three examples of the case, where group T serves as the pool class in all three cases. Taking these account I use the BP neural network in a setup of two target classes. Feature vectors include a_i and b_{ij} from estimation procedure. They are of dimensionality 14*1. Three different BP neural

networks are thus trained using feature vectors. Each network has two target classes [k3b, T], [k6b, T] or [l1b, T], respectively. In the experimental setup, a successfully classified feature vectors would thus be classified as belonging to class k3b or k6b or l1b, respectively. Class T serves as a non-target class for feature vectors (k3b or k6b or l1b, respectively).

In the testing process, all remaining feature vectors are classified by the three trained networks. Training and test set are thus disjoined.

The architecture of BP neural network used to classify the feature vectors is shown in Fig. 2. Input vector of dimensionality 14*1 are weighted and fed to hidden layer. Five neurons form hidden layer in our case, which groups subclasses into target classes. Two target classes exist here, the class of interest (k3b or k6b or l1b, respectively) and class T, serving as the "authentication" and "denial" of the class of interest.

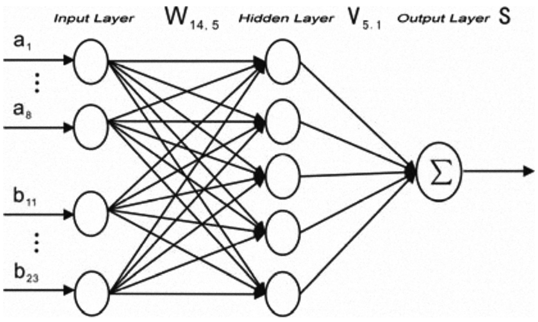

Figure 2 Architecture of BP neural network employed for the classification

(2) In the identification process, three subjects k3b, k6b and l1b are to be classified, in contrast to authentication process. This test is carried out to verify that the proposed neural network has the ability to correctly classify EEG features within a multi-person group.

IV. Results

With the selected configuration of hidden units, I run each neural network ten times with different random initializations. Then average accuracies were obtained.

A. *Authentication*

Results are tabulated in table 1-5 for imagining left hand, right hand, tongue, foot movement and average, respectively. As it can be seen in the table 1, correct positive classification scores (i.e. subject k3b classified as k3b, k6b as k6b, l1b as l1b, respectively) ranging from 75.0% to 85.0%. Table 1-4 show that authentication accuracy is dependent on motor imagery type.

Table 1. Authentication accuracy for left hand movement imagery

	k3b	k6b	l1b	Non-T
k3b	88/108=81.5%			20/108=18.5%
k6b		51/60=85.0%		9/60=15.0%
l1b			45/60=75.0%	15/60=25.0%
average		80.5%		

Table 2. Authentication accuracy for right hand movement imagery

	k3b	k6b	l1b	Non-T
k3b	89/108=82.4%			19/108=17.6%
k6b		50/60=83.3%		10/60=16.7%
l1b			46/60=76.7%	14/60=23.3%
average		80.8%		

Table 3. Authentication accuracy for tongue movement imagery

	k3b	k6b	l1b	Non-T
k3b	99/108=91.7%			9/108=8.3%
k6b		54/60=90.0%		6/60=10.0%
l1b			58/60=96.7%	2/60=3.3%
average		92.8%		

Table 4. Authentication accuracy for foot movement imagery

	k3b	k6b	l1b	Non-T
k3b	87/108=80.6%			21/108=19.4%
k6b		49/60=81.7%		11/60=18.3%
l1b			49/60=81.7%	11/60=18.3%
average		81.3%		

Table 5. Average Authentication accuracy for four types of motor imagery

	k3b	k6b	l1b
k3b	84.1%		
k6b		85.0%	
l1b			82.5%
average		83.9%	

B. *Identification*

Results are tabulated in table 6-10 for imagining left hand, right hand, tongue, foot movement and average, respectively. As it can be seen in the table 6, correct positive classification scores (i.e. subject k3b classified as k3b, k6b as k6b, l1b as l1b, respectively) ranging from 75.0% to 78.3%. Table 6-9 show that identification accuracy is dependent on motor imagery type.

Table 6. Identification accuracy for left hand movement imagery

	k3b	k6b	l1b
k3b	83/108=76.9%	18/108=16.7%	7/108=6.5%
k6b	10/60=16.7%	47/60=78.3%	3/60=5.0%
l1b	8/60=13.3%	7/60=11.7%	45/60=75.0%
average		76.7%	

Table 7. Identification accuracy for right hand movement imagery

	k3b	k6b	l1b
k3b	83/108=76.9%	17/108=15.7%	8/108=7.4%
k6b	10/60=16.7%	48/60=80.0%	2/60=3.3%
l1b	1/60=1.7%	13/60=21.7%	46/60=76.7%
average		77.9%	

Table 8. Identification accuracy for tongue movement imagery

	k3b	k6b	l1b
k3b	99/108=91.7%	8/108=7.4%	1/108=0.9%
k6b	4/60=6.7%	55/60=91.7%	1/60=1.7%
l1b	2/60=3.3%	2/60=3.3%	56/60=93.3%
average		92.2%	

Table 9. Identification accuracy for foot movement imagery

	k3b	k6b	l1b
k3b	84/108=77.8%	10/108=9.3%	4/108=3.7%
k6b	5/60=8.3%	50/60=83.3%	5/60=8.3%
l1b	8/60=13.3%	3/60=5.0%	49/60=81.7%
average		80.9%	

Table 10. Average Identification accuracy for four types of motor imagery

	k3b	k6b	l1b
k3b	80.8%		
k6b		83.3%	
l1b			81.7%
average		81.9%	

V. Conclusion

A biometric system based on EEG signals is essentially a pattern recognition system that establishes a person's identity by comparing the binary code of a uniquely specific EEG signature to the binary code of a stored characteristic. This accomplished by acquiring an EEG signature from a petitioner. The system then applies a complex and specialized algorithm to the EEG signature; it then convert into a binary code. Once the EEG signature has been converted into a binary code, it is compared to the reference signature to determine the petitioner's access or not.

The results obtained in the present work corroborate the long existing line of research showing evidence that EEG carrying individual-specific information which can be successfully exploited for purpose of person authentication and identification. An ARMA model is fitted to the EEG data and the estimated parameters are used as feature vectors. For the purpose of EEG-based person authentication and identification, a BP neural network classifier is used to classify three subjects.

The classification accuracy is dependent on the type of motor imagery movement. In the authentication process, an average 83.9% classification accuracy of the four types and three subjects was achieved. Different classification accuracies have been obtained from different movement types. The classification accuracy across movement types was relatively large in tongue movement imagery (92.8%, Table 3). In the identification process, an average 81.9% classification accuracy of the four types and three subjects was achieved. Similarly, the classification accuracy across movement types was also relatively large in tongue movement imagery (92.2%, Table 8).

The classification accuracy obtained here are not enough to allow for a direct application. In the future, investigating the performance of EEG based biometrics on a large database is worth studying.

Think of EEG signature as a key, it can open door for you and provides security to keep other not. It's key that can be customized to an individual's access needs. You can use EEG password to access your home, your account, or to invoke a customized setting for any security application.

ACKNOWLEDGMENT

This study was supported Natural Science Foundation of Jiangxi Province of China (2008GQS0003) and Foundation of Jiangxi Educational Committee (GJJ09622). The author is grateful to the anonymous reviewers who made constructive comments. He would also like to thank Bao X. C., Xiao D., Mu Z. D. Yin J. H. and Mao C. L. for various helps and suggestions.

REFERENCES

[1] M. Poulos, M. Rangoussi, V. Chrissikopoulos and A. Evangelou. "Person identification based on parametric processing of the EEG," Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems, 1999, 1:283– 286.

[2] M. Poulos, M. Rangoussi, V. Chrissikopoulos and A. Evangelou. "Parametric person identification from the EEG using computational geometry," Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems, 1999, 2:1005–1008.

[3] R. Paranjape, J. Mahovsky, L. Benedicenti and Z. Koles. "The electroencephalogram as a biometrics," Proceedings of the Canadian Conference on Electrical and Computer Engineering, 2001, 2:1363–1366.

[4] R. Palaniappan and D. Mandic. "Biometrics from brain electrical activity: A machine learning approach," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(4):738–742.

[5] S. Marcel and J del R Millán. "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(4): 743-752.

[6] A. Riera, M. Soria-Frisch, C. Caparrini, C. Grau and G. Ruffini. "Unobtrusive biometric system based on electroencephalogram analysis," EURASIP Journal on Advances in Signal Processing, 2008, 1:18-25.

[7] C. N. Gupta, R Palaniappan and S Swaminathan. "Novel analysis technique for a brain biometric system," International Journal of Medical Engineering and Informatics, 2008, 1(2): 266–273.

[8] R Palaniappan. "Two-stage biometric authentication method using thought activity brain waves," International Journal of Neural Systems, 2008, 18(1): 59-66.

[9] J. F. Hu, X. C. Bao and Z. D. Mu. "Classification of Motor Imagery EEG Based on Phase Synchronization," Microelectronics and Computer, 2008, 25(9):138-140.

[10] J. F. Hu, D. Xiao and Z. D. Mu. "Communication systems of Brain-Computer Interface based on second-order blind identification," Journal of Clinical Rehabilitative Tissue Engineering Research, 2008, 12(13):2481-2484.

[11] J. F. Hu，Z. D. Mu and D. Xiao. "Classification of motor imagery EEG signals based on energy entropy," Computer Engineering and Applications, 2008, 44(33):235-238.

[12] G Pfurtcheller. "Motor imagery and direct brain-computer communication," Proceedings of the IEEE, 2001, 89(7): 1123-1134.

[13] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan. "Brain-computer interfaces for communication and control," Clinical Neurophysiology, 113(6):767–791.

[14] B Hjorth. "An online transformation of EEG scalp potentials into orthogonal source derivations Electroencephalogr," Clin Neurophysiol, 1975, 39: 526–30.

[15] M. Poulos, M. Rangoussi, N. Alexandris and A. Evangelou. "Person Identification from the EEG using nonlinear signal classification," Methods of Information in Medicine, 2002, 41(1):64-75