Fractional Delay Time Embedding of EEG Signals into High Dimensional Phase Space

R. Damasevicius[1], I. Martisius[2], V. Jusas[1], D. Birvinskas[2]

[1]*Department of Software Engineering, Kaunas University of Technology*
[2]*Department of Computer Science, Kaunas University of Technology*
Studentu St. 50–415, LT-51368 Kaunas, Lithuania
robertas.damasevicius@ktu.lt

***Abstract**—We propose a method for reconstruction of the high-dimensional phase space of the electroencephalography (EEG) signal. The method is based on the selection of positive trajectories from the phase space, distance-adaptive sampling of the negative trajectories from the phase space, classification of trajectories in the phase space, and reconstruction of a fuzzy state of the signal for classification of EEG signals.*

***Index Terms**—Adaptive classification, delay time embedding, phase space, landmark signal, EEG, BCI.*

I. INTRODUCTION

Brain-computer interface (BCI) is a communication system that translates brain activity into commands for a computer or other digital device [1]. Most BCI systems work by reading and interpreting cortically-evoked electro-potentials via an electroencephalogram (EEG) data. The frequencies of these brain waves range from 0.5 Hz to 100 Hz, and their characteristics change dynamically depending on the activity of the human brain [2].

BCI systems require correct classification of the EEG signals for useful operation. Since the EEG signal may be considered as chaotic [3], [4], the nonlinear dynamics and chaos theory methods can be applied for analysis and classification of the EEG data. Nonlinearity of the EEG and other types of biosignals has been studied by applying nonlinear analysis methods such as Detrended Fluctuation Analysis (DFA) [5], Poincaré Plots [6], reconstruction of Local Phase Space, segmentation of the EEG signal into stationary fragments [7], application of non-linear operators to the EEG time series [8].

Nonlinear analysis usually depends on the reconstruction of the phase-space geometry of the signal from a small number of observables. The first step is to embed data in a higher dimensional phase space. The embedding method is based on two theorems.

First, Whitney Embedding Theorem [9] implies that 2M + 1 independent signals measured from a system can be considered as a map from the set of states to the 2M + 1 dimensional space, thereby reconstructing the phase space.

Next, Takens Embedding Theorem [10] proved that instead of 2M + 1 generic signals, the time-delayed versions of a generic signal would suffice to embed the M-dimensional manifold.

This paper proposes a method for reconstruction of the high-dimensional phase space of the EEG signal. The novelty of the method, first, is the concept of fractional embedding, which extends linear embedding for fractional time delays. Second novelty is the reconstruction of the fuzzy state vector of the signal which is useful for in-depth signal analysis and identification of signal landmarks [11].

II. PROPOSED METHOD

A. Preliminaries

The aim of binary classification is classification of the members of a given set of objects into two groups on the basis of whether they have some property or not.

Let us have two datasets of signal values: a positive dataset $\{X_1^P, X_2^P, ..., X_p^P\}$ and a negative dataset $\{X_1^N, X_2^N, ..., X_n^N\}$, where $X_i = \{x(t), t \in T\}$ is a time series (signal), representing some real physical process $X \in \Re$ sampled at time t, T is the index space, and p and n are the size of the positive and negative dataset, respectively.

Let us have a label function $\}: X \rightarrow [0,1]$, which indicates to which dataset a time series belongs to, i.e., $\forall X_i^P, \}(X_i^P) = 1$ and $\forall X_i^N, \}(X_i^N) = 0$.

The aim is to perform classification of unknown label time series X_u sampled from X, i.e., to find $\}(X_u)$.

B. Method

The proposed method consists of the following stages:
1. Reconstruction of the high-dimensional phase space of the signal.
2. Selection of positive trajectories from the phase space.
3. Distance-adaptive sampling of the negative trajectories from the phase space.
4. Classification of trajectories in the phase space.
5. Reconstruction of a fuzzy state of the signal.

The stages are explained in detail in the following sub-sections.

C. Reconstruction of Phase Space

The *M*-dimensional phase space $\Re^M$ is reconstructed from a single observable $x(t)$, using the M-dimensional vectors (trajectories) $Y = [x(t), x(t + \ddagger), ... x(t + (M - 1)\ddagger)]$,

where $‡$ is the delay (lag), and M is the dimensionality of the trajectory. We define the reconstruction of the phase space as a function $t_‡ : X \rightarrow Y$, where $‡$ is the delay parameter.

The classical method of delays allows only for $‡ \in \aleph$. Our novelty is the modification of the classical delay-based method to allow for fractional delay $‡ \in \Re$. The value of $x(t+‡)$ for a fractional delay $‡$ is calculated using the linear interpolation as in (1)

$$x(t+‡) = x(\lfloor t+‡ \rfloor) + \frac{x(\lceil t+‡ \rceil) - x(\lfloor t+‡ \rfloor))}{\lceil ‡ \rceil - \lfloor ‡ \rfloor}(‡ - \lfloor ‡ \rfloor). \quad (1)$$

This step allows for dimensionality reduction of original datasets, because usually $M \ll T$. The result of this step is two datasets of signal trajectories: a positive dataset $\{Y_1^P, Y_2^P, \ldots, Y_{p \cdot M}^P\}$ and a negative dataset $\{Y_1^N, Y_2^N, \ldots, Y_{n \cdot M}^N\}$.

D. Selection of Positive Trajectories

We assume that a positive dataset Y^P is a mixed set of trajectories, which consists of a subset of trajectories that are similar to Y^N as well as of a subset of trajectories that are different from Y^N.

The aim is to find a subset of positive trajectories $\hat{Y}^P$ such that the distance between $\hat{Y}_i^P$ and Y_j^N is the largest. We also assume that for any dataset X_i there exists at least one such positive trajectory such as described in (2)

$$\forall \hat{Y}_i^P, \max_j D(\hat{Y}_i^P, Y_j^N), \quad (2)$$

where $\hat{Y}_i^P \in t(X_k^P)$, $Y_j^N \in t(X_k^N)$, $D : Y \times Y \rightarrow \Re$ is a dissimilarity (distance) function between two trajectories.

Here we use the Euclidean distance, described in (3), as the dissimilarity function. We specify the selection step as a selection operator $\dagger : Y \rightarrow \hat{Y}$

$$d_{i,j} = \sqrt{\sum_{k=1}^{M} (y_{ik} - y_{jk})^2}. \quad (3)$$

E. Distance-Adaptive Sampling of Negative Trajectories

The result of the selection step is a positive trajectory dataset $\{\hat{Y}_1^P, \hat{Y}_2^P, \ldots, \hat{Y}_M^P\}$, which is significantly smaller than a negative trajectory dataset $\{Y_1^N, Y_2^N, \ldots, Y_{n \cdot M}^N\}$. Using such imbalanced datasets for classification is impractical, because the classification results become biased [12].

To balance the dataset, we perform the distance-adaptive sampling of a negative dataset Y^N as described in (4)

$$h_\lor(Y_i^N) = |Y_j^N|, \quad (4)$$

where $D(Y_i^N, Y_j^N) < \lor$, $i \neq j, Y_i^N$, $Y_j^N \in t_‡(X_k^N)$, the - neighborhood count function of trajectories. Let

$p_s(Y_i^N) = 1 - \dfrac{1}{1 + h_\lor(Y_i^N)}$ be the sampling probability of Y_i^N.

The algorithm of the distance-adaptive sampling is as follows:

1. Calculate a dissimilarity matrix D, where $D(i, j)$ is a distance (dissimilarity) between Y_i and Y_j.

2. Calculate a neighborhood array E, where $E(i)$ is the number of neighbors of Y_i which are at most at $\lor$ distance from Y_j.

3. Calculate the probability array P, where $P(i)$ is the sampling probability of Y_i.

4. Construct a sampled dataset $\hat{Y}^N$, where $Y_i^N \in \hat{Y}^N$ if $p_s(Y_i^N) > p$, and p is a Gaussian distributed random number from $(0,1)$.

F. Classification

Finally, the trajectory vectors are used as features for classification.

For training datasets, we define a label function $\hat{\}} : Y \rightarrow [0,1]$, and assign $\hat{\}}(\hat{Y}_i^P) = 1$ and $\hat{\}}(\hat{Y}_i^N) = 0$. The result of training is a support vector model which is used for classification of unknown label time series X_u as follows:

1. Reconstruct a set of trajectories in the phase state $Y^u = t_‡(X_u)$.

2. Perform classification of a set of trajectories Y^u and obtain a set of trajectory labels $L^u = \hat{\}}(Y_i^u), Y_i^u \in Y^u$.

G. Reconstruction of a Fuzzy State of the Signal

The classification result of each trajectory is used to reconstruct the state vector $S(t)$ of the EEG signal as described in (5)

$$S(t) = \frac{1}{M} \sum_{k=t}^{t+M} L_k^u. \quad (5)$$

The results are interpreted in terms of fuzzy logic as a fuzzy probability of the state (positive or negative) of the signal at time t.

Finally, a classification label is assigned to X_u as described in (6):

$$\}(X_u) = \begin{cases} 1, \max_t S(t) \geq \upsilon, \\ 0, \max_t S(t) < \upsilon, \end{cases} \quad (6)$$

where υ is the threshold probability of the state.

III. METHODS

For our experiments, Datasets Ia, Ib (‹self-regulation of SCPs›) and dataset IV (‹self-paced 1s›) from the BBCI competition datasets (http://bbci.de/competition/) were used. The datasets were taken from healthy subjects. In the first two datasets the subject was asked to move a cursor up and down on a computer screen, and received visual feedback,

while his slow cortical potentials (SCPs) were recorded.

Dataset Ia [13] consists of 135 trials belonging to class 0 and 133 trials belonging to class 1.

Dataset Ib [13] consists of 100 trials belonging to class 0 and 100 trials belonging to class 1. Each trial consists of 896 samples from each of 6 channels. The sampling rate is 256 Hz, the recording length is 3.5 s.

In dataset IV [14], the goal is to predict the laterality of upcoming finger movements (left vs. right hand) 130 ms before key press. The sampling rate is 1 kHz. Down-sampling to 256 Hz was performed before the experiments. The set consists of 159 trials belonging to class 1 and 157 trials belonging to class 0.

All datasets were randomly partitioned into 5 parts, and 5-fold cross-validation was used to evaluate the classification results.

IV. RESULTS

Figure 1 shows some examples of the EEG signals taken from Dataset Ib, channel 0.
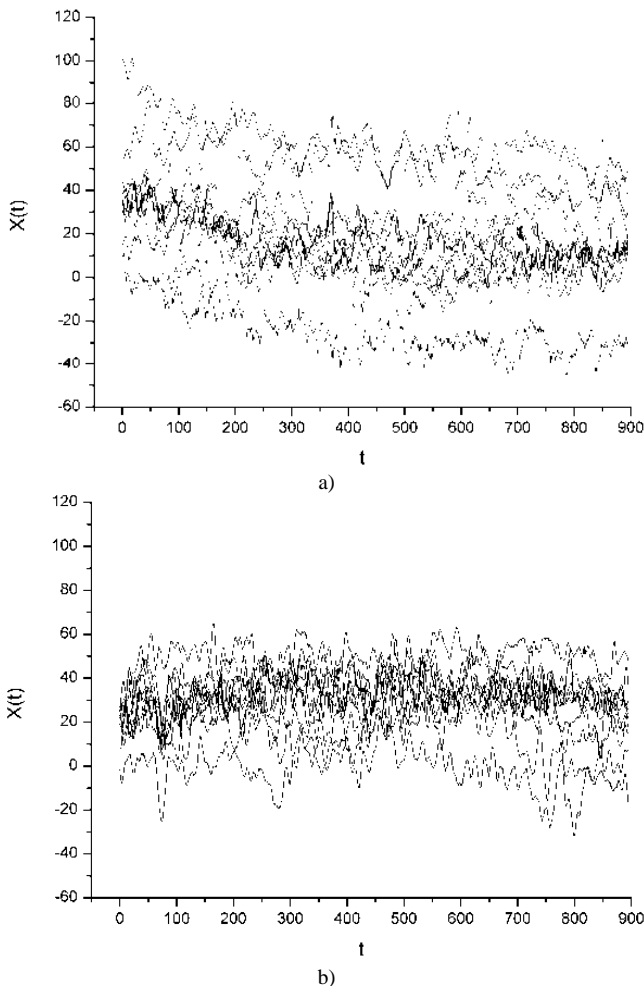

a)

b)

Fig. 1. EEG signal (Dataset Ib, channel 0): a) 20 positive signals, b) 20 negative signals.

The results of the fractional time delay embedding of the EEG signal into the high dimensional phase space is shown in Fig. 2 (dataset Ib, channel 0). The plot shows that positive dataset has a wider distribution of values than a negative dataset (compare Fig. 2(a) with Fig. 2(b)). Such extreme values (called *landmarks* [11]) can be exploited for signal classification.
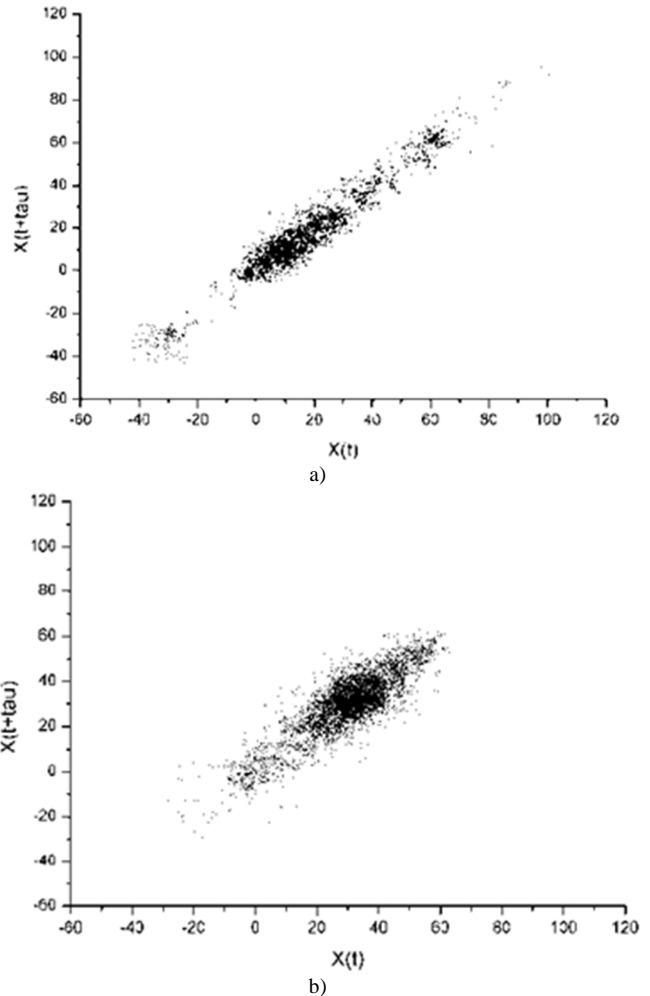

a)

b)

Fig. 2. Phase space reconstruction of EEG signals (dimension M = 2, time delay = 9.5): a) 20 positive signals, b) 20 negative signals.

V. RELATED WORK

The idea of embedding has been used previously for analysis of EEG data.

Athitsos *et al.* [15] proposed a method for approximate subsequence matching by using embedding. Embedding maps each database time series into a sequence of vectors, so that every step of every time series is mapped to a vector. Embedding is computed by applying full dynamic time warping between reference objects and each time series.

Klonowski *et al.* [3] have estimated the embedding dimension of the EEG signals of the same subjects to be between 7 and 11.

Rohrbacker [16] has found that the EEG data has been unfolded in 22-dimensional space. Anderson *et al.* [17] have used a time-embedding dimension of 50 for a neural network classifier.

Kroupi *et al.* [18] investigated how brain signals represented by the EEG data interacted when watching music clips. A non-linear measure was applied to determine the driver/driven relationship between these two modalities. The range of the embedding dimensions used for the EEG segments was between 3 to 8.

Jianbo and Jing [19] used the recurrence time based method for seizure detection. Their empirical evidence shows that embedding dimensions from 3 to 6 are fine, while time delays may be chosen from 2 to 6.

Yuan *et al.* [20] analysed the embedding dimension of

normal and epileptic EEG time series. The study has found that the embedding dimension reflects variation of freedom degree of human brain nonlinear dynamic system during seizure. The embedding dimension changes during seizure and becomes different from that of normal EEG signals: the average value of the embedding dimension of normal EEG signals is 8, while the embedding dimension of the epileptic EEG signals is 17.

VI. CONCLUSIONS

The advantages of the proposed method are: 1) the reconstruction of the fuzzy state vector of the signal allows for in-depth analysis of the signal and identification of signal landmarks; and 2) fractional time embedding allows for more flexible embedding.

The disadvantages of the proposed method are 1) the determination of the delay time τ and the dimension M, which allows to achieve best separation of trajectories from positive and negative datasets, is a computationally hard task if done by extensive search; 2) identification of signal landmarks can be time-consuming if performed manually by an expert; and 3) the method is susceptible to noise (this is the main problem of all embedding-based methods), so the accuracy of the method is hard to guarantee.

Further research will focus on proposing the computationally feasible solution to finding optimal delay time and embedding dimension values to achieve best separation of positive and negative signal trajectories considering time constraints of real-time BCI systems.

REFERENCES

[1] U. Hoffmann, J. M. Vesin, T. Ebrahimi, "Recent advances in brain-computer interfaces", *IEEE 9th Workshop on Multimedia Signal Processing, (MMSP 2007)*, pp. 17, 2007.

[2] J. Zhang, Y. Wang, R. Wang, "Application of KIII model to EEG classification based on nonlinear dynamic methods", *Int. Journal of Artificial Intelligence*, vol. 7, no. A11, 2011.

[3] W. Klonowski, W. Jernajczyk, K. Niedzielska, A. Rydz, R. Stepien, "Quantitative measure of complexity of EEG signal dynamics", *Acta Neurobiol Exp (Wars)*, vol. 59, no. 4, pp. 315–21, 1999.

[4] J. Ulbikas, A. Cenys, O. P. Sulimova, "Chaos parameters for EEG analysis", *Nonlinear Analysis: Modelling and Control*, vol. 3, pp. 1–8, 1998.

[5] R. Krishnam, H. Nazeran, S. Chatlapalli, E. Haltiwanger, Y. Pamula, S. Cabrera, "Detrended fluctuation analysis: a suitable long-term measure of heart rate variability accurate derivation of heart rate variability signals in children with sleep disordered breathing", in *Proc. 27th Annual Int. Conf. of the IEEE Eng Med Biol Soc.*, 2005, pp. 1174–1177.

[6] S. Chatlapalli, H. Nazeran, V. Melarkod, R. Krishnam, E. Estrada, Y. Pamula, S. Cabrera, "Accurate derivation of heart rate variability signal for detection of sleep disordered breathing in children", in *Proc. 26th Ann. Int. Conf. IEEE Eng Med Biol Soc.* 2004, pp. 538–541.

[7] R. Ursulean, A. M. Lazar, "Segmentation of electroencephalographic signals using an optimal orthogonal linear prediction algorithm", *Elektronika ir Elektrotechnika*, no. 1, pp. 73–76, 2007.

[8] I. Martisius, R. Damasevicius, V. Jusas, D. Birvinskas, "Using higher order nonlinear operators for SVM classification of EEG data", *Electronics and Electrical Engineering*, no. 3, pp. 99–102, 2012.

[9] H. Whitney, "Differentiable manifolds", *Ann. Math.*, vol. 37, pp. 645–680, 1936. [Online]. Available: http://dx.doi.org/10.2307/1968482

[10] F. Takens, "Detecting strange attractors in turbulence", *Lecture notes in mathematics,* vol. 898, pp. 366–381, 1981.

[11] C.-S. Perng, H. Wang, S. R. Zhang, D. S. Parker, "Landmarks: A new model for similarity-based pattern querying in time series databases", *Proc. 16th Int. Conf. Data Engineering (ICDE 2000)*, pp. 33–42, 2000. [Online]. Available: http://dx.doi.org/10.1109/ICDE.2000.839385

[12] Z. Yang, Y. Yang, W. Gang, "Classification for imbalanced dataset based on biased empirical feature mapping". *IEEE Int. Instrumentation and Measurement Technology Conf. (I2MTC)*, 2012, pp. 1645-1649.

[13] N. Birbaumer, H. Flor, N. Ghanayim, T. Hinterberger, I. Iverson, E. Taub, B. Kotchoubey, A. Kubler, J. Perelmouter, "A brain-controlled spelling device for the completely paralyzed", *Nature*, vol. 398, pp. 297–298, 1999. [Online]. Available: http://dx.doi.org/10.1038/18581

[14] J. Kubanek, K. J. Miller, J. G. Ojemann, J. R. Wolpaw, G. Schalk, "Decoding flexion of individual fingers using electrocorticographic signals in humans", *J. Neural Eng.*, vol. 6, 2009.

[15] V. Athitsos, P. Papapetrou, M. Potamias, G. Kollios, D. Gunopulos, "Approximate embedding-based subsequence matching of time series", in *Proc. Int. Conf. on Management of data (SIGMOD 2008)*, ACM, NY, USA, 2008, pp. 365–378.

[16] N. Rohrbacker, "Analysis of electroencephalogram data using time-delay embeddings to reconstruct phase space", *Dynamics at the Horsetooth*, vol. 1, 2009.

[17] C. Anderson, E. Forney, D. Hains, A. Natarajan, "Reliable identification of mental tasks using time-embedded EEG and sequential evidence accumulation", *J. Neural Eng.* vol. 8, 2011.

[18] E. Kroupi, J.-M. Vesin, T. Ebrahimi, "Driver-response relationships between frontal EEG and Respiration during affective audiovisual stimuli", in *Proc. 35th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 2911–2914.

[19] G. Jianbo, H. Jing, "Fast monitoring of epileptic seizures using recurrence time statistics of electroencephalography", *Frontiers in Computational Neuroscience*, vol. 7, 2013.

[20] Y. Yuan, Y. Li, D. P. Mandic, "A comparison analysis of embedding dimensions between normal and epileptic EEG time series", *J Physiol Sci.*, vol. 58, no. 4, pp. 239–247, 2008. [Online]. Available: http://dx.doi.org/10.2170/physiolsci.RP004708