Proceedings of the 5th WSEAS International Conference on Signal Processing, Istanbul, Turkey, May 27-29, 2006 (pp201-214)

An artificial neural network approach to the classification of inferred intracranial signals

CHRISTOS E. VASIOS[1] and GEORGE K. MATSOPOULOS[2]

[1] Athinoula A. Martinos Center for Biomedical Imaging,
Department of Radiology,
Harvard Medical School,
13th Street, Charlestown, 02129, Massachusetts, USA

[2] School of Electrical and Computer Engineering,
National Technical University of Athens,
9 Iroon Polytehneiou St., 15780, Zografos, GREECE

Abstract: Event-Related Potentials (ERPs) provide non invasive measurements of the electrical activity on the scalp that are linked to the presentation of stimuli and events. Brain mapping techniques are able to provide evidence on the solution of debatable issues in cognitive science. In this paper, an effective signal classification approach is proposed, extending the use of two inversion techniques: the Brain Electrical Tomography using Algebraic Reconstruction Technique (BET-ART) and the Low Resolution Brain Electromagnetic Tomography (LORETA). The first step of the methodology applied is the feature extraction, which is based on the combination of the Multivariate Autoregressive model with the Simulated Annealing technique, in order to extract optimum features, in terms of classification rate. The classification, as the second step of the methodology, is implemented by means of an Artificial Neural Network (ANN) trained with the back-propagation algorithm under the "leave-one-out cross-validation" scenario. The ANN is a multi-layer perceptron, the architecture of which is selected after a detailed search. The proposed methodology has been applied for the classification of First Episode Schizophrenic (FES) patients and normal controls using the intracranial activity distributions obtained by ERPs. A comparative analysis was performed using BET-ART and LORETA inversion methods. Implementation of the proposed methodology provided classification rates of up to 93.1%, for both types of input signals. Additionally, for both BET-ART and LORETA signals, the brain regions that differentiate FES patients from normal controls are located in the frontal brain area, in accordance to the related literature. The proposed methodology may be used for the design of more robust classifiers based on intracranial source distributions, which are more closely related to the underlying cognitive mechanisms responsible for the generation of the scalp-recorded biosignals.

Key-Words: Multivariate Autoregressive Model (MVAR), Neural Networks, Simulated Annealing (SA), Classification, "Leave-one-out" cross-validation, Brain Electrical Tomography using Algebraic Reconstruction Technique (BET-ART), Low Resolution Brain Electromagnetic Tomography (LORETA)

1 Introduction

Event-related potentials (ERPs) provide measurements of electrical activity on the scalp that are linked to the presentation of stimuli or events [1]. The study of ERPs is focused on parts of the waveform containing significant local maxima and minima, called peaks or components. The use of ERPs as diagnostic tools in psychiatry has been enhanced by classification systems integrated in properly designed decision-support systems (DSS). The use of scalp-recorded ERPs has been studied in the literature towards the development of classification systems [2-5]. Anderson et al. [6] explored the use of scalar and multivariate autoregression models to extract EEG features, in order to discriminate different mental tasks. These features were then classified with neural network classifier.

The inversion of cognitive ERPs to intracranial current sources provides a method to observe brain phenomena related to information processing mechanisms. Various methods are currently used, mainly computing discrete brain dipoles or dipolar layers, which generate potentials, on the surface of a

model of the intervening volume conductor that best fit the measured ERPs [7-10].

In contrast to the use of features extracted from scalp-recorded EEG or ERPs, for the development of classification systems, relatively fewer studies explore the use of features from intracranial quantities. Musha et al. [12] classified normal controls from very mild and moderately severe Alzheimer's Disease (AD) patients by means of an alpha-band resting EEG inversion, using a single shell spherical head model, to a single dipole. The goodness-of-fit of the potentials produced by the computed intracranial dipole compared to the measured potentials, defined as the alpha bipolarity, has provided the basis for discriminating among the three subjects' categories. A method, which has been more widely used for investigating differences in various classes of subjects, and more specifically in classification, is the Fast Fourier Transform (FFT) dipole approximation [13-14]. According to this method, starting from multichannel scalp EEG recordings, the brain electrical field data are modeled by a potential distribution map, using the first principal component in the frequency domain. The potential distribution is then used for single dipole source localization. Discriminant analysis is based on the strength and the position coordinates of the dipole, for various frequency bands. The method has been used for the discrimination of AD patients and subjects presenting mild cognitive impairment (MCI) [15]. The best overall correct classification of AD versus control subjects and MCI subjects was 84% and 78%, respectively, while conventional FFT provided similar results.

In this paper, a methodology for the classification of psychiatric disorders, based on intracranial electrical distributions, is presented. The methodology consists of a feature extraction module and an Artificial Neural Network (ANN) classifier. The feature extraction module comprises the use of a novel method combining the Multivariate Autoregressive (MVAR) model with the Simulated Annealing (SA) technique, as a global optimization technique, in order to detect the optimum combinations of the current sources (number and kind), the time interval to use, as well as the order of the model, leading to the classification rate achieved by the neural network classifier. The potential of the Brain Electrical Tomography using Algebraic Reconstruction Technique (BET-ART) and the Low Resolution Brain Electromagnetic Tomography (LORETA) inversion methods is comparatively investigated for the classification of First Episode Schizophrenic (FES) patients against normal controls. The results of this investigation are exploited in order to design a DSS with optimized classification performance, for the case of FES patients, and also in order to investigate brain areas that are crucial for FES patients.

2 System Design

2.1 The Proposed Classification System

The proposed classification system consists of two modules: the feature extraction module and the classification module, as shown in Fig. 1.

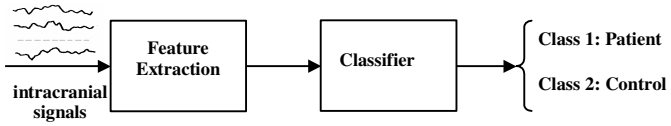

Fig. 1: Block diagram of the proposed system for the classification of intracranial current source waveforms into two classes: patients and normal controls.

The inputs to the first module are current source waveforms, as computed from each inversion process. The appropriate features are extracted and processed by the feature extraction module, and then fed to the classification module. The output of the system encodes two classes: patients or control subjects.

2.2 Feature Extraction Module

In the present work, the MVAR model is used for feature extraction. When using the MVAR model in ERPs classification problems, for the construction of the feature vector, a number of parameters have to be selected, regarding the number and the kind of signals (i.e. intracranial sources) whose waveforms will be modeled, the time interval of the waveforms to be modeled and the order of the model to be used. An exhaustive search for the selection of the parameters, that achieve the best classification rate, seems practically very difficult, as it creates a very large space for all parameter combinations. A further disadvantage of the MVAR model is the dependence of the model coefficients on the used input signals; thus, any modification of the input signals requires recalculation of the MVAR coefficients. When the classification is based on intracranial current source waveforms, the same problems exist concerning the selection of the number and kind of the source positions, the time interval of the waveforms to be modeled and the order of the model to be used.

For these reasons, a new method for the extraction of MVAR coefficients from computed intracranial current source waveforms is proposed in the present work. The method combines the MVAR model with a global optimization method, the SA technique

[16], in order to detect optimum combinations of sources (number and kind), time interval and model order, leading to the best classification performance of a neural network classifier.

The implementation of the MVAR model to current source signals is based on the principle that the signals are described by a linear filter fed with noise. According to this model, each value of the signal can be estimated using the values of the preceding *p* samples, as follows [17-18]:

$$x(k) = -A(1)x(k-1) - A(2)x(k-2) + \cdots - A(p)x(k-p) + e(k) \qquad k=p,\ldots,N \quad (1)$$

where N is the maximum number of samples available. The procedure works for samples with the index *(p, p+1, ...)*, i.e. starting after the p minimum number of initialization inputs. In Eq. (1), $x(k)$ is a $d-$dimensional vector of data at sample with index *k* and *e(k)* is a $d-$dimensional vector of random input (noise). Furthermore *A(i), i=1,...,p* are the $d \times d$ matrices of the AR coefficients to be estimated from *x(k), k=1,...,N* and *p* is the model order. These coefficients construct the feature vector of each subject.

According to the proposed feature extraction methodology, an optimum combination of current sources, in terms of number and kind, is obtained by implementing the MVAR model in conjunction with the SA technique, which provides the advantage of finding possible global minimum (or maximum) in contrast with other local optimization methods, which require a good initial guess and are often trapped to local minima (or maxima) [19-20]. This optimum selection is tested for different model orders, within a pre-defined interval, based on the performance of the Fuzzy C-Means classifier (FCM) [21].

More specifically, in feature extraction from intacranial source data, an initial random selection of inputs (in terms of kind and number) is considered. For this selection, the MVAR coefficients are extracted, constructing the feature vector for each subject. These coefficients are fed to a classifier, based on the FCM algorithm, and the classification rate is then calculated. After several temperatures, as defined by the SA schedule, an optimum combination of sources is extracted corresponding to the best classification rate achieved. The proposed feature extraction methodology is presented in the form of pseudocode in the following:

Step 1: Define the model order p.
Step 2: Search for the optimum combination of inputs using the SA technique
  Step 2.1: Define the kind and number of inputs
  Set initial Temperature
  Random selection of initial combination of inputs
  For i=1 to a number of temperatures do
  Begin
    For j=1 to maximum number of combinations per temperature
    Begin
      Step 2.2: Selection of next combination of inputs based on the current combination of inputs and the current temperature
      Step 2.3: Calculation of MVAR Coefficients
      Step 2.4: Calculation of classification rate, using the FCM algorithm
      Step 2.5: Acceptance of the current combination based on the Boltzmann distribution
    End
    Reduction of Temperature
  End

It must be pointed out that the FCM algorithm was selected in Step 2.4, as an objective function of the SA technique, since it requires minimum parameter trimming compared to neural networks.

The selection of the next combination of input signals depends on the current one and the current temperature. The higher the temperature, the smaller the number of inputs that participate in each change. Given a combination of inputs, for the choice of the next combination, one of the following changes were taken place: a)insertion of inputs, b)abstraction of inputs, c)alteration of inputs, d)insertion and alteration of inputs, and e)abstraction and alteration of inputs. In cases of insertion or alteration of inputs, the lower the temperature, the smaller the distance between the new input and the current combination of inputs.

According to the aforementioned MVAR model, a feature vector is constructed with a dimensionality of $p \times d \times d$, where *p* is the model order and *d* is the number source signals.

2.3 Classification Module: selection of neural network structure

The classification module contains a multi-layer perceptron ANN trained with the back-propagation algorithm. The selection of the topology of the ANN is a methodological aspect that was investigated in the present work. Various methodologies for the selection of the number and the size of hidden layers in ANNs have been used, including evolutionary

strategies and genetic algorithms [22-23], network pruning techniques [24], network growing techniques [25], as well as extensive network architecture search [26][1].

In the present work, we opted for an extensive network architecture search strategy, scanning combinations of network structure parameters, in order to compare the performance of 3-layered and 4-layered networks with one or two output neurons. Specificity was computed as the percent ratio of the correctly classified controls to the total number of controls. Sensitivity was computed as the percent ratio of the correctly classified patients to the total number of patients. Negative predictive value was computed as the percent ratio of correctly classified controls to the total number of subjects classified as controls and positive predictive value the percent ratio of correctly classified patients to the total number of subjects classified as patients. The overall classification rate (stated simply, in the following, as classification rate - CR) was computed as the percent ratio of correctly classified subjects to the total number of subjects used in the study. After a preliminary investigation, concerning sources as input signals, it was found that satisfactory results, i.e. an overall classification rate of greater than 90%, without either sensitivity or specificity being less than 85%, were achieved using four or three source combinations, with an order of 4. Therefore, we chose to search for the optimum number of hidden layer neurons using a combination of four sources and an order of 4, leading to a feature vector of 64 (see end of Section 2.2 - Feature Extraction Module). Initially 3-layer neural networks with one and two output neurons and hidden layer neurons ranging from 4 to 40 (with steps of 4) were tested and afterwards 4-layer neural networks with the first hidden layer neurons ranging from 4 to 40 (with steps of 4). For each number of neurons in the first hidden layer, the neurons in the second hidden layer varied from 4 to 40 (with steps of 4).

The results of the tests suggested broadly similar performance for 3-layered and 4-layered networks with one or two output neurons. Furthermore, given that the input layer consisted of 64 neurons, the performance of the network was not significantly influenced by the number of neurons in the first hidden layer (which, of course, is the sole hidden layer in 3-layered ANNs), as long as that number was between 12 and 28 (approximately equal or more than 1/5 and equal or less than 2/5 of the number of input neurons). Reducing the neurons in the first layer to less than 12, led to a gradual reduction of the performance. It should be stressed that the selection of the number of first hidden layer neurons is purely empirical and it was obtained experimentally using the set of data presented in the present study. Based on the above empirical test results, and taking into account that minimizing the size and number of the hidden layers, while maintaining acceptable performance as the major goal of the search strategy, the networks implemented in the present study were 3-layered, with the ratio of neurons at the input and hidden layer around 5 and 1 neuron in the output layer.

As a result of the above network structure selection investigation, the classification module is implemented with an ANN consisting of three layers (see Fig. 2).

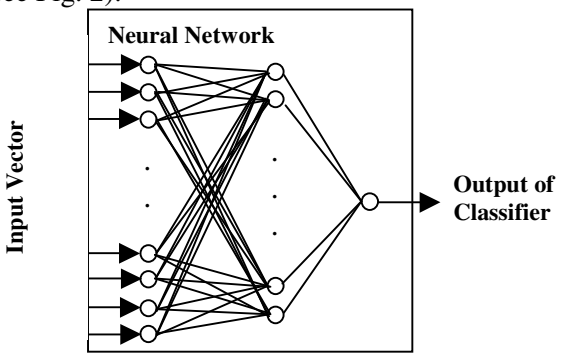

Fig. 2: Architecture of the neural network classifier.

The input layer consists of a number of neurons equal to the number of the selected features. The hidden layer contains a number of neurons equal to one fifth of the input neurons. The output layer consists of one neuron, encoding the two classes of the subjects: patient and normal (0=patient and 1=normal).

3 Measurement and computational methods

3.1 Subjects

Fourteen (14) never medicated FES patients (8 men and 6 women) with mean age 29 (±7) years were matched for age and sex to 30 healthy controls (20 men and 10 women) with mean age 31 (±3) years. The mean educational level was 11 (±3) and 12 (±3) years of schooling, for the patients and controls, respectively.

All patients met DSM-IV criteria (American Psychiatric Association, 1994) for schizophrenic disorder, paranoid type. The diagnosis was verified independently by two psychiatrists. Age at onset was defined as the earliest age at which medical advice was sought for psychiatric reasons or at which subjective distress or impairment of functioning was observed [27]. The controls were recruited from hospital staff and local volunteer groups. They were free of psychiatric and physical

illness. All participants had no history of any neurological or hearing problems. All participants were right-handed as assessed by the Edinburg Inventory. Written informed consent was obtained from both patients and controls.

3.2 Stimuli and ERP recording procedure

Patients and controls were evaluated by a computerized version of the digit span Wechsler test [28-29]. The subjects sat in an anatomical chair placed inside an electromagnetically shielded room. A single sound of either high (3,000 Hz) or low frequency (500 Hz) was presented to the subjects, who were asked to memorize the numbers that followed. The warning stimulus lasted 100 msec. A 1-sec interval followed and then the numbers to be memorized were presented. Two to nine one-digit numbers were presented in each trial. The time interval between administered numbers was one second. At the end of the number sequence presentation, the signal tone was repeated and subjects were asked to recall the administered numbers as quickly as possible. The numbers were recalled by the subject in the same (low frequency tone) or in the opposite order (high frequency tone) than that presented to him/her. Each test session involved 26 repetitions of the trial, 13 repetitions with low frequency tone and 13 with high frequency tone. The sequence of the warning stimuli was pseudorandom. Warning stimuli and numbers were delivered using earphones at 65 dB. ERPs were recorded during the 1-sec interval between the warning stimulus and the first administered number, for each of the 26 test repetitions.

Referential ERP recordings were performed using Ag/AgCl electrodes (resistance < 5 kΩ) at positions Fp1, Fp2, F3, F4, C3, C4, C3-T5/2, C4-T6/2, P3, P4, O1, O2, Pz, Cz and Fz according to the 10-20 International system of Electroencephalography [30]. Data was digitized at a sampling rate of 500 Hz and band-pass filtered from 0.05 to 35 Hz. During the administration of the stimuli, the subjects had their eyes closed in order to minimize eye movements and blinks. Recordings were averaged for each electrode, by a computerized system, as a pre-processing denoising step of the procedure.

The study was concentrated on the time interval from 500 to 800 msec, corresponding to the P600 ERP component. The focus on this component was based on its importance in reflecting information processing in the brain [31]. Furthermore, the P600 component is accepted as reflecting the completion of any synchronized operations, concerning a decision taken after the presentation of a warning stimulus and target detection. Specifically, its amplitude is considered as an index of the cost of processing, while its latency is considered as a function of the onset and the duration of processes [32].

3.3 Intracranial Distributions Computation

The computation of intracranial distributions was carried out using both the BET-ART and the LORETA inversion methods.

A) BET-ART

According to this method, in order to solve the inverse problem, a three-layered spherical human head model was considered simulating brain, skull and scalp media, with corresponding radii r_1, r_2, r_3 and conductances σ_1, σ_2, σ_3 [33]. The brain region supposed active was divided into n small volumes Δv_k, called voxels, centered at points $\underline{r}_k=(r_k,\theta_k,\varphi_k)$, $(r_k<r_1)$, carrying an unknown current source density ρ_k, $k=1,...,n$ (Amperes/m^3). Scalp potentials V_i were measured at $m=15$ points $\underline{r}_i=(r_3,\theta_i,\varphi_i)$, $i=1,...,m$, corresponding to the ERP electrodes used in the present study. The inversion problem was formulated in matrix form as follows:

$$A\cdot\underline{X} = \underline{Y} \tag{2}$$

where $\underline{Y}=[y_i]$, $y_i=V_i$, $i=1,...,m$, $\underline{X}=[x_k]$, $x_k=\rho_k\Delta v_k$, $k=1,...,n$, $A=[a_{ik}]$, $a_{ik}=G_3(\underline{r}_i,\underline{r}_k)$ and G_3 is the Green's function connecting the unknown source position at $\underline{r}_k$ with the observation point $\underline{r}_i$. The current quantities x_k correspond to point current activities located at the center of each voxel, representing the whole voxel activity.

The solution of Eq. (2) was provided by the Algebraic Reconstruction Technique-1 (ART-1) algorithm [34], which is an iterative procedure seeking the minimum norm solution of the inverse problem. The algorithm can be stated as follows:

Initial Values (Iterative step q=0): Start with an assumed distribution $\underline{X}=\underline{X}^0$, where $\underline{X}^0$ is an arbitrary element of the set of all the linear combinations of $\underline{a}_i^T$'s.

Typical iterative step (q≥1): Apply the recursive formula:

$$\underline{X}^{(q+1)} = \underline{X}^{(q)} + r^{(q)}\,\underline{a}_i^T\,(y_i-\langle\underline{a}_i,\underline{X}^{(q)}\rangle)/\|\underline{a}_i\|^2 \tag{3}$$

where $\underline{a}_i=(a_{i1},...,a_{in})$ is the i-th line of matrix A, $\underline{a}_i^T$ is the transpose of $\underline{a}i$, $\langle\underline{a}_i,\underline{X}_{(q)}\rangle$ is the inner product of vectors $\underline{a}_i$ and $\underline{X}^{(q)}$, $\|\underline{a}_i\|^2=\langle\underline{a}_i,\underline{a}_i^T\rangle$ and $r^{(q)}$ are relaxation parameters. In Eq. (3), it is $i = (q \bmod m) + 1$, so that the lines of A, and the elements of $\underline{Y}$ were used cyclically, and therefore $\underline{X}$ was examined after each full cycle, i.e. $m=15$ iterations. The

iterations were stopped when the parameter $D(L)=||\underline{X}^{(L+1)m}-X^{Lm}||/||\underline{X}^{Lm}||$, expressing the rate of change between successive reconstructions (where L is the number of cyclic iterations executed), reached a low threshold of 0.001.

In order to implement the inversion procedure, the geometrical (r_1, r_2, r_3) and electrical $(\sigma_1, \sigma_2, \sigma_3)$ model parameters as well as the coordinates of the external electrode positions and the current source positions under investigation were selected. The digitized and averaged ERP data, for each subject, were regrouped to represent the spatial sample of the head surface potential function, for each sampled time moment, resulting in 500 ERP sets. The inversion of these sets to current values was then performed. The region where the unknown current sources were computed was restricted into a spherical shell, located at a radius of $r_S<r_1$, corresponding to the outer layers of the cortex [34]. A number of 60 current source positions distributed throughout the spherical shell, were investigated (Fig. 3).

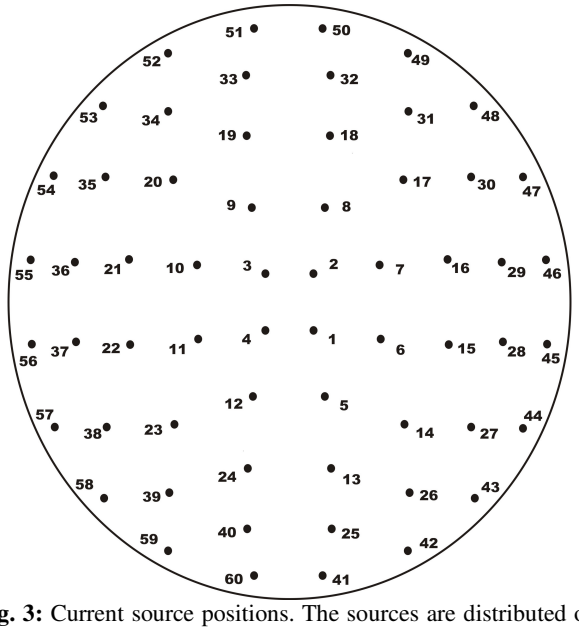

Fig. 3: Current source positions. The sources are distributed on a spherical shell layer, which is projected with the front of the head corresponding upwards.

B) LORETA

LORETA was used to compute the 3-dimensional intracerebral distributions of current density. This method was implemented by Pascual-Marqui [35] at the KEY Institute for Brain-Mind Research, University Hospital of Psychiatry, Zurich, Switzerland. The algorithm solves the inverse problem assuming related orientations and strengths of neighboring neuronal sources. Mathematically this assumption is implemented by finding the 'smoothest' of all possible activity distributions.

The LORETA version used in the present study [36], considered a three-shell spherical head model that was finally registered to the Talairach brain atlas [37]. Based on the digitized Talairach brain atlas and the probability atlas of the Brain Imaging Centre (Montreal Neurologic Institute), solution space was restricted to cortical gray matter and hippocampus.

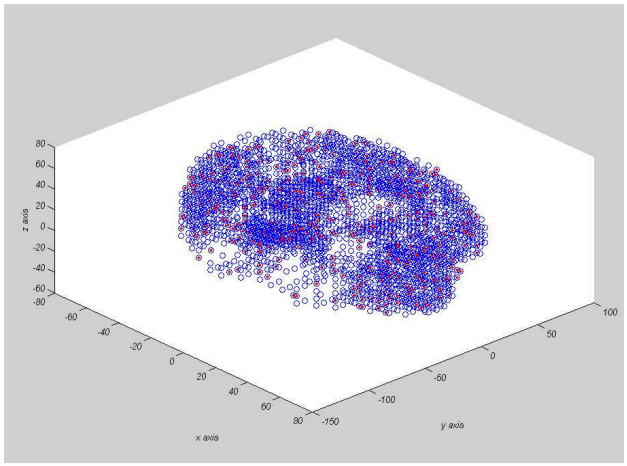

Fig 4: Three-dimensional representation of the solution space of LORETA consisting of 2,394 voxels at the spatial resolution of 7 mm. The 478 sources selected after the dimensionality reduction are also presented in red color.

The spatial resolution of the method was 7mm and the solution space consisted of 2,394 voxels, as presented in Fig. 4. According to LORETA algorithm the current density was computed at each voxel as the linear weighted sum of the scalp electric potentials. Thus, LORETA combines the high time resolution of the EEG/ERP with a source localization method that permits truly three-dimensional tomography of the brain electrical activity.

3.4 Intracranial Distributions Computation

When the MVAR model is used, in ERP source modeling, it requires the definition of several parameters such as the number and kind of signals, the time interval of the examined waveforms and the order of the model used. Even though in our study we considered a fixed time interval (500-800 msec), the search space constructed by the combination of the remaining aforementioned parameters, seems practically non-manageable, especially for the case of the LORETA inversion, which computes a great number of intracranial signals (up to 2,394). Consequently, a pre-processing step is required before the feature extraction module for each of the inversion methods, in order to reduce the computational intensity of the procedure.

3.4.1 Source activity computed using BET-ART

As a pre-processing step, the implementation process proceeded into three distinct Phases:

In ***Phase-1***, ERPs were inverted to 60 current source waveforms, covering the full set of source positions. The algorithm was allowed to create source position combinations from 40 positions, specifically positions 1-4, 6-11, 15-22, 28-37 and 45-56, i.e. the positions corresponding mainly to occipital and parietal regions were omitted from the allowed combination search space.

In ***Phase-2***, the number of positions from which the algorithm was allowed to create combinations was further restricted to 22 positions, specifically positions 2, 3, 8, 9, 17-20, 30-35 and 47-54, corresponding mainly to frontal, central and temporoparietal regions.

In ***Phase-3***, the allowed combination search space was as in *Phase-2*, but ERPs were inverted to current source waveforms restricted to the above set of 22 positions.

The parameters used for the inversion problem solution were $r_S=8.7$ *cm*, $r_1=9.0$ *cm*, $r_2=9.7$ *cm*, $r_3=10$ *cm*, $\sigma_1=\sigma_3=0.33$ *S/m*, $\sigma_2=0.0042$ *S/m* [38-42]. ART1 algorithm was implemented starting from an initial current source distribution $\underline{X}=\underline{X}^0=0$, i.e. equal to the null vector, since no a priori information existed about the distribution and an unbiased solution was sought. As previous research has shown [34], when no estimate on the level of noise present in the recorded potentials exists, the best selection for the relaxation parameters is $r^{(q)}=1, \forall q$, and this value was used in the present work.

In all phases, the order of the model used varied from 3 to 15 and the number of source positions in each combination varied from 2 to 8. The MVAR parameters were extracted for the time interval from 500 to 800 msec, in correspondence to the ERPs' P600 time interval. The parameters of the SA technique were determined experimentally and set to: initial guess=random, initial temperature=5, percentage of temperature reduction at each iteration=10%, number of temperatures=40, maximum number of combinations per temperature=40. Consequently, in accordance with the methodology given in Section 2.2 - Feature Extraction Module, for each model order that was investigated, an initial random combination of two source positions was selected and then the SA algorithm was implemented, leading to a selected feature vector for a source combination (kind and number) that might differ from the initial one. Then the same procedure was repeated for an initial random combination of 3 up to 8 source positions.

3.4.2 Source activity computed using LORETA

A pre-processing step is initially performed, in order to reduce the total number of 2,394 input sources as obtained by LORETA. The source space was divided into small cubic neighbours by placing the source under investigation in the centre of the cube and the rest neighbour sources at the vertices. The waveforms of all the sources belonging to this cube were then compared by means of the correlation criterion. Based on this criterion the tests indicated that adjacent sources have similar waveforms. Such a finding seems quite reasonable because of the LORETA's choice of the smoothest inverse solution and also of the relatively small electrodes/sources ratio (16/2,394 ≈ 1/150). Since similar waveforms do not affect the classification performance we can reduce the initial 2,394 input sources by a factor of 5 ending up to 478 sources (see Fig. 4) uniformly distributed in the cortical gray matter and hippocampus.

The parameters of the SA technique were determined experimentally and set to: initial guess=random, initial temperature=5, percentage of temperature reduction at each iteration=5%, number of temperatures=50, maximum number of combinations per temperature=40. As in the case of BET-ART sources, the order of the model used varied from 3 to 15, the number of source positions in each combination varied from 2 to 8 and the MVAR parameters were extracted for the time interval from 500 to 800 msec.

3.5 Classification Module

The back-propagation algorithm with adaptive learning rate and momentum has been used in order to train the ANN [43]. The initial weights of the neurons have been randomly selected in the range [–1.0, +1.0]. The log-sigmoid and tan-sigmoid activation functions have been used for the hidden and the output layer, respectively. The values of the learning rate and the momentum have been estimated using a process of trial-and-error, until no further improvement in classification could be obtained.

In order to test the performance of the network, in a reliable manner, taking into account the limited number of control and patient subjects available, while aiming to avoid overtraining and achieve an acceptable generalization in the classification, the leave-one-out cross-validation procedure was adopted [9]. According to this procedure, the neural network is trained using all the patients and control subjects, except from one (no mater if is a patient or a control subject), which will be used for testing. The generalization ability of the specific network is

tested using the single excepted subject. The aforementioned training-testing procedure is repeated using a different subject for testing, until all subjects are used once each. Using the leave-one-out cross-validation procedure, the resulting ANNs present slight differences between each other, by inference of the slight variation of the training and validation sets and testing subject in each one. The classification performance parameters are computed by the aggregate sums of correctly classified or misclassified controls and patients. The main alternative to cross-validation procedures is the static-split procedure, according to which the whole subject set is divided into one training and one test set and the classification performance parameters are computed once for the testing set. In a variation of the static-split procedure, the subject set is divided into three sets, the training, the validation and the test set. The ANN is constructed based on the training set, but training is stopped when an error measure is minimized on the validation set, and then the classification performance parameters are computed once for the testing set. Therefore, using cross-validation procedures provides the opportunity to make multiple tests based on the existing subject set, employing the entire data for both training and testing, instead of one test of performance in the static-split procedures, based on the use of part of the data for training and another part for testing, which leads to a loss of available information. Consequently, the classification performance parameters computed using cross-validation might provide a more realistic approximation of the performance of the network, when it will encounter new data different from the data available in the present study, than the performance parameters computed using one fixed part of the subject set for testing in the static-split procedures.

4 Experimental results

4.1 BET-ART data

Classification results obtained with the *MVAR/SA* method, using BET-ART data, are presented in Table 1.

For *Phase-1* of the feature extraction and classification process the highest classification rate (72.7%) was achieved for the source position combinations {8 17 18 20 32} for model order of 5, misclassifying 7 out of 30 normal controls and 5 out of 14 FES patients. The dimension of the feature vector produced by the feature extraction module was 125.

For *Phase-2* the highest classification rate, 86.3%, was achieved for the source position combination {8 18} and for model orders of 6 and 8, respectively. The dimensions of the feature vectors produced by the feature extraction module were 24 and 32, respectively. For both combinations, 4 out of 30 controls and 2 out of 14 FES patients have been misclassified by the proposed system. In Fig. 5(a) we present the topographical distribution of the source positions for this combination.

Table 1: Performance of the MVAR/SA feature extraction method implemented on BET-ART data, for *Phase-1*, *Phase-2* and *Phase-3*. The first column corresponds to the model order (p), the second to the resulting dimension of the feature vector (D), while the two last columns present the classification rate (CR %) achieved by each source position combination and the ratio of misclassified controls to total (N=30) controls and misclassified patients to total (N=14) patients.

Phase-1				
Source positions	***p***	***D***	***CR (%)***	***Misclassified subjects***
8 17 18 20 32	5	125	**72.7**	7/30 - 5/14
8 17 18 19 31	8	200	70.4	8/30 - 5/14
3 4 15 16 33 35	4	144	70.4	9/30 - 4/14
17 19 20 35 36 37	5	180	68.1	7/30 - 7/14
Phase-2				
Source positions	***p***	***D***	***CR (%)***	***Misclassified subjects***
8 18	4	16	84.0	4/30 - 3/14
8 18	6	24	**86.3**	4/30 - 2/14
8 18	7	28	84.0	4/30 - 3/14
8 18	8	32	**86.3**	4/30 - 2/14
8 18	9	36	84.0	4/30 - 3/14
8 9 17 20	4	64	72.7	7/30 - 5/14
8 16 18	6	54	70.4	8/30 - 5/14
2 3 8 9 17 20	4	144	70.4	9/30 - 4/14
Phase-3				
Source positions	***p***	***D***	***CR (%)***	***Misclassified subjects***
8 18 32	4	36	**93.1**	1/30 - 2/14
8 18 32 50	4	64	**93.1**	1/30 - 2/14
3 8 18 32	4	64	**93.1**	1/30 - 2/14
2 17 32 51	4	64	90.9	2/30 - 2/14
33 34 35	7	63	86.3	2/30 - 4/14
3 17 19 33	4	64	86.3	3/30 - 3/14
8 18 32 49	4	64	84.0	4/30 - 3/14
19 30 34 53	4	64	84.0	3/30 - 4/14
17 20 32 52	7	112	72.7	7/30 - 5/14

Finally, for *Phase-3*, a classification rate of 93.1%, was achieved for the source position combinations

{8 18 32}, {8 18 32 50} and {3 8 18 32} for model order 4. The dimensions of the feature vectors produced by the feature extraction module were 36, 64, and 64 respectively. For these combinations, 1 out of 30 controls and 2 out of 14 FES patients have been misclassified by the proposed system. In Fig. 5(b-e), the topographical distribution of the source positions for these three combinations as well as for the combination {2 17 32 51} with a classification rate of 90.9%, are presented.

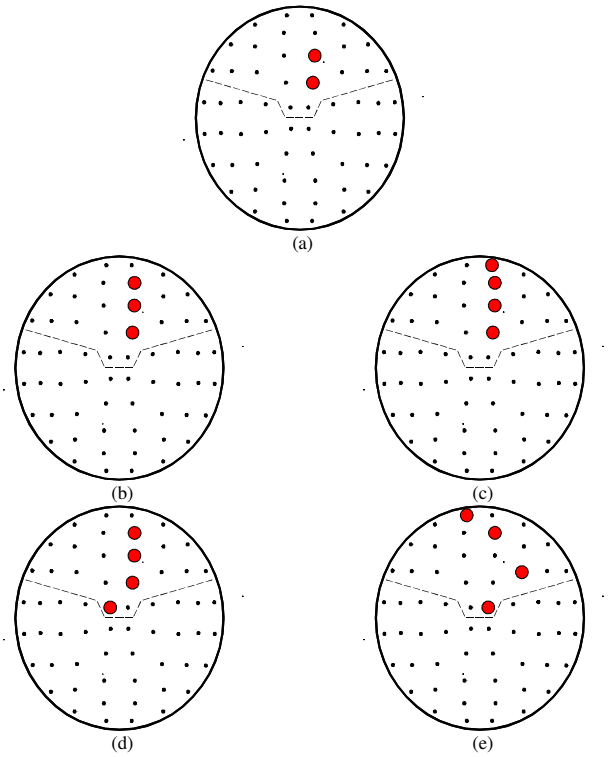

Fig. 5: Topographical distribution of current source positions: (a) For the position combination {8 18}, providing the best classification rate (86.3%) in *Phase-2* of the feature selection and classification implementation processes. (b-e) For the position combinations {8 18 32}, {8 18 32 50}, {3 8 18 32} and {2 17 32 51} providing a classification rate higher than 90% in *Phase-3*. In each figure, the sources are distributed on a spherical shell layer, which is projected with the front of the head corresponding upwards. The region that was included in the source position search space is located upwards from the dashed line.

In Table 2, the sensitivity, specificity, negative predictive and positive predictive values are explicitly given for the optimum source combinations of each phase. The overall classification rates of Table 2 are the same as those given in the corresponding rows of Table 1 for the best current source combination of each phase.

The computational time for the MVAR feature extraction process of all phases was 14 ± 1 min using a personal computer possessing an Intel Pentium 4 CPU, running at 2 GHz, with 1 GB RAM. Then, the computational time for the totality of the 44 training-testing rounds needed to complete the leave-one-out cross-validation procedure was 1 min and 28 sec.

Table 2: Sensitivity, specificity, negative predictive and positive predictive values calculated for the best source combination of each phase using the BET-ART method.

Phase-1: {8 17 18 20 32}, p=5

	Controls	**FES**		
Controls	23	7	**76.6**	***Specificity***
FES	5	9	**64.2**	***Sensitivity***
	82.1	**56.2**	**72.7**	***Overall classification rate***
	Negative predictive value	***Positive predictive value***		

Phase-2: {8 18}, p=6

	Controls	**FES**		
Controls	26	4	**86.6**	***Specificity***
FES	2	12	**85.7**	***Sensitivity***
	92.8	**75.0**	**86.3**	***Overall classification rate***
	Negative predictive value	***Positive predictive value***		

Phase-3: {8 18 32}, {8 18 32 50}, {3 8 18 32}, p=4

	Controls	**FES**		
Controls	29	1	**96.6**	***Specificity***
FES	2	12	**85.7**	***Sensitivity***
	93.5	**92.3**	**93.1**	***Overall classification rate***
	Negative predictive value	***Positive predictive value***		

4.2 LORETA data

Classification results obtained with the *MVAR/SA* method are presented in Table 3.

The highest classification rate (93.1%) was achieved for the source combination {766, 1521}. Sensitivity, specificity, negative predictive and positive predictive values are the same as in the best performance *Phase-3* case for the BET-ART data. For combinations {1476, 1661}, {571, 366}, {946, 1691}, {571, 1366} and {1336, 1431, 1921} the classification rate was 90.9% and for combination {881, 1266, 1496, 2226} was 88.6%. The dimensions of the feature vectors produced by the feature extraction module were 16, 20, 20, 20 16, 45

and 64 respectively. In Fig. 6 we present the topographical distribution of the source positions for optimum combinations.

Table 3: Performance of the *MVAR/SA* feature extraction method implemented on LORETA source data. The first column corresponds to the model order (p), the second to the resulting dimension of the feature vector (D), while the two last columns present the classification rate (CR %) achieved by each source position combination and the ratio of misclassified controls to total (N=30) controls and misclassified patients to total (N=14) patients.

Source positions	***p***	***D***	***CR (%)***	***Misclassified subjects***
766, 1521	4	16	**93.1**	1/30 - 2/14
1476, 1661	5	20	90.9	2/30 - 2/14
571, 1366	5	20	90.0	2/30 - 2/14
946,1691	5	20	90.9	3/30 – 1/14
571, 1366	4	16	90.9	2/30 - 2/14
1336, 1431, 1921	5	45	90.9	2/30 - 2/14
881,1266,1496,2226	4	64	88.6	2/30 - 3/14

The computational time for the feature extraction process of all phases was 21 ± 1 min using a PC (Pentium 4, 2 GHz with 1 GB RAM). The computational time for the classification process, as in the case of the BET-ART data, was in the range of seconds.

The computational time for the MVAR feature extraction process was 21 ± 1 min using a personal computer possessing an Intel Pentium 4 CPU, running at 2 GHz, with 1 GB RAM. The computational time for the totality of the 44 training-testing rounds needed to complete the leave-one-out cross-validation procedure was, as in the case of the BET-ART data, 1 min and 28 sec.

5 Discussion

The requirement to use information existing in concurrently recorded ERP source waveforms, leads to the creation of an unpractical large search space for selecting the MVAR model providing the best classification rate. The combination of the MVAR model with the SA optimization technique, as proposed in the present work, provides a principled way to reduce the computational complexity of the search process. Nevertheless, even using the SA technique, we followed a gradual approach concerning the degrees of freedom allowed in the search process, so that a compromise between satisfactory classification rates and computational complexity is reached. The classifier that has been used is cross-validated, so as to overcome, as far as possible the limited number of patients subjects and also to provide a classifier that is less prone to overfitting the existing data.

In *Phase-1* of the feature extraction and classification process for the current sources (sub-section 3.4.1), 20 positions corresponding mainly to occipital and parietal regions were omitted from the allowed combination search space. In *Phase-2* the search space was further restricted to 22 mainly frontal, central and temporoparietal positions. The focus of the present study on sources reflecting activity of frontal, central and temporoparietal brain regions, has been based on evidence from previous investigations suggesting the involvement of those regions in schizophrenic symptomatology [44].

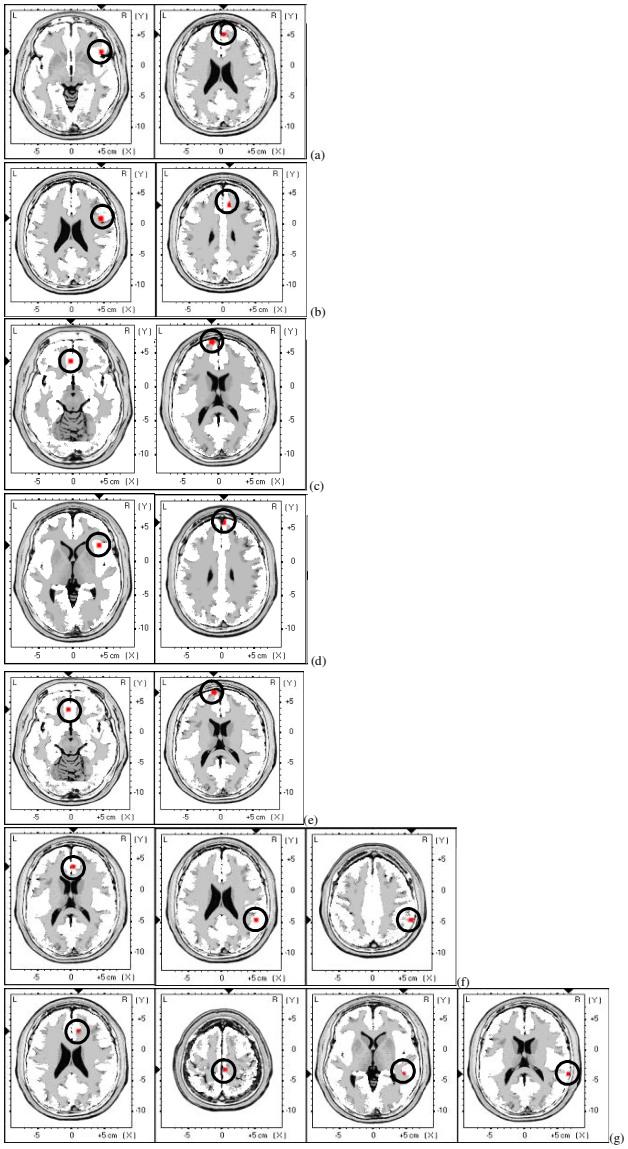

Fig. 6: Horizontal planes of optimum source combinations, registered to the cortical gray matter and hippocampus of the Talairach brain atlas. (a):{766, 1521}, (b):{1476, 1661}, (c):{571, 1366}, (d):{946,1691} (e):{571, 1366}, (f):{1336, 1431, 1921}, (g):{881, 1266, 1496, 2226}

The performance in the current source classification case of *Phase-2* was improved as compared to Phase-1. The larger domain search in *Phase-1* did not have to end necessarily with the same or similar results with the reduced domain search that took place in *Phase-2*, since the search, concerning the possible combinations of source positions, was not exhaustive and the optimality of the final selection, for each model order, depends on the performance of the SA technique (see end of sub-section 3.4.1 - Source activity computed using BET-ART). By restricting the source position search space to frontal, central and temporoparietal positions during the transition from *Phase-1* to the *Phase-2*, the method became more «focused» on a topographically reduced search space and there existed the possibility that the MVAR/SA algorithm, having a smaller space to search, would be less easily trapped into less than optimal local minima, as might have been the case for the larger search domain in *Phase-1*. At this point it should be noted that when similarly topographically restricted search spaces were investigated, covering other brain regions, no significant improvement in the performance of the system, as compared to *Phase-1* was found. The transition from *Phase-2* to *Phase-3* of the implementation process for the current source classification problem provided a further improvement. In *Phase-3*, not only the search space was reduced, as in *Phase-2*, but the whole source activity was assumed to take place in the 22 positions defined in *Phase-2*. This assumption implies that no significant electrical activity related to the experimental task condition takes place in any other position apart from the considered set of 22 positions. This might not be possibly justified solely on indications of brain studies in schizophrenia, which nevertheless stress the importance of some of the regions corresponding to these positions [44]. *Phase-3* investigation was undertaken in an effort to provide a set of current source waveforms that would be searched by the MVAR/SA method making use, through the inversion process, of the whole information present in the initial ERP curves. It remains to be elucidated in future studies whether the use of the whole information present in the initial ERP curves is a reason for the improvement found in the classification results.

Apart from using a sub-set of the available current source positions for creating the search space, during the implementation process, the proposed system allowed a maximum number of 8 source positions for the current source classification case, to be combined for the calculation of the MVAR coefficients. This decision was also taken in order to reduce the computational complexity of the search process. After several trials, it was found that optimal classification results corresponded to combinations consisting of less than 8 source position. As a result, further expansion of the process, by combining more than 8 source positions, was deemed not necessary.

In the case of the LORETA method, the compromise between satisfactory classification rates and computational complexity was taken in one step, by subsampling the initial number of sources, after a principled step-wise approach. In this case, no restrictions were placed on which brain regions would be investigated, so as to make use of one of the main advantages of the currently publicly available LORETA software implementation, i.e. the correspondence of source positions to specific Talairach coordinates and brain voxels.

The use of data from two inversion techniques has been used to compare the results they provide, both concerning the classification rate, as well as the indications about the intracranial source positions that seem to be implicated in the best classification cases. A first conclusion that may be reached, through the data set used in the present study, is that comparable classification results are achievable. Furthermore the intracranial signals of right frontal regions provide the best results in both methods. Both of the above similarities in performance are interesting, taking into account the different electric models used by the methods. The above comparison should also be done having in mind the limitations of the BET-ART method concerning spatial resolution [34].

The present study concentrated on the time interval corresponding to the P600 ERP component. The motivation for selecting the P600 component is based on research indicating its relation to cognitive processes, such as its association to mnemonic binding processes, by which internal and external aspects of information are linked into a coherent representation [45], leading to the conceptualization of P600 as an index of information processing that assigns a specific response to a specific stimulus [46].

The rationale for investigating, the capacity of current source waveforms in providing satisfactory classification performance, is that intracranial currents computed by the scalp-recorded ERPs, provide information on the non-observable electrical phenomena taking place in the brain, related to the cognitive mechanisms induced by the experimental task used in the ERP recording procedure. Therefore, the existence of source positions that are

repeatedly present in combinations providing satisfactory classification results, for various model orders and implementation phases, such as positions 8 and 18, may indicate the significance of the brain regions corresponding to those positions, in differentiating between normal and pathological mechanisms in schizophrenia, in agreement to research results concerning the involvement of those regions in schizophrenia [44]. On the other hand, though features extracted from intracranial quantities might be more closely related to the actual pathophysiological processes, attention should be given to the fact that overlearning may still occur and be harmful, even if the used features are closely related to the modeled processes, as long as they do not cover exhaustively any possible input instances.

6 Conclusions and future plans

In this paper, a methodology for the classification of intracranial signals was proposed for the case of FES patients where the existing markers are inconclusive. The proposed methodology combines the MVAR/SA method for the feature extraction of the input signals with a cross-validated ANN for the classification. Two different input signals had been used to test the performance of the system: intracranial waveforms computed using BET-ART and LORETA methods. The maximum obtained classification rate of 93.1% showed that specific source combinations could be used as inputs to a DSS in order to successfully classify FES patients from normal controls.

The methodology was based on a number of constraints. These constraints refer to (a) the number of source positions, which constitute the allowed search space, (b) the time interval of current source signals modeled and (c) the model order in the MVAR/SA method and the number and kind of the sources, and (d) the limited number of subjects available for the study. Especially concerning the last constraint, it should be noted that patient samples are often unavoidably restricted due to the kind of psychopathological entities that are investigated, such as FES and in order to overcome as much as possible this limitation the leave-one-out cross validation procedure was used.

Further research is currently carried out in order to incorporate other psychiatric disorders in the system, for example drug addicts and patients suffering from Obsessive Compulsive Disorder (OCD). The proposed approach could, in principle, be applied for more complex biosignal classification tasks, e.g. for the case of individuals that could be affected by several neuropsychiatric conditions, without the prior diagnosis reducing the problem to a dichotomy. This would require, for example, the construction of an ANN with as many output layer neurons as the neuropsychiatric conditions for which the system would be designed to handle, plus one for the normal subjects class, and a very extended subject set for training and testing the network. The set should include all the possible classes of subjects, i.e. classes representing only one condition and classes representing possible comorbidities. This is a daunting task, especially concerning the collection of recordings from a significant number of matched groups of subjects, with and without comorbidities. Furthermore, as the classes to be differentiated by the system are augmented, it remains to be verified whether the MVAR model will continue to fit the various classes in a manner useful for classification purposes. As a first tentative step in this direction, ongoing work in our laboratories aims at extending the methodology proposed in the current study to pairs of neuropsychiatric conditions. Another important aspect of ongoing research concerns the comparative evaluation of additional ERP inversion techniques that have been proposed in the literature, in order to compare the classification performance provided by the various parameters modeling brain electrical phenomena, according to each inversion technique.

7 Acknowledgements

The authors would like to thank the Department of Psychiatry at the Eginition Hospital, University of Athens, Greece, for the collaboration regarding the data for testing the proposed classification method.

References

[1] Johnson, Jr R, (Ed), (1995) *Event-Related Brain Potentials and Cognition*, In Handbook of Neuropsychology, 10(14), Elsevier, Amsterdam, The Netherlands.

[2] Sveinsson, J. R., Benediktsson, J. A., Stefansson, S. B. and K., D. (1997), *Parallel principal component neural networks for classification of event-related potential waveforms*, Med Eng Phys, **19,** 15-20.

[3] Palaniappan, R. and Paramesran, R. (2002), *Using genetic algorithm to identify the discriminatory subset of multi-channel spectral bands for visual response*, Appl Soft Comput**, 2,** 48-60.

[4] Kalatzis, I., Piliouras, N., Ventouras, E., Papageorgiou, C. C., Rabavilas, A. D. and Cavouras, D. (2004), *Design and Implementation of an SVM-based Computer Classification System for*

Discriminating Depressive Patients from Healthy Controls using the P600 Component of ERP Signals, Comput Meth Prog Bio, **75,** 11-22.

[5] Wang, Y., Zhang, Z., Li, Y., Gao, X., Gao, S. and Yang, F. (2004), *BCI Competition 2003—Data Set IV: An Algorithm Based on CSSD and FDA for Classifying Single-Trial EEG*, IEEE T Biomed Eng, **51,** 1081-86.

[6] Anderson, C., Stolz, E. and Shamsunder, S. (1998), *Multivariate Autoregressive Models for Classification of Spontaneous Electroencephalogram During Mental Tasks*, IEEE T Biomed Eng, 45, 277-86.

[7] Scherg, M., Bast, T. and Berg, P. (1999), *Multiple source analysis of interictal spikes: goals, requirements, and clinical value*, J Clin Neurophysiol, 16, 214-24.

[8] Koles, Z. J. (1998), *Trends in EEG source localization*, Electroen Clin Neuro, 106, 127-37.

[9] Schenker, B. and Agarwal, M. (1996), *Cross-validated structure selection for neural networks*, Comput Chem Eng, 20, 175-86.

[10] Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D. and Halgren, E. (2000), *Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity*, Neuron, 26, 55-67.

[11] Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., Grave de Peralta, R., Gonzalez, S., Seeck, M. and Landis, T. (2001), *Electric source imaging of human brain functions*, Brain Res Rev, 36, 108-18.

[12] Musha, T., Asada, T., Yamashita, F., Kinoshita, T., Chen, Z., Matsuda, H., Uno, M. and Shankle, W. R. (2002), *A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage Alzheimer's disease*, Clin Neurophysiol, 113, 1052-8.

[13] Dierks, T., Strik, W. K. and Maurer, K. (1995), *Electrical brain activity in schizophrenia described by equivalent dipoles of FFT-data*, Schizophr Res, 14, 145-54.

[14] Pizzagalli, D., Koenig, T., Regard, M. and Lehmann, D. (1999), *Affective attitudes to face images associated with intracerebral EEG source location before face viewing*, Cognitive Brain Res, 7, 371-7.

[15] Huang, C., Wahlund, L., Dierks, T., Julin, P., Winblad, B. and Jelic, V. (2000), *Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study*, Clin Neurophysiol, 111, 1961-7.

[16] Kirkpatrick, S. G., C. D.; and Vecchi, M. P. (1983), *Optimization by Simulated Annealing.*, Science, 220, 671-680.

[17] Priestley, M. B. (1981) Spectral Analysis and Time Series, Academic Press.

[18] Kay, S. M. (1987) Modern Spectral Estimation: Theory and Application, Prentice-Hall, Englewood Cliffs, NJ.

[19] Ingber, L. (1993), *Simulated Annealing: Practice Versus Theory*, Math Comput Model, 18, 29-57.

[20] Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992) Numerical Recipes in C: the Art of Scientific Computing, Cambridge University Press, Cambridge.

[21] Rezaee, M. R., Nyqvist, C., Zwet, P. M. J., Jansen, E. and Reiber, J. H. C. (1995), *Segmentation of MR images by a fuzzy c-mean algorithm*, Computers in Cardiology, 21-24.

[22] Arifovic, J. and Gencay, R. (2001), *Using Genetic Algorithm to Select Architecture of a Feedforward Neural Network*, Physica A, 289, 574-94.

[23] Baumgart-Schmitt, R., Herrmann, W. M., Eilers, R. and Bes, F. (1997), *On the use of neural network techniques to analyse sleep EEG data. First communication: application of evolutionary and genetic algorithms to reduce the feature space and to develop classification rules*, Neuropsychobiol, 36, 194-210.

[24] Reed, R. (1993), *Pruning algorithms - a survey*, IEEE T Neural Networ, 4, 740-47.

[25] Jerez-Aragones, J. M., Gomez-Ruiz, J. A., Ramos-Jimenez, G., Munoz-Perez, J. and Alba-Conejo, E. (2003), *A combined neural network and decision trees model for prognosis of breast cancer relapse*, Artif Intell Med, 27, 45-63.

[26] Chen, Y., Kopp, G. A. and Surry, D. (2002), *Interpolation of wind-induced pressure time series with an artificial neural network*, J Wind Eng Ind Aerod, 90, 589-615.

[27] McGuffin, P., Farmer, A. and Harvey, I. (1991), *A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system*, Arch Gen Psychiatry, 48, 764-70.

[28] Wechsler, D. (1955) Wechsler Adult Intelligence Scale, Psychological Corporation, New York.

[29] Papageorgiou, C., Liappas, I., Asvestas, P., Vasios, C., Matsopoulos, G. K., Nikolaou, C., Nikita, K. S., Uzunoglu, N. and Rabavilas, A. (2001b), *Abnormal P600 in heroin addicts with prolonged abstinence elicited during a working memory test*, Neuroreport, 12, 1773-8.

[30] Jasper, H. H. (1958), *The ten twenty electrode system of the International Federation*, Electroen Clin Neuro, 10, 371-375.

[31] Papageorgiou, C., Kontaxakis, V. P., Havaki-Kontaxaki, B. J., Stamouli, S., Vasios, C., Asvestas, P., Matsopoulos, G. K., Kontopantelis, E., Rabavilas, A., Uzunoglu, N. and Christodoulou, G. N. (2001a), *Impaired P600 in neuroleptic naive patients with first-episode schizophrenia*, Neuroreport, 12, 2801-6.

[32] Friederici, A. D., von Cramon, D. Y. and Kotz, S. A. (1999), *Language related brain potentials in patients with cortical and subcortical left hemisphere lesions*, Brain, 122 (Pt 6), 1033-47.

[33] Uzunoglu, N., Ventouras, E., Papageorgiou, E., Rabavilas, A. and Stefanis, C. (1991), *Inversion of Simulated Evoked Potentials to Charge Distribution Inside the Human Brain Using an Algebraic Reconstruction Technique*, IEEE T Med Imaging, 10, 479-84.

[34] Ventouras, E., Papageorgiou, C., Rabavilas, A., Uzunoglu, N. and Stefanis, C. (1999), *Brain Electrical Tomography Using Algebraic Reconstruction Techniques*, Biotechnol Biotechnol Eq, 13, 113-118.

[35] Pascual-Marqui, R. D., Michel, C. M. and Lehmann, D. (1994), *Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain*, Int J Psychophysiol, 18, 49-65.

[36] Pascual-Marqui, R. D. (1999), *Review of Methods for Solving the EEG Inverse Problem*, Int J Bioelectromag, 1, 75-86.

[37] Talairach, J. and Tournoux, P. (1988) Co-planar stereotaxic atlas of the human brain, Thieme, New York.

[38] Geddes, L.A. and Baker, L.E. (1967), *The specific resistance of biological material - a compendium of data for the biomedical engineer and physiologist,* Med Biol Eng, 5, 271–93.

[39] Ventouras, E. (1994), *Development of a Tomographic Imaging Technique and Simulation Using Artificial Neural Networks of the Intracranial Sources of Event-Related Electrical Activity*, Ph.D. Thesis, National Technical University of Athens, Athens, Greece.

[40] Ventouras, E., Uzunoglu, N., Koutsouris, D., Papageorgiou, C., Rabavilas, A. and Stefanis, C. (1995), *Intracranial Electrical Current Source Mapping Derived from Event-Related Potentials Data*, Innov Tech Biol Med, 16, 27-41.

[41] Fuchs, M., Wagner, M., Wischmann, H.-A., Kohler, T., Theißen, A., Drenckhahn, R., Buchner, H. (1998), *Improving source reconstructions by combining bioelectric and biomagnetic data*, Electroen Clin Neuro, **107**, 93–111.

[42] Mosher, J.C., Baillet, S. and Leahy, R.M. (1991), *EEG Source Localization and Imaging Using Multiple Signal Classification Approaches*, J Clin Neurophysiol, 16, 225-38.

[43] Haykin, S. (1994) Neural Networks, MacMillan, New York.

[44] Tekin, S. and Cummings, J. L. (2002), *Frontal-subcortical neuronal circuits and clinical neuropsychiatry: an update*, J Psychosom Res, 53, 647-54.

[45] Guillem, F., Bicu, M., Pampoulova, T., Hooper, R., Bloom, D., Wolf, M.-A., Messier, J., Desautels, R., Todorov, C., Lalonde, P., Debruillea, J.B. (2003), *The cognitive and anatomo-functional basis of reality distortion in schizophrenia: A view from memory event-related potentials*, Psychiatry Res, 117, 137–58.

[46] Falkenstein, M., Hohnsbein, J., Hoormann, J. (1994), *Effects of choice complexity on different subcomponents of the late positive complex of the event-related potential*, Electroen Clin Neuro, 92, 148-60.