CLASSIFICATION OF CHAOTIC SIGNALS USING HMM CLASSIFIERS: EEG-BASED MENTAL TASK CLASSIFICATION

Soroosh Solhjoo[1], *Ali Motie Nasrabadi*[2], *and Mohammad Reza Hashemi Golpayegani*[3]

[1, 3]Biomedical Engineering Faculty, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
[2] Biomedical Engineering Department, Engineering Faculty, Shahed University, Tehran, Iran
[1, 2] Neuropsychiatry Research, Research Center for Intelligent Signal Processing, Tehran, Iran
phone[1]: (+98-21) 64542394, fax[1]: (+98-21) 8063547, email[1]: soroosh@bme.aut.ac.ir

ABSTRACT

Mental task classification using brain signals, mostly electroencephalogram (EEG), is an approach to understand human brain functions. As EEG seems to be chaotic, it is important to verify the capability of probabilistic and statistical processing tools (such as HMM-based classifiers) in working with chaotic signals. At first, we study the performance of HMM's in classification of different classes of synthetically generated chaotic signals. Then performance of such classifiers in EEG-based mental task classification is studied. Results show good performance in both cases.

Keywords: *Chaos, Hidden Markov Models (HMM), EEG-Based Mental Task Classification*

1. INTRODUCTION

Introducing EEG signal, hopes to use it for communications and environment control with no need to limbs and peripheral nerves arose [1]. Nowadays, research groups are trying to provide such communication channels, namely Brain-Computer Interfaces, for patients suffering from severe motor control disabilities [1-3].

One approach to provide such an interface is mental task classification; i.e. each segment of EEG signal should be assigned to its appropriate class among the predefined classes of mental tasks. Different types of classifiers have been used for this purpose; such as, neural classifiers [2, 4-9] and HMM-based classifiers [10-13].

There are some evidences showing EEG signals' chaotic behavior [14-17]. As an approach to better understanding the nature of EEG signal and evaluating its chaotic behavior as well as assessment of compatibility of HMM's with chaos, this study is set to verify if chaotic signals can be well modeled and classified by these statistical and probabilistic models, HMM's which have shown very good results in speech processing and have been used growingly in EEG signal processing.

A dynamical chaotic system may be modeled as

$$s_{n+1} = f(s_n)$$

where s_n is the state of system at time n, and f is a nonlinear function. These states can be observed through $y_n = h(s_n)$ [18]. Modelling f and h by probability functions leads to HMM models of chaotic systems.

At first, we have used different types of synthetically generated chaotic signals, such as logistic map, tent map, and Lorenz model. For each type of these chaotic signals, we have assigned a parameter as the characteristic parameter of each class, and based on that parameter, the signals of that class are produced, while other parameters are constant in all classes of that type.

After transient state, signals are used for modeling and classification; i.e. the chaotic system is in its basin of attraction and the HMM model is to be the model of the basin of attraction.

Performance of classifiers is evaluated regarding the level of chaocity and the difference between levels of chaocity of those classes of signals which are to be classified. The criterion for the level of chaocity is assigned to be the Lyapunov exponent of the chaotic signal.

Different structures of discrete HMM (dHMM) and multi-Gaussian HMM (mHMM) classifiers have been studied to determine the best model for representing each type of chaotic signals (if it exists). Maximum performance in each experiment is reported.

Finally, we have studied different structures of HMM classifiers in classification of raw EEG signals related to different mental tasks.

The cross-validation procedure used in this study is PCV (randomized Permutation Cross Validation) and .

2. DATA, METHODS AND RESULTS

2.1 Synthetically Generated Chaotic Signals

We have studied classification of datasets of logistic map, tent map, and Lorenz model using dHMM and mHMM classifiers in this section.

For logistic map we have

$$x_{i+1} = a \times x_i \times (1 - x_i).$$

The parameter a is selected as the characterizing parameter of each class. At first we produced 8000 samples of data, and then 1000 initial samples were deleted to ignore the transient segment. 7000 remained samples were segmented into 70 segments of 100 samples.

To compute Lyapunov exponent, we have generated 600 points of logistic map with a random initial point for 20 times. Each time, Lyapunov exponent has been computed after deleting the transient phase. Then mean of these 20 quantities was selected as the Lyapunov exponent of that class of data.

For tent map, we have:

$$x_{i+1} = r \times (1 - 2 \times |x_i - 0.5|).$$

The parameter r is selected to be the characteristic parameter of each class. Data samples have been provided as for logistic map. Tables 1-3 show the characteristic parameters and relevant

Lyapunov exponents for datasets. (Each row shows information for each study on logistic and tent maps).

As logistic map and tent map are both discrete one-dimensional signals, use of a multi-dimensional chaotic model with continuous dynamicity; e.g. Lorenz model seems to be needed. This model produces three-dimensional signal, and combining the samples of these three dimensions, we have produced samples of our classes of signals. For Lorenz model we have:

$$\dot{x} = p(y - x)$$
$$\dot{y} = -y - xy + rx$$
$$\dot{z} = xy - bz$$

The parameter r was selected as the characteristic parameter. b = 8/3, and p = 10 were constant parameters. The initial points of the signal were selected randomly.

While natural signals are distorted by correlated noises, Gaussian noise of variance 0.01 has been added to synthetic chaotic signals to make them nearer to natural signals and the above studies have been performed for new set of signals.

For purpose of cross-validation, 50 segments of data were assigned randomly to train and 20 to test the classifier for 10 times; in each time the initial point of signal was selected randomly.

2.1.1 dHMM

The number of states of models was between 1 and 5, and the number of observable symbols (levels of quantization)/state was powers of two between 2^2 and 2^8. In training of HMM classifiers, initial probability matrices have been selected randomly. Maximum number of EM algorithm iterations in training the HMM models has been selected to be 50.

At first, we have studied the performance of dHMM in classification of signals produced by logistic map. Studies have shown if the difference between Lyapunov exponents is less then more complicated structures are required to result in a good (nearly 100%) classification.

In order to generalize the results, dHMM has been used to classify different classes of signals produced by tent map equation, as well. Different structures of dHMM classifiers showed to be 100% capable in classification of classes of tent map signals.

After that, classification of signals as logistic map or tent map has been considered. Results have shown the high capability (100%) of dHMM in classification of these two types of signals.

In the case of classification of Lorenz signals, for two classes with $r1$ = 90 and $r2$ = 80, the best classification percentage was 65.75 ± 6.24 % for dHMM with one state and 64 observable symbols/state.

For logistic map, the percentage of classification decreases a little when the noise is added to the signal. For example, for two datasets of logistic map with $a1$ = 3.90 and $a2$ = 3.80, the best classification percentage is 97.75 ± 1.84 % for models with 7 states and 16 observable symbols/state. For $a1$ = 3.90 and $a2$ = 3.925, we have 94.25 ± 5.78 % for models with 8 states and 32 observable symbols/state. In classification of signals to tent and logistic classes, we have nearly complete (99 %) classification accuracy for models of simple structures. For tent map classification the best performances are more than 96 %.

In both past steps, more complicated models lead to worse performance. This shows the importance of the quantization step in working with dHMM which applies as a filter to data and fits the model to it.

Table 1 **Lyapunov Exponents of logistic maps**

a1	L1	a2	L2	|L1 – L2|
3.9	0.4944	3.8	0.4309	0.0635
3.9	0.4944	3.925	0.5349	0.0405
3.9	0.4944	3.875	0.4583	0.0361
3.825	0.4020	3.8	0.4309	0.0289

Table 2 **Lyapunov Exponents of logistic map, tent map**

a	L	r	L	|L1 – L2|
3.9	0.4944	0.6	0.1823	0.3121
3.8	0.4309	0.6	0.1823	0.2486
3.9	0.4944	0.8	0.4700	0.0244

Table 3 **Lyapunov Exponents of tent maps**

r1	L1	r2	L2	|L1 – L2|
0.7	0.3365	0.8	0.4700	0.1335
0.9	0.5878	0.8	0.4700	0.1178
0.9	0.5878	0.7	0.3365	0.2513

For signals of Lorenz model, the performance was low and adding noise made it lower: for classes with $r1$ = 90 and $r2$ = 80, we have 61.9 ± 6.4023 % for dHMM with 5 states and 128 observable symbols/state.

2.1.2 mHMM

For chaotic signals, that a very small deviation (e.g. resulted by quantization) can lead to change of basin of attraction, mHMM is expected to produce better results than dHMM. Therefore, we have studied performance of mHMM-based classifiers in the same cases above. We have used mHMM's with between 1 and 10 states and between 1 and 10 Gaussian mixtures/state (100 different structures). Covariance matrix has been considered diagonal. Gaussian mixtures have been initiated by k-means and then trained by EM for at most 5 iterations.

At first logistic maps with $a1$ = 3.90, and $a2$ = 3.875 were considered. In the case of no noise, there is a very high accuracy of 98.5 ± 3.74% for 9 states and 1 Gaussian/state as the maximum accuracy. The best classification percentage for noise-contaminated data was 74.50 ± 11.55% for 10 states and 2 Gaussians/state. Because this structure was on the limit of the number of states (10 states), it is probable that there may be a structure with better performance out of the applied ranges. So we continued our study for structures having up to 30 states and 2 Gaussians/state. The best performance showed to be that of the structure with 19 states: 84.75 ± 7.45% [see Figure 1].

Decreasing the parameter a of the logistic map to $a1$ = 3.80 and $a2$ = 3.825, we have 79.25 ± 10.31% of correct classification for the simple structure of 2 states and 1 Gaussian/state. Conversely, there is a decrease in classifier performance increasing the chaocity of data to $a1$ = 3.90 and $a2$ = 3.925, as the best performance is 71.00 ± 5.39 % for the structure with 1 state and 6 Gaussians/state, however the difference between Lyapunov exponents has been increased [see Table 1].

The classifiers showed 100% accuracy in classification of different classes of data to logistic map and tent map.

In classification of tent maps, the classifiers were well capable as well. Accuracy of 91.75 ± 7.08% for classification of two datasets of tent map with $r1$ = 0.7, $r2$ = 0.8 for the structure with 9 states and 2 Gaussians/sate, and 97.50 ± 1.94% for $r1$ = 0.9, $r2$ = 0.8 for the same structure, and 100% for $r1$ = 0.9, $r2$ = 0.7 for several structures were reached.

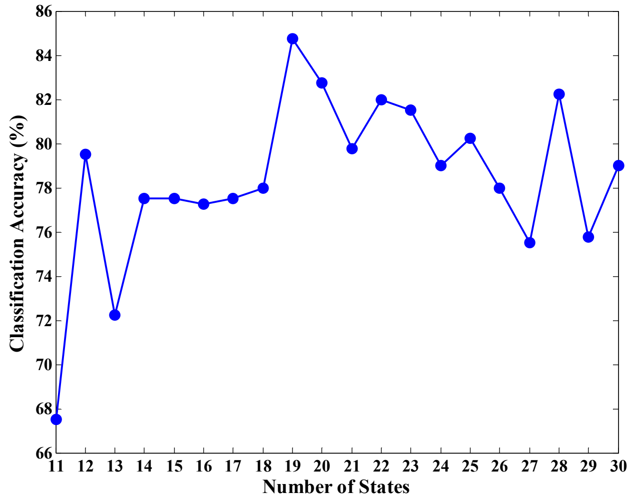

Figure 1 **Classification Accuracy for models of different number of states with 2 Gaussian mixtures as the mean of results of 10 times of implementation of classification of noise-contaminated logistic maps with *a1* = 3.90, and *a2* = 3.875 (cross-validation PCV).**

We had good results for Lorenz model. Correct classification percentage of Lorenz data with *r1* = 90 and *r2* = 80 was at most 84.50 ± 3.72% for 1 state and 4 Gaussians/state mHMM. Increasing the length of signals leads to 99.7 ± 0.9% for signals with length three times before. Increasing the distance of chaocity of two classes of signals, having *r1* = 90, *r2* = 60, the maximum percentage of correct classification was 98.40 ± 1.02% for 1 state and 1 Gaussian/state mHMM; i.e. using just a Gaussian function we can classify these two classes of signals.

2.2 EEG-Based Mental Task Classification

We have used EEG data set provided by Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz (Gert Pfurtscheller et al.) for BCI competition 2003.

EEG signals were acquired during performing mental tasks in a feedback session by a normal subject (female, 25y). The subject sat in a relaxing chair with armrests. The task was to control a feedback bar by means of imagery left or right hand movements according to the cues shown to the subject. The order of left and right cues was random. All data were collected on one day with several minutes break in between runs [19].

There have been several studies on selecting the best EEG-channels for the purpose of classification of mental tasks [20, 21] and it has been shown that for recognition of imagery hand movements, two or three channels of data over the motor cortex are enough to be processed. Mentioned dataset provides data only from 3 channels C3, C4, and Cz. Sampling frequency is 128Hz, and the data is filtered between 0.5 and 30 Hz. There are 140 trials of 9 second length. In each trial the first 2s was quiet; at t = 2s an acoustic stimulus indicates the beginning of the trial, and a cross '+' was displayed for 1s; then at t = 3s an arrow (left or right) was displayed as the cue. At the same time the subject was asked to move a bar into the direction of the cue [19]. Therefore, only the period between t = 4s and t = 9s of each trial was considered for classification study.

We have randomly chosen 100 trials as training- and 40 trials as test-dataset. This process has been repeated 20 times on randomly separated training and test data and at last the results of all these trends have been averaged to provide the total classification accuracy percentage.

In addition to EEG dataset related to movement imagery, mentioned here, we have studied raw EEG signal mental task classification using data recorded during hypnotism sessions in RCISP[1] by Abootalebi et al. In that case, application of mHMM on raw data from all channels has been used for classification of mental tasks, relaxation and imagination, in hypnosis and normal states. Electrodes were placed according to 10-20 standard. EEG signal was sampled at 256Hz and filtered using an elliptic filter with band-width of 0.5-30Hz and then down-sampled to 128Hz. Data of 4 subjects with high hypnotizability have been used. For each subject, a 1-minute segment of each task was selected to be processed. Each 1-minute segment was divided into 12 5-second segments.

2.2.1 dHMM

At first, we have considered classification of signals using 1s windows with no overlap. So HMM's are to model signals with 1s length. In the case the result of classification is assigned to be the average of results of the classifier for different seconds of data, we have 67.63 ± 6.66% accuracy for the classifier with 1 state and 4 observable symbols/state.

If we consider the result of classifier for all seconds of data, we can see that the classification percentage is decreasing. So the best way is to use just the first second for classification. Reasons for the observed decrease in classification results can be: 1. the user lose his concentration on the task, and 2. existed feedback in the BCI system has affected the signals in a way that the probabilistic classifier of HMM cannot be well used for their classification. To increase the classifier performance, the length of window was made half. The best classification percentage showed to be for the classifier with 2 states and 16 observable symbols/state according to second 0.5s segment of data, which is 77.13 %.

2.2.2 mHMM

Data preparation is the same as previous section. Initials models are selected according to k-means, and maximum iteration for performing EM algorithm to train the classifier was set to be 5.

Averaging the results of classification of different seconds, we can see 70.25 ± 9.35% of correct classification for the classifier with 5 states and 10 Gaussians/state. As well, it was shown that for data classification according to just one second of data, the classification percentage decreases using later seconds of data. For classification according to 1st second of data the performance was 76.25%, which is for classifier with 10 states and 6 Gaussians/state.

If windows are selected to be 0.5s then the classification performance is 63.75 ± 7.23% for classifier with 8 states and 4 Gaussians/state. Using first 0.5s segment, there is an increase in classification performance to 77.5% for classifier with 8 states and 2 Gaussians/state.

Studies on RCISP data led to not very good percentage of correct classification for most of subjects in hypnosis state (100, 79.17, 90.91, 71.43% for 4 considered subjects), while results in normal state were satisfactory (91.57, 95.83, 100, 100%, for mentioned 4 considered subjects, respectively).

[1] Research Center for Intelligent Signal Processing, Tehran, Iran

3. DISCUSSION AND CONCLUSIONS

In the case of dHMM, results show that by decreasing the difference between levels of chaocity of different classes, more complex structures are needed to classify them. As well, if the quantities of Lyapunov exponents of two classes are less then classification accuracy is more.

When tent map is considered, dHMM leads to 100% classification rate. It shows that the map producing chaotic signal is more important than the chaosity level of signal and HMM classifier can fit to tent map better than logistic map.

It has been shown that in lots of cases, making models more complex does not lead to better performance of the classifier.

To find the optimum structure, a number of structures should be studied and then, according to the resulted classification percentages, the optimum structure will be specified. There is no optimum structure for signals of same map with different characteristic parameters (different classes) according to our study. But for each two classes of signals we can find the optimum structure according to results shown in the table of classification accuracy for different structures, provided the selected structure is not located on the edge of the table; in such cases, we should continue our study to find a structure with a maximum accuracy.

In the case of Lorenz signals, increasing the length of signal, leads to better performance. This verified the fact that it is needed to consider chaotic signals in long term.

Performance of these classifiers was studied in mental task classification. In our study, we have reached more than 75% of correct classification for imagery movement (77.5% accuracy considering only the first half-second data of each trial) and nearly 100% for RCISP data.

In this study we wanted to show the capability of HMM, a probabilistic classifier, in classification of raw EEG signals, as a class of chaotic signals. It should be noticed that in the case of using features, such as frequency features, the classification accuracy will be much better. As we have studied this case for Graz dataset, the results were up to 90% of correct classification. As well, features such as fractal dimensions and Hjorth parameters [11] have been used with HMM for mental task classification with very good performance.

REFERENCES

[1] J.R. Wolpaw *et al.* Brain-Computer Interface Technology: A Review of the First International Meeting. IEEE Trans. on Neural Systems and Rehabilitation 2000; 8:164-173.

[2] T. Ebrahimi, J.-M. Vesin, G. Garcia. Brain-Computer Interface in Multimedia communication. IEEE Signal Processing Mag. 2003; 20:14-24.

[3] G. Guger, G. Edlinger, W. Harkam, I. Niedermayer, G. Pfurtscheller. How many People are Able to Operate an EEG-Based Brain-Computer Interface (BCI)? IEEE Trans. on Neural Systems and Rehabilitation 2003; 11:145-147.

[4] S. Solhjoo, M.H. Moradi. Mental Task Recognition: A Comparison between Some of Classification Methods, BIOSIGNAL 2004 International EURASIP Conference 2004; 24-26.

[5] J.R. Millan *et al.* A Local Neural Classifier for the Recognition of EEG Patterns Associated to Mental Tasks. IEEE Trans. on Neural Networks 2002; 13:678-686.

[6] J.R. Millan *et al.* Neural Networks for Robust Classification of Mental Tasks, The 22nd Annual International Conference of the IEEE. Engineering in Medicine and Biology Society 2000; 2:1380-1382.

[7] K. Tavakolian, S. Rezaei, A. M.Nasrabadi. Classification of Different Mental Tasks using Neural Network. World Scientific and Engineering Academic and society 2004; 1:2:343-347.

[8] K. Tavakolian *et al.* Effects of Different Feature Vectors and Neural Network Topology on EEG Mental Tasks Classification, IASTED International Conf. in Biomedical Eng. 2003; 39-43.

[9] R. Palaniappan, R. Paramesran, S. Nishida, N. Saiwaki. A New Brain-Computer Interface Design Using Fuzzy ARTMAP. IEEE Trans. on Neural Systems and Rehabilitation 2002; 10:3:140-148.

[10] G. Pfurtscheller, C. Neuper. Motor Imagery and Direct Brain-Computer Communication. Proceedings of the IEEE 1989; 77:2:257-286.

[11] B. Obermaier, C. Guger, C. Neuper, G. Pfurtscheller. Hidden Markov Models for Online Classification of Single Trial EEG Data. Pattern Recognition Letters 2001; 22:12:1299-1309.

[12] B. Obermaier, C. Munteanu, A. Rosa, G. Pfurtscheller. Asymmetric Hemisphere in an Offline Brain-Computer Interface. IEEE Trans. on Systems, Man and Cybernetics 2001; 31:4:536-540.

[13] S. Zhong, J. Ghosh. HMMs and Coupled HMMs for Multi-channel EEG Classification. International Joint Conference on Neural Networks IJCNN'02 2002; 2:1154-1159.

[14] R. Ferri *et al.* Chaotic Behavior of EEG Slow-Wave Activity during Sleep. Electroencephalography and clinical Neurophysiology 1996; 99:539-543.

[15] A. Murata *et al.* Analysis of Chaotic Dynamics in EEG and Its Application to Assessment of Mental Workload. The 20th International Conference of the IEEE Eng. in Med. & Biology 1998; 20:3.

[16] S.N. Sarbadhikari, K. Chakrabarty. Chaos in the brain: a short review alluding to epilepsy, depression, exercise and lateralization. Medical Engineering & Physics 2001. 23:445–455.

[17] H. Korn, P. Faure. Is there chaos in the brain? II. Experimental evidence and related models. C. R. Biologies 2003; 326:787–840.

[18] C. Myers, A. Singer, F. Shin, E. Church. Modeling Chaotic Systems with Hidden Markov Models. IEEE International Conference on Acoustics, Speech, and Signal Processing 1992; 4:565-568.

[19] A. Schlögl. Data set: BCI – experiment. http://ida.first.fraunhofer.de/projects/bci/competition/Graz_description.doc.

[20] K. Tavakolian, A. M.Nasrabadi, M.R. Hashemi Golpayegani, C. Lucas, S.K. Setarehdan, M.A. Khalilzadeh. Using Genetic Algorithm to Reduce Channels and Features in EEG Mental Task Classification. Proceedings of 7th IASTED ASC 2003; 795-800.

[21] K. Tavakolian, A. M.Nasrabadi, S. Rezaei, Selecting Better EEG Channels for Classification of Mental Tasks. IEEE International Symposium on Circuits and Systems 2004; 537-540.