A TREMOR MODEL SELECTION BASED ON ARTIFICIAL NEURAL NETWORKS CORRECT CLASSIFICATION RATE

M.C. Serban, D.M. Dobrea

“Gh. Asachi” Technical University, Iasi, Romania

serbanm@etc.tuiasi.ro, mdobrea@etc.tuiasi.ro

Abstract: In this paper, an autoregressive (AR) model and, respectively, an adaptive nonlinear Markov process amplitude algorithm were used to model and simulate the physiologic tremor signal obtained during a visual repetitive stimulation. In order to select the best between the two proposed models – linear, and, respectively, nonlinear stochastic process – the visual reflex detected in the tremor signal was used. Thus, knowing that a significant response is obtained in the tremor signal at double the stimuli frequency, the tremor recorded for three different stimuli frequencies (5 Hz, 10 Hz and 15 Hz) were classified with a multilayer perceptron (MLP) and a support vector machine (SVM) artificial neural network. The selection of the most appropriate model was done based on the correct classification rate of the classifiers. Hence, the best rate gave the model that better fitted the original tremor data.

Introduction

So far, physiological tremor (PT) was subject to numerous studies aimed to find its real mechanisms and origins. But, even now, we still know very little about it. Physiological tremor – defined as the involuntary, oscillatory movement of parts of the body, mainly the upper limbs –, is a complex signal resulting from interactions between several mechanical and neural factors. Up till now some models were proposed for the complex system generating the tremor. The tremor signal’s modeling is of great importance since a well fitting model might yield insight into the underlying process. Thus, a better understanding of the physical mechanisms generating the tremor signals and a superior comprehension of the causes of the tremor signals changes could be achieved.

The *linear stochastic autoregressive* (*AR*) *process* was one of the first proposed models [1] for the tremor data. An improved model was then suggested in [2]. This was the *linear state space model* (*LSSM*) and, unlike the AR model, it takes into account the observational noise too. Compared to the AR model, the LSSM model proved to be a better one [1] while the *autoregressive moving average* (*ARMA*) *processes*, which are generalized AR processes including past driving noise terms in the dynamics, yielded only to comparable results [4]. All the above models are linear stochastic processes and their implementation results led to the conclusion that the PT should mainly be considered a realization of a linear stochastic process [3]. In terms of physics and physiology this means that the PT could be regarded as the output of a linear damped oscillator driven by uncorrelated muscle activity.

Nevertheless, there are also papers pointing out that tremor may include significant nonlinear [3],[5], even chaotic components [6], which may play a major role in diagnosis and rehabilitation. If in [6] the chaotic behaviour was demonstrated in the tremor of healthy persons, in [3],[5] the nonlinear behaviour is considered to be the characteristic of only the pathological forms of tremor (i.e. Parkinsonian tremor, essential tremor etc.). In other words, pathological tremors represent non-linear processes. If these processes are stochastic or determinist chaotic this is still an opened question. Hence, if in [3][5][7] the processes are proved to be stochastic in nature, in [8] a chaotic approach is proposed in order to differentiate between two types of pathologic tremors.

However, beside all these hypotheses there is growing evidence that complex bodily rhythms arise from stochastic, nonlinear biological mechanisms continuously interacting with a constantly fluctuating environment. Thus, many physiological processes are neither linear stochastic processes nor nonlinear deterministic ones.

In Figure 1 it can be seen that the two paradigms, nonlinear-deterministic and linear-stochastic behaviour, are extreme positions in the area spanned by the properties “nonlinearity” and “stochasticity”. Some models that allow a connection to be made between the nonlinear stochastic model approach and particular real world phenomenon fill the gap between them. One of these models is equally the nonlinear Markov process amplitude (NMPA) model.

In this paper the linear as well as the nonlinear stochastic processes are reviewed as possible model candidates for the hand physiological tremor. The

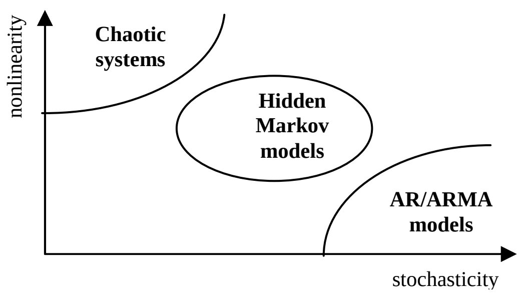

Figure 1: Systems’ areas covered by analysis methods

analysis capitalizes on the results reported in [9]. Here, the observed changes in the frequency characteristics of the tremor's signal – due to visual repetitive stimuli (5 Hz, 10 Hz and 15 Hz) –, demonstrated a significant connection between visual pathways in the central nervous system (CNS) and the regions basically governing tremor. Knowing that a considerable coherence was revealed in the tremor signals at double the stimuli frequency, we further took advantage on the central driving oscillations confined in the PT signal. Thus, the tremor was modeled using first an AR model and then, nonlinear Markov process amplitude (NMPA) model [10]. Finally, the better fitting model was chosen using, as a criterion, the correct classification rate of two artificial neural networks that had as inputs the estimated parameters of the two models.

Materials and Methods

Two subjects were admitted in this study. The subjects were healthy, with no known neurological or endocrine pathology. The subjects have been explained all procedures and gave written consent regarding the participation in the study. The entire procedure of the tremor acquisition was unobtrusive for the subjects, with no physical contact, due to the sensor capability [11].

The subject's elbow was fixed by a mechanical support in order to preserve the tremor characteristic unaffected by the hand fatigue influence in the last part of the recordings. In order to isolate them from all kind of surroundings stray stimuli, all the recordings took place in a quiet room without any source of light. All the time the subjects looked to a computer display. The stimuli consisted in a circle, of 2 cm radius, placed in the middle of the display changing its luminosity between a black background followed by a white flash. The pattern of the stimuli changes was a symmetric rectangular wave with the desired selectable frequency (5Hz, 10Hz, 15Hz). The subjects had no visual control of their hand position. Each recording last 98.4 s, but only the first 32.8 s and the last 32.8 s of the tremor signal were kept. After the first time segment of 32.8s a visual stimuli at one of the specified frequencies (5Hz, 10Hz, 15Hz) was presented to the subject. The recording sessions were scheduled several days until the acquiring of the entire data set was finished (77 recordings: 30/5Hz, 29/10Hz, 18/15Hz). The sampling rate was 250 samples per second and we got 8.200 samples per each acquired segments of the recording.

Each time series was pre-filtered using a 4 ÷ 40 Hz linear band-pass filter. Further, it was normalized to unity variance and zero mean value and it was divided into 119 sliding windows of 128 samples each, overlapped by 64 samples.

A first implemented model for the tremor data generating system was an autoregressive (AR) process of 4th order. To apply the AR framework to tremor signals, we assumed that a linear filter described the process of tremor generation and that this filter was fed with a white noise signal, w[13]:

$$x[n] = -\sum_{i=1}^{N} a_i \cdot x[n-i] + w[n] \qquad (1)$$

In eq. (1) the current observation, $x[n]$, is represented as a weighted linear combination of past observations, $x[n-i]$ (i=1,…,N), plus a random, uncorrelated input. In our case N was 4 and the AR coefficients, a_i (i=1,…,N), were calculated by solving the Yule–Walker equations given in eq. (2).

$$R_x \cdot \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (2)$$

Here, σ^2_w is the variance of the white noise signal, $w[n]$, and R_x, given by:

$$R_x = \begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-N] \\ R_x[1] & R_x[0] & \cdots & R_x[-N+1] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[N] & R_x[N-1] & \cdots & R_x[0] \end{bmatrix} \qquad (3)$$

represents the correlation matrix for a random vector whose N components are the outputs of the random process $\boldsymbol{x}$ ($\boldsymbol{x}$ = [$x[0]$, $x[1]$, …, $x[N-1]$]).

Further, a second model was implemented for the original tremor data. The tremor modeling using *adaptive nonlinear Markov process amplitude* (*ANMPA*) reported in the present article is an implementation of the nonlinear Markov process amplitude model (NMPA) proposed in [10] for nonlinear coupling interaction of spontaneous EEG. The model parameters were determined adaptively with the least mean square (LMS) algorithm. A version of this algorithm for a first-order Markov process amplitude model is presented in [12].

In this paper, the NMPA model was assumed to appropriately decompose the frequency components of the tremor signal into some spontaneous oscillations (a priory specified) and the nonlinearly coupled frequencies (self-coupling oscillations and, respectively, cross-coupling oscillations). More exactly, two oscillatory waves (m_1 and m_2) passing through a nonlinear square system generates two kinds of harmonic frequencies: self-coupling harmonics ($2m_1$ and $2m_2$) and, respectively, cross-coupling harmonics ($m_1 \pm m_2$). Having these we can write the NMPA model as in eq. (4). Here, $y(n)$ is the estimated tremor signal assumed to be composed of K different oscillations (x_j, j=1÷K), m_j is the dominant jth frequency, ϕ_j is the initial phase (in this paper, it is equal to zero), ε^s_j is the self-coupling coefficient of the jth model oscillation, ε^{c1}_{ij} and ε^{c2}_{ij} are the cross-coupling coefficients of the coupled frequency m_i-m_j and m_i+m_j, respectively, n is the time index and $a_j(n)$ is the model amplitude of the first order Markov process.

$$\begin{cases} y(n)=\sum_{j=1}^{K} a_j(n)x_j(n)+\sum_{j=1}^{K}\varepsilon_j^s a_j(n)\alpha_j(n)+ \\ \qquad \sum_{\substack{i,j=1,K \\ i\neq j}}[\varepsilon_{ij}^{c1}a_i(n)a_j(n)\beta_{ij}(n)+\varepsilon_{ij}^{c2}a_i(n)a_j(n)\theta_{ij}(n)] \\ x_j=\sin(2\pi m_j n+\phi_j) \\ \alpha_j=\sin(2\pi 2m_j n+2\phi_j+\pi/2) \\ \beta_{ij}=\cos\left[2\pi(m_i-m_j)n+\phi_i-\phi_j\right] \\ \theta_{ij}=\cos\left[2\pi(m_i+m_j)n+\phi_i+\phi_j\right] \end{cases} \tag{4}$$

The next estimate of the model amplitude $a_j(n+1)$ is given by eq. (5):

$$\begin{cases} a_j(n+1)=\gamma_j(n)a_j(n)+\mu_j(n)\xi_j(n) \\ 0<\gamma_j(n)<1 \end{cases} \tag{5}$$

where: $\xi_j(n)$ is the independent increment of Gaussian distribution with zero mean and unity variance, μ_j is the coefficient of the random process and γ_j is the coefficient of the first-order Markov process.

As tremor signal is highly nonstationary the least mean square algorithm was used in order to adaptively estimate some of the model parameters (see Figure 2). The error squared, $e(n)^2 = [s(n)-y(n)]$ (where $s(n)$ is the tremor signal to be modeled), was used as an estimate of the mean square error cost function J, defined as $J = 1/2\cdot E\{e(n)^2\}$.

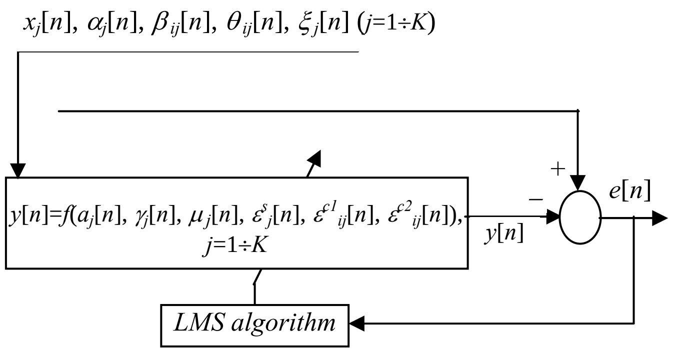

Figure 2: Block diagram of the LMS nonlinear AMPA

The model's parameters (a_j, γ_j, μ_j, ε^s_j, ε^{c1}_{ij}, ε^{c2}_{ij}, i,j=1÷K, $i\neq j$) were adaptively adjusted using for this a steepest descent type algorithm. For each iteration and for each of the above parameters the following formula was applied:

$$\tau_j(n+1)=\tau_j(n)-\eta_\tau \hat{\nabla}_{\tau_j} J(n) \tag{6}$$

Here, τ is a symbol denoting each of the above parameters of the model, η_τ is a small positive constant called the adaptive learning rate and $\nabla_{\tau j}\, J(n)$ is the gradient approximation of $J(n)$ that is defined as:

$$\hat{\nabla}_{\tau_j} J(n)=\begin{bmatrix} \dfrac{\delta e^2(n)}{\delta\tau_i(n-1)} \\ \vdots \\ \dfrac{\delta e^2(n)}{\delta\tau_K(n-1)} \end{bmatrix} \tag{7}$$

Applying the eq. (6) and (7) one can obtain the following adjusting formulas for the NAMPA model parameters (see eq. (8), (9), (10), (11) and (12)):

$$\begin{aligned} \gamma_j(n+1)=\gamma_j(n)+\eta_\gamma e(n)\{&a_j(n-1)x_j(n)+ \\ &\varepsilon_j^s(n)a_j(n-1)\alpha_j(n)+ \\ &\sum_{\substack{i=1\\ i\neq j}}^{K}[\gamma_i(n-1)a_i(n-1)+ \\ &\mu_i(n-1)\xi_i(n-1)][\varepsilon_{ij}^{c1}(n)\beta_{ij}(n)+ \\ &\varepsilon_{ij}^{c2}(n)\theta_{ij}(n)]a_j(n-1)\} \end{aligned} \tag{8}$$

$$\begin{aligned} \mu_j(n+1)=\mu_j(n)+\eta_\mu e(n)\{&\xi_j(n-1) \\ &x_j(n)+\varepsilon_j^s(n)\xi_j(n-1)\alpha_j(n)+ \\ &\sum_{\substack{i=1\\ i\neq j}}^{K}[\gamma_i(n-1)a_i(n-1)+\mu_i(n-1) \\ &\xi_i(n-1)][\varepsilon_{ij}^{c1}(n)\beta_{ij}(n)+\varepsilon_{ij}^{c2}(n) \\ &\theta_{ij}(n)]\xi_j(n-1)\} \end{aligned} \tag{9}$$

$$\begin{aligned} \varepsilon_j^s(n+1)=\varepsilon_j^s(n)+\eta_{\varepsilon^s}e(n)[&\gamma_j(n-1) \\ &a_j(n-1)+\mu_j(n-1)\xi_j(n-1)]\alpha_j(n) \end{aligned} \tag{10}$$

$$\begin{aligned} \varepsilon_{ij}^{c1}(n+1)=\varepsilon_{ij}^{c1}(n)+\eta_{\varepsilon^{c1}}e(n)[&\gamma_i(n-1) \\ &a_i(n-1)+\mu_i(n-1)\xi_i(n-1)] \\ &[\gamma_j(n-1)a_j(n-1)+\mu_j(n-1) \\ &\xi_j(n-1)]\beta_{ij}(n) \end{aligned} \tag{11}$$

$$\begin{aligned} \varepsilon_{ij}^{c2}(n+1)=\varepsilon_{ij}^{c2}(n)+\eta_{\varepsilon^{c2}}e(n)[&\gamma_i(n-1) \\ &a_i(n-1)+\mu_i(n-1)\xi_i(n-1)] \\ &[\gamma_j(n-1)a_j(n-1)+\mu_j(n-1) \\ &\xi_j(n-1)]\theta_{ij}(n) \end{aligned} \tag{12}$$

In order to select the better between the two proposed models – linear, and, respectively, nonlinear stochastic process – the tremor recorded for three different stimuli frequencies (5 Hz, 10 Hz and 15 Hz) were classified with a multilayer perceptron (MLP) and, respectively, a support vector machine (SVM) artificial neural network. The selection of the most appropriate model was done based on the correct classification rate

ISSN: 1727-1983 © 2005 IFMBE

of the classifiers. Thus, the best rate gave the model that better fitted the original tremor data.

Multilayer perceptrons (MLPs) [13] are distributed, adaptive, generally nonlinear learning machines consisting in many different interconnected processing elements (PEs). Moreover, they are feedforward, supervised, neural networks trained with the standard backpropagation algorithm [13]. Frequently used for pattern classification problems, network topologies with one or two hidden layers can approximate virtually any input-output map.

Learning in neural networks is done by propagating the information back through the network and by changing the weighting factors (weights) at each processing element (PE) to reduce output errors. Thus, the weights are adjusted directly from the training data without any assumptions about the statistical distribution of the data. In order to prevent the network to memorize the training-set data a cross-validation set is frequently used as a stop criterion in the training process. Thus, for a given network size the generalization is maximized, fact that allows the incomplete or noisy inputs to be completely recovered by the network. Very important in the MLPs (that are semiparametric classifiers) are the network topology and the activation functions assigned to the PEs. These two central issues in neural network design have great impact in the classification performance of the network.

Support vector machines are state of the art learning machines based on statistical learning theory. This type of artificial neural networks map the input training data set onto a high dimensional state using for this some function ϕ. If this feature space is sufficiently large, then the patterns become linearly separable and a simple perceptron network can do the classification in this new space. A key point of the theory consists in the dot product $<\phi(x), \phi(x_i)>$ in the feature space that is replaced with a nonlinear function $k(x, x_i)$ called kernel. The kernel Adatron algorithm [14] represents the "on-line" version of the quadratic optimization approach used for SVMs. By determining, in the feature space, a separating hyperplane that maximizes the *margin* or distance between the hyperplane and the closest data point (support vector) belonging to different classes, the SVMs minimize an upper bound to the Bayes error. This means that a better generalization is obtained for SVMs, making them very useful in pattern recognition.

Results

In this paper the frequency components of the tremor signals were decomposed into five spontaneous oscillations (5 Hz, 10 Hz, 15 Hz, 20 Hz, 30 Hz) and their corresponding self- and cross-coupling oscillations. The five oscillations were selected in order to detect and model – for each particular stimuli frequency (5 Hz, 10 Hz and, respectively, 15 Hz) – the effect of the visual stimulation upon the tremor signal characteristics.

The AR model of 4th order was implemented in Matlab, using the Yule-Walker equations given by eq. (2).

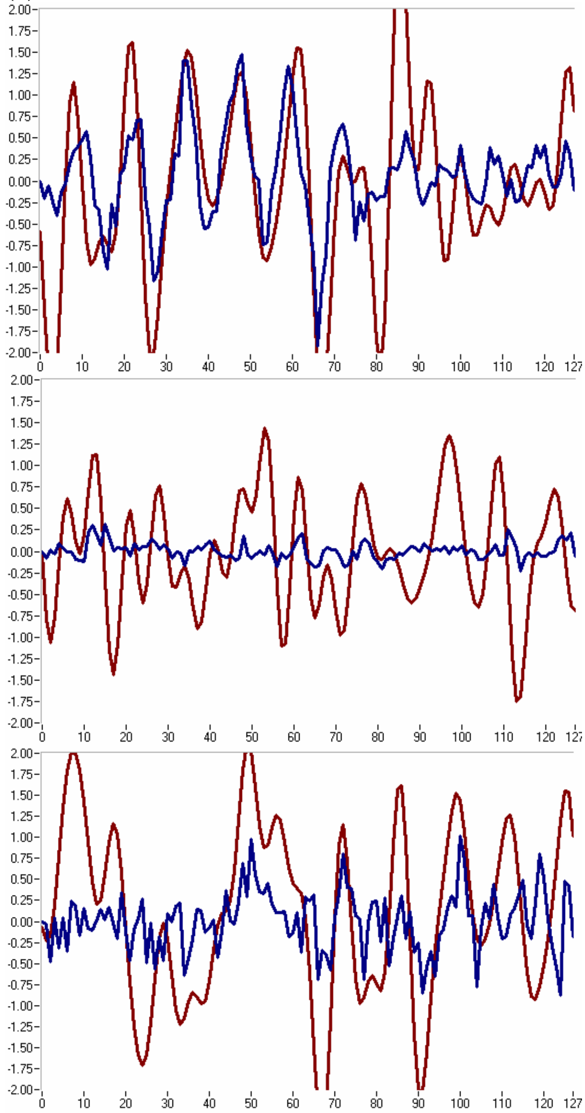

Figure 3: Original (red) and modeled (blue) tremor data (some examples)

The 4 AR coefficients of the signal sliding windows formed the input feature vectors corresponding to the autoregressive model. The adaptive nonlinear Markov process amplitude model was implemented in LabWindows CVI 5.5. The adaptive learning rates (η_γ, η_μ, η_{ε_s}, η_{ε_c1}, η_{ε_c2}) were set to fixed values throughout training. The learning was stopped after 500 epochs of training. The feature vectors were chosen among different combinations, beginning with 5 components given by the proposed five spontaneous oscillations and ending with feature vectors of 50 components (spontaneous oscillations, self-coupling oscillations and some cross-coupling oscillations). In Figure 3 some examples with one of the best, one of the worst and one of the moderate modeling results obtained with ANMPA model are illustrated.

ISSN: 1727-1983 © 2005 IFMBE

The MLP network was the first implementation for the classifier. The network had two hidden layers with *tanh* activation functions for the hidden PEs. The output layer was compounded of three neurons corresponding to the three different classes (tremor signals acquired under 5 Hz, 10 Hz and 15 Hz visual stimuli). The SoftMax activation functions used for these last PEs are similar to the *tanh* functions, except that the outputs are scaled by the total activation at the output layer. Thus, the sum of the outputs is 1 and the output of the MLP is providing the a posteriori probability of the class given the data.

The second implementation of the classifier was done using a SVM. The kernel function was a Gaussian function, fact that avoided the explicit computation of the pattern projections into the high-dimensional space (the inner product of Gaussians is still a Gaussian).

The feature vectors were divided into two datasets: the training set (90% of the total data) and the cross-validation set (10%), the last being used as a training stop criteria.

The classifier's correct classification rates using an AR model and a MLP network were 84.6% for 5 Hz class, 67.2% for 10 Hz class and 21.4% for 15 Hz class. In the case of adaptive nonlinear Markov process amplitude model and the same MLP network topology the classifier's performances constantly underperformed the AR model. When a SVM network was used the performances obtained for both models of tremor data slightly increased but, once again, the AR model outperformed the ANMPA model.

Discussion

In our opinion the weak results, obtained using the ANMPA model, are mainly due not to some model's weaknesses but rather to an inappropriate usage of it. In fact, the ANMPA model is seeking to describe the behavior of very complicated time series by using the ensembles of some simpler functions (sinusoidal functions with different frequencies and amplitudes). In this idea, the main characteristics of the model are assumed to be related to some few a priori specified spontaneous spectral oscillations. In our case the tremor data were modeled using five spontaneous oscillations (5 Hz, 10 Hz, 15 Hz, 20 Hz, 30 Hz) and their corresponding self- and cross-coupling oscillations.

A first problem with this proposed model is the lack of any rigorous algorithm to determine the optimum number of signal's dominant frequencies. In this study, the ANMPA tremor model was developed under the hypothesis of the nonlinearly coupled frequency components in the tremor signals.

Second, regarding the self- and cross-coupling oscillations frequency corresponding to the main spontaneous oscillations someone can observe a lot of redundant spectral information that unable the model to correctly identify the tremor signal characteristics.

Conclusions

Based on the results presented above we could conclude that the stochastic tremor components dominate the tremor time series and, as a consequence, the tremor series span a space region more closed to ARMA zone (figure 1) characterized by light linearity and stochastic character then that of nonlinearly stochastic processes. Inspite of these, the ANMPA model should not be considered a completely unappropriate model for the tremor data. The frequencies selected to model the data could be the cause for the bad results obtained with this model.

Also, the main problems of the ANMPA model are similar with those resulting from every model's estimation using elementary functions:

- the choice of the basis functions,
- the selection of the elementary function's parameters, and
- the determination of the model's coefficients.

If for the model coefficients' determination there are some very well defined and studied methods the choice of the number of basis functions and the choice of parameters for the elementary functions are very challenging problems.

To solve these problems in the future research we will use a genetic algorithm in order to determine the optimum spontaneous oscillations frequencies.

Acknowledgment

This work was partially supported by Romanian National Council for Research in High Education Grant theme 19, code CNCSIS 493, 2005.

References

[1] RANDALL J. E. (1973): 'A Stochastic Time Series Model for Hand Tremor', *J. Appl. Physiol.,* **34**, pp. 390-395

[2] GANTERT C., HONERKAMP J. and TIMMER J. (1992): 'Analyzing the Dynamics of Hand Tremor Time Series', *Biol. Cybern*., **66**, pp. 479-484

[3] TIMMER J. (1998): 'Modeling Noisy Time Series: Physiological Tremor', *Int. J. Bif. & Chaos in Applied Sciences and Eng.*, **8** (**7**), pp. 1505-1516

[4] PHADKE M. and WU S. (1974): 'Modeling of Continuous Stochastic Processes from Discrete Observations with Application to Sunspots Data', *J. Am. Stat. Ass.*, **69**, pp. 325-329

[5] TIMMER J., HÄUßLER S., LAUK M. and LÜCKING C.H. (2000): 'Pathological Tremors: Deterministic Chaos or Nonlinear Stochastic Oscillators?', *Chaos*, **10**, pp. 278-288

[6] TEODORESCU H. N., CHELARU M., KANDEL A., TOFAN I. and IRIMIA M. (2001): 'Fuzzy Methods in Tremor Assessment, Prediction, and Rehabilitation', *Artificial Life in Medicine*, **21** (1-3), pp. 107-130

[7] GAO J. B. and TUNG W.W. (2002): 'Pathological Tremors as Diffusional Processes', *Biol. Cybern.*, **86**, pp. 263–270

[8] KOYAMA M., IOKIBE T. and SUGIURA T. (2000): 'Quantitative Symptom Discrimination of Parkinson's Disease by Chaotic Approach', *IEICE Trans. Fundamentals*, E83-A (2), pp. 337-342

[9] SERBAN M.C., DOBREA D.M. and TEODORESCU H.N. (2004): 'Evidence for the Central Oscillators in the Physiological Tremor Generation Process', CD Proc. of 4th European Symposium in Biomedical Engineering, Patras, Greece, 2004

[10] BAI O., NAKAMURA M., IKEDA A. and SHIBASAKI H. (2000): 'Nonlinear Markov process amplitude EEG model for nonlinear coupling interaction of spontaneous EEG', *IEEE Trans. on Biomed. Eng.*, **47** (9), pp. 1141 – 1146

[11] DOBREA D.M., TEODORESCU H.N. and MLYNEK D. (2002): 'An Interface for Virtual Reality Applications', *Romanian J. of Information Science and Technology*, **5** (3), pp. 269–282

[12] AL-NASHASH H., AL-ASSAF Y., PAUL J. and THAKOR N. (2004), 'EEG Signal Modeling Using Adaptive Markov process Amplitude', *IEEE Trans. on Biomed. Eng.*, **51** (5), pp. 744 – 751

[13] CARLING, A. (1992): 'Introducing Neural Networks', Wilmslow: Sigma Press, UK, pp. 147-154

[14] CRISTIANINI N., CAMPBELL C. and BURGES C. (2002): Kernel Methods: Current research and Future Directions', *Machine Learning*, **46**(1/3), pp 5-9

ISSN: 1727-1983 © 2005 IFMBE