Epileptic Seizure Prediction Using Hybrid Feature Selection Over Multiple Intracranial EEG Electrode Contacts: A Report of Four Patients

Maryann D'Alessandro, *Member, IEEE*, Rosana Esteller, *Member, IEEE*, George Vachtsevanos, *Senior Member, IEEE*, Arthur Hinson, *Member, IEEE*, Javier Echauz, *Member, IEEE*, and Brian Litt, *Member, IEEE*

***Abstract*—Epileptic seizure prediction has steadily evolved from its conception in the 1970s, to proof-of-principle experiments in the late 1980s and 1990s, to its current place as an area of vigorous, clinical and laboratory investigation. As a step toward practical implementation of this technology in humans, we present an individualized method for selecting electroencephalogram (EEG) features and electrode locations for seizure prediction focused on precursors that occur within ten minutes of electrographic seizure onset. This method applies an intelligent genetic search process to EEG signals simultaneously collected from multiple intracranial electrode contacts and multiple quantitative features derived from these signals. The algorithm is trained on a series of baseline and preseizure records and then validated on other, previously unseen data using split sample validation techniques. The performance of this method is demonstrated on multiday recordings obtained from four patients implanted with intracranial electrodes during evaluation for epilepsy surgery. An average probability of prediction (or block sensitivity) of 62.5% was achieved in this group, with an average block false positive (FP) rate of 0.2775 FP predictions/h, corresponding to 90.47% specificity. These findings are presented as an example of a method for training, testing and validating a seizure prediction system on data from individual patients. Given the heterogeneity of epilepsy, it is likely that methods of this type will be required to configure intelligent devices for treating epilepsy to each individual's neurophysiology prior to clinical deployment.**

***Index Terms*—Epileptic seizure prediction, feature selection, genetic algorithms, multiple channels and features.**

Manuscript received April 16, 2002; revised November 8, 2002. This work was supported in part by the Whitaker Foundation, in part by the Epilepsy Foundation, in part by the American Epilepsy Society, in part by the University of Pennsylvania Research Foundation, and in part by the National Institutes of Health (NIH) under Grant R01NS041811-01 and Grant MH-62298RO1. *Asterisk indicates corresponding author.*

*M. D'Alessandro was with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA. She is now with the Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: maryann0625@hotmail.com).

R. Esteller and J. Echauz are with NeuroPace, Inc., Mountainview, CA 94043 USA.

G. Vachtsevanos is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

A. Hinson is with the Naval Surface Warfare Center, Dahlgren, VA 22448 USA.

B. Litt is with the Departments of Bioengineering and Neurology, University of Pennsylvania and the Hospital of the University of Pennsylvania, Philadelphia, PA 19104 USA.

Digital Object Identifier 10.1109/TBME.2003.810706

I. Introduction

In humans, epilepsy is the second most common neurological disorder, next to stroke, affecting 50 million people worldwide. Of these individuals, 25% do not respond to available therapies [1]. There is currently an explosion of interest in predicting epileptic seizures from intracranial electroencephalogram (IEEG) that has its roots in experimental and theoretical work first published in the 1970s. Despite over 40 years of investigation into the physiology of epilepsy, it still is not possible to explain how and over what time spontaneous clinical seizures emerge from the relatively normal brain state observed between them [2].

An exciting application of seizure prediction technology is its potential for use in therapeutic epilepsy devices to trigger intervention to prevent seizures before they begin. Since the early 1970s, approximately 22 patents addressing epileptic seizures have been filed. [3]–[18]. Most are open loop or triggered open loop systems addressing stimulation for preventing clinical seizures after seizure detection, while seven claim seizure prediction. To date, the vagus nerve stimulator Neurocybernetic Prosthesis is the only device that has been FDA approved for implantation to treat partial onset seizures. This device is an open loop system that operates blindly based on the amplitude and duration of stimulation set by the medical doctor [19]. More intelligent "closed loop" devices based upon seizure detection or prediction are under development, as are the appropriate algorithms to drive them. Fig. 1 provides bar graphs depicting the number of patents related to epilepsy and seizure prediction publications since the early 1970s addressing seizure prediction.

Seizure prediction has been investigated by type to include prediction by studying preictal features, prediction by fast detection, prediction by classification, and prediction by probability estimation. Only research involving multiple channels and multiple features will be identified here, since a synergy of multiple channels and multiple feature prediction is the heart of this work. The technical community's exponentially growing interest in epilepsy applications is clear. One reason for this surge in interest over the past 5 years is the realization that epilepsy surgery is unlikely to benefit more than a minority of patients who cannot be adequately treated by medication, and that progress in engineering and neuroscience is making therapeutic epilepsy devices more realizable.

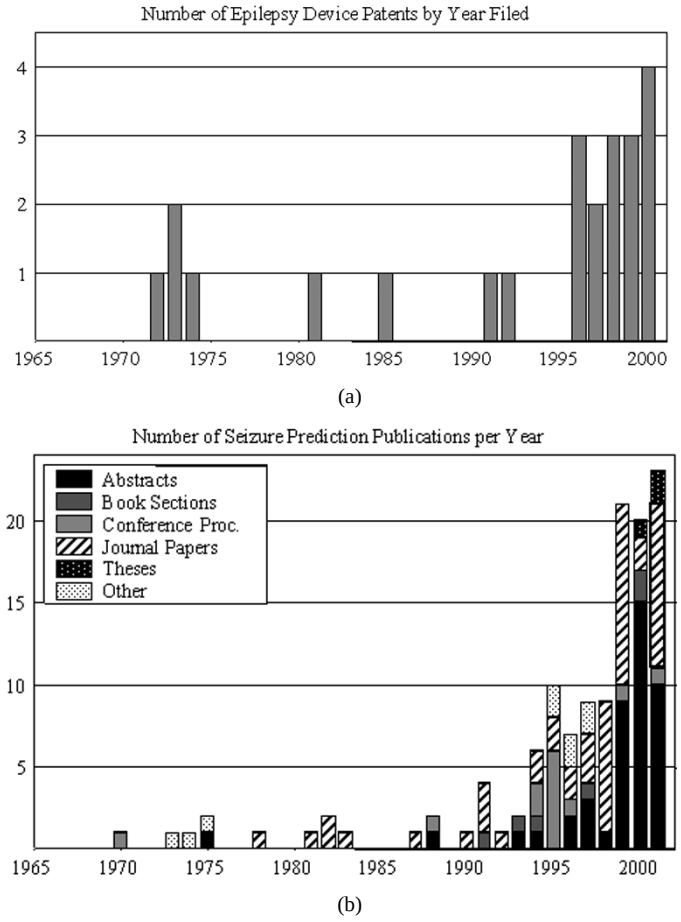


Fig. 1. (a) Bar graph of patent literature addressing epileptic seizures. (b) Bar graph of journal publications.

Studies in seizure prediction vary widely in their theoretical approaches to the problem, validation of results, and the amount of data analyzed. Some relative weaknesses in this literature are the lack of extensive testing on baseline data free from seizures, the lack of technically rigorous validation and quantification of algorithm performance in many studies, and the absence of methods for distilling multichannel and multifeature data into its most useful components for prediction. In the study presented, preictal data segments and 40 h of baseline data were reserved for testing to demonstrate the feasibility of the approach. Since the research is limited ethically by the requirements for presurgical analysis, it is not possible to test the resultant feature on data gathered months or years later from the same subject. As new developments are pushing forward in this field, it is expected that in the future, the availability of implantable devices will allow prospective studies on seizure prediction where an individualized feature set can be validated over months or years on the same subject.

Some studies, such as those of Petrosian, report seizure prediction after analyzing one channel of electroencephalogram (EEG) from an intracranial depth electrode in one patient [20]–[22]. In these studies, using univariate techniques no analysis of baseline data far removed from the seizure was undertaken. A potential pitfall of conclusions based upon such limited data is that quantitative changes identified prior to seizure onset may not be specific to the preseizure period, but may occur at other times as well, unrelated to epileptic events. Validation of prediction algorithms on long, continuous sets of clinical data, representing all states of awareness, is an important part of more recent seizure prediction studies.

A number of promising quantitative features derived from the EEG, each with different theoretical bases, have demonstrated utility for seizure prediction. Iasemidis and Sackellares were the first group to apply nonlinear dynamical techniques, particularly methods based upon the principal Lyapunov exponent (PLE), for predicting seizures beginning in the late 1980s. This group has demonstrated evidence of seizure precursors in a variety of data sets, ranging from one to multiple channels and epochs spanning minutes to hours [23]–[27]. In their research, seizure onset prediction was reported from one to sixty minutes prior to electrographic seizure onset in some data sets. Since the mid 1990s, Lehnertz and Elger *et al.* have expanded work in nonlinear dynamics and seizure prediction to larger data sets, greater numbers of patients, and a variety of epilepsy types, utilizing parameters based upon the correlation dimension. They report distinguishable preseizure patterns from 4.25 – 25 minutes prior to seizure onset [28]–[32] in some studies and evidence of what is likely to be preseizure synchronization of activity up to hours before seizure onset [33]. More recently, Le Van Quyen and Baulac *et al.* have anticipated seizures using a nonlinear parameter they call dynamical similarity [34]–[36]. Starting with just a few patients, this group has steadily expanded the number of patients they have analyzed and the length of data for each patient. Approximate prediction times agree with those of the previous two groups, suggesting that they may be observing similar or the same physiological phenomena in different ways.

Using an approach based upon physiologic changes in the EEG usually associated with epilepsy, such as spikes, slowing, subclinical seizures and increased energy in the signal, our research team postulated that seizure generation is composed of a cascade of EEG events that develop over hours in temporal lobe epilepsy. This conclusion was based upon findings that accumulated signal energy, the rate of subclinical seizure-like bursts and bursts of long-term energy all increased as seizures approached [2]. In addition, we found patient-specific seizure onset and pre-seizure patterns, suggesting that algorithms tailored to individual patients may offer some advantage for seizure prediction. Again, prediction horizons were similar to those described using nonlinear techniques, giving further credible evidence that quantitative measures that appear to anticipate EEG onset of seizures are unlikely to be artifact.

The successful “proof-of-principle” demonstration of a preseizure period, validated by multiple research groups and techniques has moved research in this area on to more in-depth studies exploring the temporal and spatial characteristics of the preseizure period, as well as its underlying mechanisms. Recent collaborative discussions between groups working in this area have identified several important related areas of investigation. First, in order to be able to assess and compare seizure prediction methods, acceptable performance metrics for these types of studies must be agreed upon. In a recent published work on seizure prediction [37], investigators in the field cautioned against overstating performance results of seizure pre-

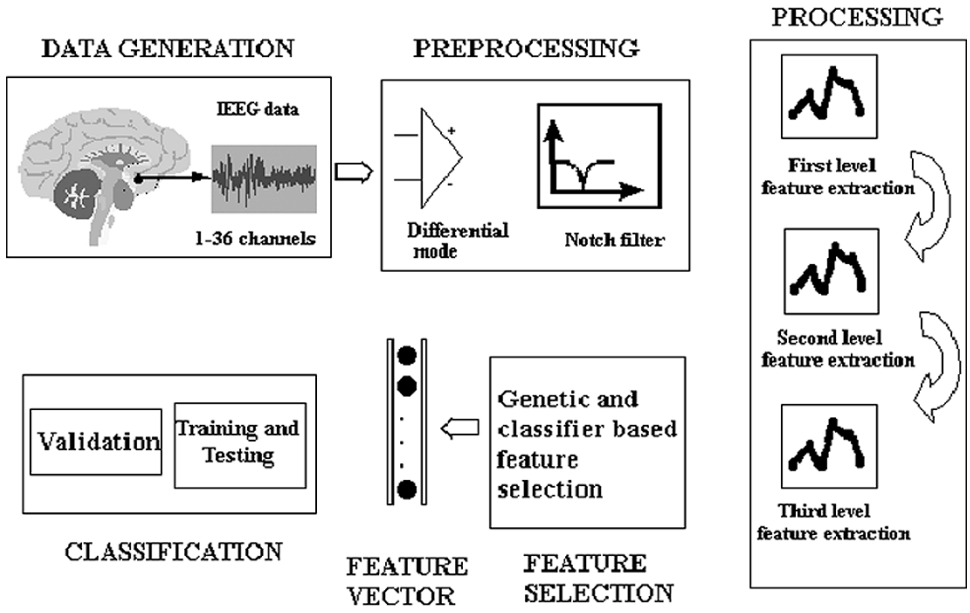


Fig. 2. Schematic drawing of the hybrid genetic and classifier based feature selection process for epileptic seizure prediction.

diction algorithms, and advised that rigorous reporting of sensitivity and specificity, among other performance criteria, are necessary to validate the utility of different experimental approaches. Second, the need for analyzing long epochs of data free from seizures was emphasized. Without this work, eventually leading to prospective, on-line trials of prediction algorithms, the specificity of seizure precursors to the preictal period will always remain suspect. Finally, great interest has been stimulated by findings that seizure precursors appear to spread spatially in the temporal lobe over time, even crossing to "entrain" the contralateral temporal region as temporal lobe seizures approach. These factors, as well as the need for a method to tailor algorithms to specific patients to achieve more optimized results prompted this study, in which we demonstrate that patient specific combinations of multiple channels and multiple quantitative features can predict seizures with reasonable accuracy after supervised training on appropriate preseizure, seizure and baseline data.

II. Methods

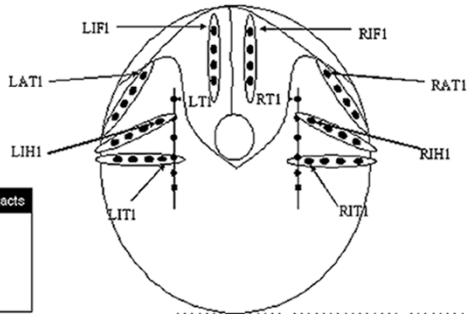
The objective of this study was to conduct a preliminary evaluation based on quantitative EEG analysis to determine if a genetic search process is capable of identifying appropriate patient specific features and electrode sites for predicting seizures from a large set of candidate features and intracranial electrode recording sites. Numerous candidate features revealed in the literature were considered, including those from time domain analysis, frequency domain analysis, and nonlinear dynamics. From these features, a subset was selected as potentially useful for seizure prediction. The transformation of potentially relevant features was approached heuristically, based on previous research, expert visual assessment, and analysis of processed signals. The genetic search process was applied to a portion of available preictal and baseline data, while the remaining preictal records and 40 h of consecutive baseline data were used to validate this approach. The features and electrode sites selected were not meant to provide the best features for seizure prediction, but to provide a potential solution on a patient-specific basis to the problem of predicting seizures in the minutes prior to their onset on the IEEG (a 10-min "prediction horizon"). This prediction horizon was chosen as a reasonable beginning preseizure interval that would allow time for preemptive intervention and straightforward validation. We also chose to study quantitative IEEG features that could be related to neurophysiology, and potentially the mechanisms underlying seizure generation.

To investigate whether a hybrid genetic- and classifier-based feature selection process could reveal features capable of predicting epileptic seizures, after data were collected; we implemented the steps shown in Fig. 2.

A. Data Generation

1) Subjects: Patients ranged in age from 21–53, had suspected mesial temporal lobe epilepsy, were implanted with bilateral amygdalo-hippocampal depth electrodes, subdural strip electrodes, and were hospitalized for 3- to 14-days for continuous video-EEG monitoring during evaluation for epilepsy surgery. Most of the patients whose data were analyzed were tapered off of their antiepileptic drugs during the hospital stay to induce seizures. Some or all of the intracranial electrode contacts identified in Fig. 3 were monitored for each patient. Patients were selected consecutively from a database of patients whose clinical and neurophysiological information were stored for analysis.

2) Recording Procedures: We collected data on a standard Nicolet 5000 video EEG acquisition system utilizing a 12-bit analog-to-digital converter and sampled at a rate of 200 Hz with bandpass filter settings of 0.1–100 Hz. Synchronization of video and EEG was achieved and stored for offline analysis of clinical onsets on SVHS video tapes. EEG data were downloaded to



Patient	Electrode contacts
A	36
B	20
C	28
D	28

Fig. 3. Intracranial electrode contacts monitored for each patient. LT: Left temporal lobe; LIF: left inferior frontal lobe; LAT: left anterior temporal lobe; RT: right temporal lobe; RIF: right inferior frontal lobe; RAT: right anterior temporal lobe; RIH: right inferior hippocampal formation.

compact disk for quantitative analysis. Sleep-wake cycles, medication administration, and pertinent patient behaviors were reviewed and catalogued for each patient over the entire hospital stay.

3) Data Selection: We analyzed four patients, comprising 46 preictal records, and 160 h of baseline data. An average of 70% of all available preictal lead seizures were used for training, while testing included both the remaining 30% of available preictal records and artificially generated preseizure records. Lead seizures were defined as seizures that were separated by at least 4 h from the end or beginning of another seizure. This distinction was made based upon evidence that seizures originating in the temporal lobe may influence the IEEG within this period of time [2]. Approximately 4 h of baseline data/patient were used for training, while 40 h/patient were used for testing. We clipped 10-min data epochs before each seizure onset from all IEEG channels from the original raw data during training to address the short 10-min prediction horizon. Data epochs were considered "baselines" if they were spaced at least 4 h away from the onset or ending of an electrographic seizure.

4) Artificially Generated Preseizure Records: Baseline data comprised the overwhelming majority of data in this research. In the available data set, there were an adequate number of records for each patient for training and testing. However, the number of preictal samples available for each patient ranged between two and four records for training and two or three records for testing. Artificially generated surrogate data sets were obtained to increase the available test set for validation. The use of surrogate data is useful to validate results obtained from experimental data when no theoretical basis is available.

A new form of jittering based on a uniform distribution with an adaptive range was used to obtain the preictal surrogates. Surrogate samples were obtained by randomly selecting data from samples in the test set while maintaining the time index of the preictal recording. Surrogate samples were chosen randomly with replacement and were obtained from the uniform distribution of numbers between the range of values in the available samples. Samples were randomly selected among data from the same awareness state. Using this technique, the probability density function (PDF) was adaptive, depending on the data.

5) Terminology: In this research, an expert epileptologist visually reviewed the EEG for each seizure before quantitative analyses of these data were conducted, and marked each seizure at the following times and identified the corresponding locations: 1) the *focus channel*: the spatial location of the earliest seizure onset on IEEG, or if multiple channels were involved at onset, the focus channel was the electrode measuring maximal onset; 2) the unequivocal electrographic onset (UEO): this marked the time at which an EEG pattern typically associated with seizures first became unquestionably clear and could be identified independent of knowing a seizure followed; and 3) the earliest electrographic change (EEC): this point in time was found by identifying unequivocal electrical seizure onset on the EEG and then moving backward in time to the point at which the first clear, sustained change from the patient's EEG baseline prior to the seizure was detected. The unequivocal clinical onset was marked when visual evidence indicating the presence of a seizure became evident. In this research, the clinical onset was identified by the expert epileptologist visually reviewing the video tapes after the data were gathered.

B. Preprocessing

To minimize the common mode artifact while maintaining the integrity of the signal, data were analyzed in a "bipolar montage," in which the signals from spatially adjacent channels were subtracted to obtain the differential mode signal. After removing the common mode artifact, a 60-Hz digital notch filter was used to remove line noise.

C. Processing

Processing utilized a three-step approach that included extracting first-level features from the raw data, extracting second-level features from first-level features, and extracting third-level features from second-level features. Candidate features were selected based on criteria described in [38], expertise, observations, and our understanding of EEG signal characteristics. To evaluate most features, it is important to maintain stationarity of the data segment. Statistical tests reveal quasi-stationarity of the EEG signal anywhere from 1 s (200 points) to several minutes [51]. Because seizures spread so quickly, a displacement as small as possible that does not provide too much variability is desired. We experimented with values ranging from 0.25 s to 5 s and observed that a displacement of 500 points (2.5 s) and the window length to 2000 points (10 s) provided reasonable propagation resolution of seizure precursors and the ability of multichannel analysis to effect prediction. These values were used for the first-level feature extraction for all tests and agree with the definition of stationarity found in the literature and preliminary prediction results. For the second- and third-level feature extraction a 24 point (1 min) window and a displacement of 1 point (2.5 s) were used.

Numerous features revealed in the literature were considered from time domain analysis, frequency domain analysis, and nonlinear dynamics. From those features considered, a set of features was selected as potentially useful for seizure prediction that had computational requirements reasonable for real-time implementation. While this list of features was not considered to be exhaustive, it was chosen to be sufficient for proof of principle for the multifeature method. This method provides a structure based upon multiple quantitative features

and multiple electrode sites that can be used to optimize seizure prediction for a particular group of features. As noted above, the transformation of potentially relevant features was approached heuristically, based on previous research, expert visual assessment and analysis of processed signals.

Selected features were also chosen to be computationally efficient and have potential for on-line implementation in low-power, implantable environments. The six selected first-level quantitative features derived from the intracranial EEG for this analysis are described below.

1) First-Level Features:

a) Curve length: An in depth theoretical fractal dimension (FD) analysis was conducted by Esteller, yielding the fractal dimension algorithm proposed by Katz as the most promising for detecting unequivocal seizure onset on IEEG, and in some cases even predicting it [39]. FD meets the computational requirements for real time implementation; however, the computational burden can be further reduced by eliminating the logarithmic computations and analyzing the curve length of the signal defined by the sum of the lengths of the line segments between successive samples of a signal. This feature was originally introduced by Olsen in [11] as the "line length" prior to being described as the "curve length" in [4]. The mathematical representation of the curve length in its discrete form is

$$CL\,[n] = \sum_{ik=1+(n-1)(N-D)}^{n(N-D)+D} |x(i-1) - x(i)| \qquad (1)$$

where $CL[n]$ is the running curve length of the time series $x(n)$, N is the length of the sliding observation window expressed in number of points, n is the discrete time index, and D is the overlap. The curve length provides results nearly equivalent to the Katz FD algorithm, while providing less computational burden. The curve length is useful for observing amplitude and frequency changes and dimensionality of the signal. In addition, the fractal dimension adds nonlinearities and may yield negative numbers not found in the curve length [40].

b) Energy: The accumulated energy (AE) provided promising results for seizure prediction in all the patients analyzed in [41] and in [2] using the focus channel. However, unless AE is converted to a resetting AE, where it is initialized blind to the seizure onset, it is not a practical feature for an online pattern recognition system. Consequently, energy is considered as a first-level feature and subsequent feature levels are expected to provide predictive preseizure indicators. Let the sequence $x(n)$ be a preprocessed and fused input signal, then the instantaneous energy of $x(n)$ is given by $x^2(n)$. Considering that a sliding window is used, the energy of the signal becomes the average power over the window mathematically defined as

$$E[n] = \frac{1}{N} \sum_{i=(n-1)N+1}^{nN} x(i)^2 \qquad (2)$$

where

- N size of the sliding window expressed in number of points;
- n set $1, 2, 3, \ldots$ (the discrete time index).

The moving average of the energy defined above is with zero overlap. If an overlap of D points is allowed, then the average energy becomes

$$E_D[n] = \frac{1}{N} \sum_{i=1+(n-1)(N-D)}^{n(N-D)+D} x(i)^2 \qquad (3)$$

where: E_D is the average energy or moving average of the power with D points of overlap. An overlap of 1500 points was used in this paper at the first level of feature extraction while an overlap of 23 points was used at levels two and three.

c) Nonlinear Energy: Esteller [41], Zaveri [42], and Gotman [43] all used Teager's algorithm to analyze EEG signals: Esteller with the intent to detect the seizure onset; Zaveri to observe how the seizure propagates after the ictal onset; and Gotman and Agarwal to provide indicators as to the spectral content of the signal. This algorithm was presented by Kaiser who was searching for a measure of energy proportional to both signal amplitude and frequency [44]. This algorithm is also recognized as "nonlinear energy" because it has also shown to be useful for determining the instantaneous frequency and amplitude envelope of AM-FM signals [45].

Zaveri used both Teager's algorithm and the classical definition of energy to investigate propagation. He found Teager's algorithm to provide more favorable results. Gotman and Agarwal found the nonlinear energy operator useful for providing an indication as to the spectral content of the signal since it is sensitive to spectral changes. For the input signal $x(n)$, in its discrete form, the nonlinear energy (NE) operator is represented by

$$\mathrm{NE}\,[n] = x^2\,(n) - x\,(n-1)\,x\,(n+1)\,. \qquad (4)$$

The NE is an instantaneous feature, such that it provides one value for each value of original data. After the NL is obtained, the feature is weighted with a Hanning window; then the mean of the windowed data, $NE_w[n]$, is taken over the desired sliding window. After windowing, the average nonlinear energy is then

$$\mathrm{ANE} = \frac{1}{N} \sum_{n=1+(k-1)(N-D)}^{k(N-D)+D} NE_w[n] \qquad (5)$$

where:

- ANE$[k]$ average NE at time k;
- N desired window length;
- D overlap in number of points;
- k discrete time index equal to $1, 2, 3, \ldots$

The algorithm is sensitive to both amplitude and frequency changes, and is computationally efficient, and simple to calculate.

d) Spectral Entropy (SE): The main objective of entropic feature analysis is to quantify regularity and order in the signal. Researchers focusing on prediction generally limit their analysis to the bipolar focus channel since this channel appears to provide the most evidence regarding the seizure. However, a preliminary analysis of the SE feature indicates that the focus channel may present a constant exhibition of abnormal activity, thereby compromising the analysis. This finding, presented in [46], has prompted the investigation of all channels for prediction indicators. This phenomenon also was observed in the chaotic domain as measured by the PLE by Iasemidis *et al.* [47].

This group found that the ipsilateral side of the brain carries different chaotic properties than the side contralateral to the focus.

In 1979, Powell and Percival introduced the concept of SE, based on the peaks of the Fourier spectrum, as a measure of regularity [48]. In 1991, Inouye *et al.* applied SE to surface EEG to compare the surface EEG signals during rest with those during mental arithmetic to analyze the degree of EEG regularity. Since the spectrum of irregular EEG tends to be 1/f while the spectrum of rhythmic EEG has a peak in the power spectrum, Inouye claimed that the SE is a useful means for observing the degree of EEG irregularity [49]. The information theory concept of entropy introduced by the late Claude Shannon in 1948 is represented by

$$H(x) = E_x\left\{i(x)\right\} \tag{6}$$

and

$$E_x\left\{i(x)\right\} = -\sum_{x\in X} p(x)\log_2 p(x) \tag{7}$$

where E_x is the expectation with respect to x, $i(x)$ is the self-information, and $p(x)$ is the discrete probability distribution function (PDF). To find the SE, first the discrete-time form of the spectrum is found from

$$P(k) = \frac{1}{NT}\left|TX(k)\right|^2 \tag{8}$$

and

$$X\left(k\right) = \sum_{n=0}^{N-1} x\left[n\right]\exp\left(\frac{-j2\pi kn}{N}\right) \tag{9}$$

where $X(k)$ is the discrete Fourier transform, T is the sampling interval, and N is the length of the sliding observation window. The SE then is expressed as

$$\mathrm{SE}(m) = -\sum_{m=1+(k-1)(N-D)}^{m(N-D)+D} P\left(k\right)\log_2 P\left(k\right) \tag{10}$$

where $SE(m)$ is the running SE of the time series $x[n]$, N is the length of the sliding observation window, and m is the discrete time index. This analysis complements visual EEG analysis by an expert epileptologist and may provide further insight into the underlying mechanisms of ictogenesis.

e) Sixth Power: The sixth power indicator calculates the sixth power of the signal and we have found it to be empirically useful for observing small amplitude differences in the IEEG, though no comparison with other exponents has been undertaken. The sixth power was selected over lower powers because it more effectively emphasized the amplitude changes when compared with the lower powers and this was confirmed over a preliminary evaluation using a sample of data from a subset of patients in the database. Magnitude differences are increased with the power of the signal, since small signals increase less than signals with larger amplitudes. The sixth power is the sixth power of each data point and is expressed as

$$SP_D[n] = \frac{1}{N}\sum_{i=1+(n-1)(N-D)}^{n(N-D)+D} x(i)^6 \tag{11}$$

where $SP(n)$ is the running sixth power of the time series $x(n)$, N is the length of the sliding observation window, and n is the discrete time index.

f) Energy of the Wavelet Packets: In Esteller's work, the "absolute value of the fourth Daubechies (daub4) wavelet coefficient averaged" provided promising results for detecting UEO [41]. The decision to use the daub4 wavelet as a potential feature was made by conducting a visual screening of sample IEEG signals to determine which mother wavelet would provide the best separation between classes. In another experiment [50], Petrosian selected the daub4 as the mother wavelet for his study on predicting seizures using a recurrent neural network and the high-pass and low-pass decompositions of the daub4 wavelet because it has good localizing properties in the time and frequency domains.

Using the daub4 wavelet, there are 2^n wavelet packets and 2^{2n-1} combinations. We use a five-level decomposition, yielding 32 wavelet packets. The wavelet packet decomposition should be a useful feature since the EEG is a combination of the numerous processes occurring in the brain; the results of the individual processes often are lost in the sum of all combinations. The wavelet packet decomposition will decompose the signal into 32 frequency bands between 0 and 100 Hz. Splitting the signal into the separate bands should reveal hidden details not evident in the original signal. The GA evaluates each wavelet packet separately, and the best packet is selected from the 32 resultant outputs.

2) Second- and Third-Level Features: The second-level features were selected after visually observing the focus channel of a random sample of first-level feature plots and class conditional PDFs representing seizures from five patients. The visual analysis identified characteristics of the 10 min prior to UEO that, if extracted from the first-level features, may provide prediction indicators. The second- and third-level features found useful from this screening include:

•minimum	•maximum	•median
•mean	•variance	•std. deviation
•skewness	•kurtosis	•slope
•integral	•derivative	•sum.

All seizures and all patients did not necessarily exhibit potential prediction indicators identified above; however, all potential second-level features are considered in this research for developing the objective feature vector.

D. Feature Selection

1) Genetic Feature Selection: A hybridization of genetic and classifier based feature selection was employed in this paper to address multiple feature and multiple channel analysis. The first phase of the approach reduced the possible 25 872 features to one feature in each of the six first-level feature families. A "family" is defined here as any feature transformation derived from that particular first-level feature. The feed forward approach then was applied using the best feature in each family.

Genetic algorithms (GAs) are smart search processes inspired by biological evolution [52]. Precisely, each possible solution is represented by a coded string of bits or a "chromosome." The GA chromosome used in this paper is shown in Fig. 4. A novel

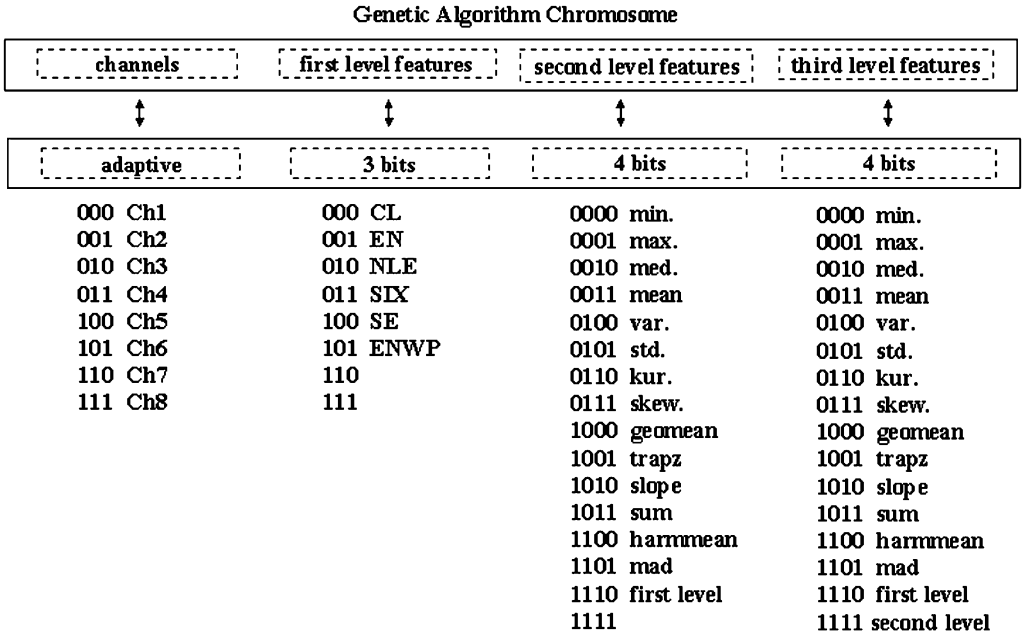


Fig. 4. GA chromosome representing the binary representation used to select electrode contacts and features for evaluation.

adaptive chromosome is applied to identify reasonable features to serve as inputs to the classifier based selection process. Each chromosome is comprised of a string of bits or "genes" whose content is called an "allele." There are four "genes" in this paper: 1) electrode contacts or channels; 2) first-level features; 3) second-level features; and 4) third-level features. The chromosome is represented by a string of bits as shown in the figure. The number of bits required for the number of channels is patient dependent, while the other "genes" have fixed bit lengths. The effectiveness or "fitness" of each individual feature is measured using Fisher's discriminant ratio (FDR) to evaluate the preseizure and the no-preseizure classes. It is a statistical rank method that determines feature effectiveness by computing a value based on the mean and standard deviation of the two classes compared. FDR is an ideal measure to use when classes are Gaussian and uncorrelated [38] and was selected after it was determined that it could adequately describe the derived feature combinations. The objective function could have been any quantitative or qualitative measure that adequately described the data. FDR was chosen as a proof-of-concept. The FDR between two data sets is found as follows:

$$\mathrm{FDR}_k = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{\left(\mu_k^i - \mu_k^j\right)^2}{\left(\sigma_k^i\right)^2 + \left(\sigma_k^j\right)^2} \tag{12}$$

where

FDR_k	FDR for the k th feature;
N	number of classes (two in the present research);
μ_k^i	mean of the kth feature for the ith class;
σ_k^i	is the standard deviation of the kth feature for the ith class.

First, the feature was calculated for each preseizure training record and was compared with each baseline training record and the average FDR values served as the objective value for each channel/feature combination.

After the "fitness" was calculated for each solution, a *new solution set* was pseudorandomly generated from the original set with higher fitness solutions given preference. The evolutionary concept was mimicked since the less fit solutions did not survive the future generations. The process was repeated until a solution dominated the population. Ideally, the dominant solution was near the globally optimal solution.

The resultant chromosomes were weighted based on their fitness values, and the *roulette wheel selection (RWS)* method was used to select surviving features. The probability of *crossover* remained constant at 70%, and the probability of *mutation* at 10%. A constrained crossover approach permitted crossover within each gene, and prohibited crossover across genes. That is, for each iteration, only one element within the first, third, or fourth genes could crossover at a time. The *stop criterion* was set to the maximum number of generations which was set equal to the population size. The population size was equal to three times the length of the chromosome. Therefore, a larger chromosome would mean there would be more feature combinations to try, and the population size and number of maximum generations would be increased to compensate. The maximum number of generations could be increased without significant loss in computation time since the GA was designed to remember values computed in the past. Therefore, the GA had the tendency to go from generation to generation rather quickly toward the final generations.

2) Classifier Based Feature Selection: The forward sequential approach was applied to the surviving features found in the first feature selection phase. The GA selected features in each domain were tested with a probabilistic neural network (PNN) to obtain the best performing feature to which the forward sequential approach would be applied to obtain the best performing feature vector. The following figure of merit (FOM)

TABLE I
FINAL FEATURES SELECTED AND PERFORMANCE FOR FEATURE SELECTION

Table 1 Final features selected and performance for seizure prediction.

Patient	IEEG Focus Channel	Selected IEEG feature/channel for prediction	Specificity	Sensitivity	Prob. Error	Avg. Prediction Time (min)	False positives per hour
A	RT5-6	mean of median of curve length channel (LIF3-4)	100.00%	75.00%	0.01	5.56	0
B	RT2-3	max. of mean abs. deviation of energy channel (RT4-5)	85.00%	62.50%	0.13	3.25	0.36
C	RIF2-3	min. of median of curve length channel (LT1-2)	90.63%	62.50%	0.13	1.81	0.4
D	LIT1-2	minimum of sum of energy channel (LIT3-4)	86.25%	50.00%	0.09	3.2	0.35
AVERAGE VALUES			90.47%	62.50%	0.09	3.455	0.2775

was developed to determine surviving features to which the forward sequential approach would be applied

$$\mathrm{FOM} = \frac{0.55(\mathrm{TP}-\mathrm{FN})}{\mathrm{TP}+\mathrm{FN}} + \frac{0.45(\mathrm{TN}-\mathrm{FP})}{\mathrm{TN}+\mathrm{F}P} \qquad (13)$$

where

- TP number of true positives or correct "preseizure" classifications;
- FN number of false negatives or incorrect "preseizure" classifications;
- TN number of of true negatives or correct "no preseizure" classifications;
- FP number of false positives or incorrect "no preseizure" classifications.

The FOM is designed to select surviving feature combinations based on performance of the training set. The FOM is empirically derived and involves intermediary judgments in addition to the quantitative decision. Since the number of baselines and seizures varies among patients, these proportions are taken into account in the FOM. The FOM yields values ranging from -1 to 1, with the best values being those closest to "1". Correct preseizure classifications are equally weighted, while the "no preseizure" or baseline classifications are equally weighted, but given less weight than the preictal classifications.

A true positive (TP) is declared in the training data if the longest data segment where the classifier correctly classified the preictal data stream is greater than or equal to the average length of all correct preictal classifications and all incorrect baseline classifications in the training records. A false negative (FN) is declared in the training data if the longest data segment where the classifier correctly classified the preictal data stream is less than the length of the shortest possible TP. The number of training records used to calculate the average value is equal to the number of preictal and baseline training records used. A record is considered a false positive (FP) if the longest data segment where the classifier incorrectly classified the baseline data stream is greater than or equal to the length of a true positive (TP) record. A record is considered a true negative (TN) if the classifier output correctly identifies the record as a baseline data segment. The integration of decisions [53] concept is applied for defining the classification lengths for the test data. Precisely, a TP is declared if the record is 0.9 times the length of the shortest TP in the training data.

E. Classification

At the classifier stage, the probabilistic neural network (PNN) classifier assigns the output of the feature vector into the class "preseizure" or "no-preseizure." The PNN is used in this paper since the decision regions observed in the one and two dimensional scatter plots are often nonlinear and not explicitly defined. We expect the decision regions to ultimately converge to the optimal decision regions for the selected feature vector by using the PNN.

Split sample or "hold-out" techniques are used for the validation stage. To use split-sample validation, a representative sample (test set) of the data is randomly selected and not used in any way during training. After training, the network is run on the test set, which represented approximately 30% of available data in this study.

III. RESULTS

Table I provides a summary of the results obtained in this study. An average probability of prediction or block sensitivity of 62.5% was achieved with an average block false positive rate of 0.2775 FP predictions/h, corresponding to 90.47% specificity. The focus channel, historically used for evaluation in seizure-prediction research, was not selected as the best channel for predicting seizures in any patient. In all patients a single third-level feature was determined to be the final feature "vector" necessary to predict seizures. In two patients, the analysis selected an energy-based feature to predict seizures while two other patients yielded a curve length-based feature.

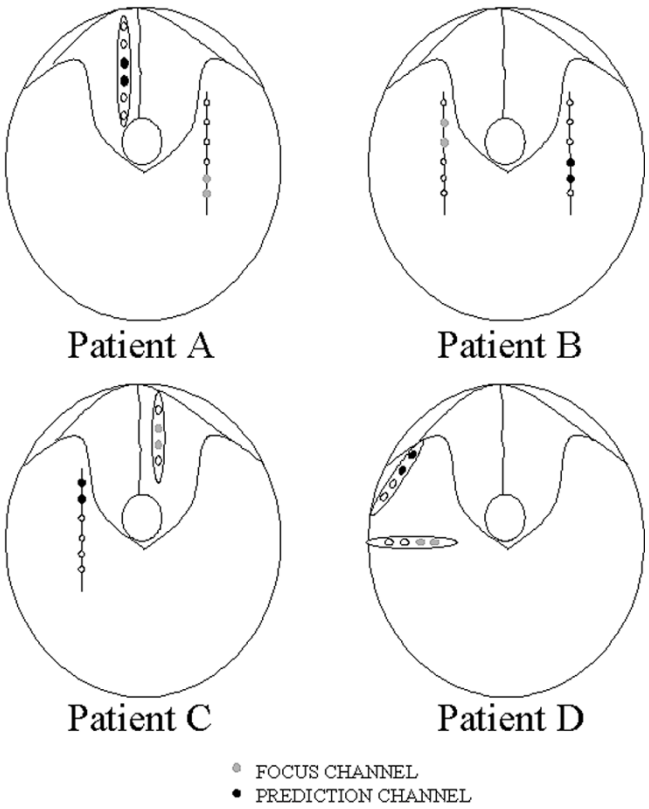


Fig. 5. Electrode contact regions where focus channels and selected prediction channels are located.

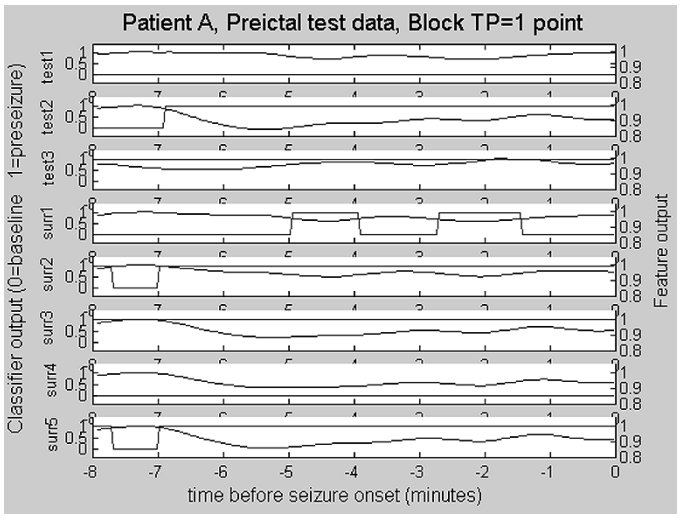
A prediction could be declared an average of 3.45 min before UEO. Each patient's results are described separately. A pictorial representation of the results for each patient is shown in Fig. 5

The processing time required for the implementation of the entire methodology over a 10 minute data segment using MATLAB® 5.3 on a stand-alone 450-MHz Pentium III ranged from 3.41 – 3.89 s, averaging 3.56 s for all four patients analyzed. This processing included the bipolar analysis, notch filtering, three levels of processing, and classification. This processing time is far below that required for real time implementation. The processing time reported here is the time required after off-line selection of the final feature vector for prediction.

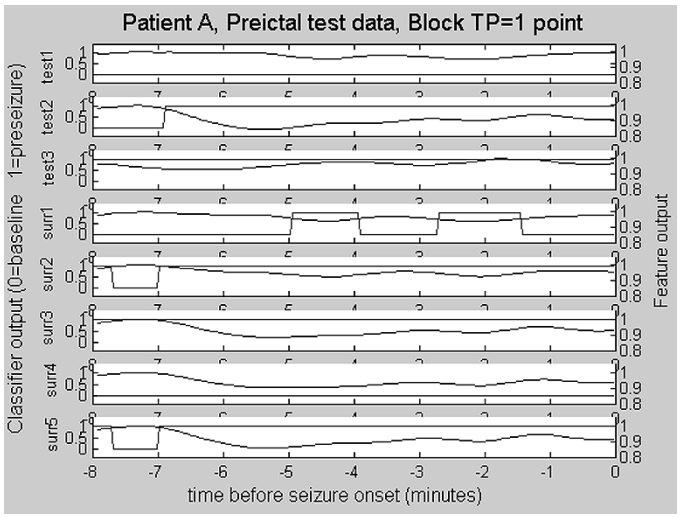
Fig. 6 provides the results from the output of the classifier for the preictal and baseline test data for this patient. Three preictal test records were available and five surrogate preictal records were created for this patient. One feature data point was required to predict a seizure for this patient. The block length required for prediction varied from patient to patient and was dependent on the training data outputs.

A. Patient A

The epileptologist found that this patient's seizures all arose from the right hippocampus, in the more posterior contacts (RT5 and RT6). Most seizures began with 3–7 s of generalized flattening and loss of background, followed by slow, semi-regular rhythmic delta ($<$ 4 Hz) activity over the right temporal region, particularly in the focus contacts. The rhythmic activity



(a)



(b)

Fig. 6. Classifier output and normalized signals for patient A. (a) Plots representing the preictal test data, up to 8 min before UEO. Three preictal test records were available and five surrogate preictal records were created for this patient. One feature data point was required to predict a seizure for this patient. The block length required for prediction varied from patient to patient and was dependent on the training data outputs. The "0" and "1" values on the y axis indicate a baseline and preictal classification, respectively. (b) Baseline test data plots. No FPs were reported for this patient's test data.

was followed by the onset of 20-25 Hz activity in the focus region, then rhythmic spiking. One seizure was followed by late revival of this activity in the left temporal region, but this was after the right temporal seizure had mostly subsided. Finally, an interesting result is that the "best" channel for prediction for this patient was actually contralateral (on the other side of the brain) to the focus channel, while the best feature was a curve length based feature.

B. Patient B

Patient B is the only patient of the four patients who was classified as having independent, bilateral seizure onsets. Three of this patient's seizures arose from left temporal and left anterior temporal electrodes, and one arose from the right temporal electrodes. The best channel for seizure prediction was actually

contralateral to the left temporal lobe where three of the four seizures originated. The best feature for this patient was an energy based feature.

C. Patient C

Patient C's seizures all seemed to begin on the right side, somewhere between the right temporal and inferior frontal region. The epileptologist suspected that the electrodes may not have been positioned directly in the ictal onset zone, because the EEC associated with seizures was not well localized to a single or distinct small group of electrodes. The pattern from seizure to seizure was consistent, even when clusters of seizures occurred. Long periods of left temporal spiking stopped and there was suppression of the EEG background of a second or two, followed by a burst of higher amplitude, rhythmic theta (4–7.5 Hz for approximately 5-s durations) activity over the RT2 and RIF 2–4 electrodes. In most seizures this was followed by a few seconds of focal beta activity in the posterior hippocampal electrodes on the right (RT4–6) then rhythmic activity that spread rapidly throughout the right temporal, inferior frontal and inferior temporal areas at the time of EEC. It was difficult to select an ictal onset zone (IOZ) channel for this patient. It was estimated to be somewhere between the RT2 or RT4–6, RIF2–4 and RIT2–4 electrodes. A volume of tissue in this region was likely responsible for these findings.

Patient C yielded a best channel contralateral to the focus. After the third level of feature extraction is achieved, clear distinguishability between the preictal and baseline records is evident in the best channel, while the preictal awake and baseline asleep records appear indistinguishable in the focus channel. These results are similar to *patient A* results. The best feature for this patient was a curve length based feature.

D. Patient D

The expert epileptologist noted that the first cluster of seizures was very focal, short, and did not spread outside of the left temporal lobe. The UEO for all seizures was exhibited in the left temporal and left inferior temporal electrode contacts. Of note, the "best" channel for prediction in this time horizon in this patient was not contralateral to the epileptic focus. Rather it was on the anterior surface of the ipsilateral temporal region. The best feature for this patient was an energy based feature.

The time for the classifier to identify a correct classification ranged from 2.48 s (1 data point) to 5.74 min (139 data points) over the four patients. Generally, the shorter the time required by the classifier to provide a true output, the better the performance of the predictor. Since patient A only required one point to declare a true positive, the average prediction performance for this patient exceeded the performance of the other three patients. Patient C required the classifier to declare a true positive consecutively for 5.74 min. Consequently, since only 8 min were available after three levels of feature extraction, a prediction horizon of less than 3 min was possible with this patient. The required block lengths were not known until the final stages of this research. These results show that approximately 20 min before UEO should be evaluated to address the ten minute prediction horizon. The prediction horizon addressed in this manuscript was minutes before UEO.

IV. Conclusion

All implanted IEEG electrode contacts monitored were used in this study. In all four patients the electrode contact most closely associated with the majority of seizures was not selected as the "best" overall channel. The bilateral patient results (patient C) yielded a channel associated with the UEO for predicting 25% of this patient's seizures. The method presented looks not for absolute preseizure changes, but rather changes in the preseizure EEG compared with remote baseline EEG. Since the EEG in the epileptic focus is abnormal at baseline, it is possible that the change from baseline is smaller minutes prior to seizure onset than the change from a "normal" brain region that is "entrained" just prior to seizure onset. In this way it is still possible that the absolute preseizure change may be greatest in the epileptic focus, but the relative change from baseline may be greatest elsewhere. The fact that the "best" channel for prediction in three of the four patients was actually contralateral to the focus channel may indicate that the cascade of events that leads to seizure onset may somehow require activation of contralateral and/or deep structures (e.g., thalamus) before actual seizure initiation can occur.

The heterogeneity of epilepsy makes identification of a neurological origin of seizures difficult to generalize. Although the focal channel is the optimal channel to **detect** the electrographic onset of seizures, the selection of channels other that the focal channel for seizure **prediction** is consist with the findings from other investigators [26], [54], [55]. The manifestation of seizures is largely dependent on the location of the brain region(s) involved and the extent to which neuronal connections are damaged and interacting with normal brain tissue and excitability. No two patients exhibit identical seizure patterns, but generally seizures of individual patients exhibit similar patterns. This is an important finding, as it may give hints to physiological mechanisms underlying seizure generation.

In patient D, the "best" channel for prediction in this time horizon was on the anterior surface of the ipsilateral temporal region. This may be somehow related to the location and pattern of spread of this subject's seizure precursors and seizures themselves. In this patient, seizures originated in the left inferior temporal neocortex, where they stayed for a short time prior to spreading to mesial temporal and other regions. This discrete focality and neocortical onset may have recruited other neurons in a different pattern than in other patients. This difference from the other patients also provides an interesting opportunity for further study.

Related to the above explanation, several investigators have reported that there is "recruitment," "entrainment," or "increased coherence" between activity in both temporal lobes prior to seizure onset in temporal lobe epilepsy [26], [54], [55]. Since these regions may be more "normal" than the epileptic focus in their baseline function, pre-ictal changes in these parts of the "epileptogenic zone" may appear more dramatic compared with baseline, and be more robust, than those that

appear in the focus area minutes prior to seizure onset. This is an important reason why the best regions for predicting seizures in this study were outside the of the ictal onset zone in 3 patients, as the method used for seizure prediction did not look for absolute changes, but rather changes in the minutes prior to seizure onset compared with baselines far removed from seizures. Activity in the ictal onset zone is abnormal at baseline, and may not change from baseline as much several minutes prior to seizure onset, as the more normal contralateral regions that are entrained prior to seizures. The regions selected outside of the ictal onset zone, however, may not demonstrate changes that occur much earlier, for example hours prior to seizure onset [2]. It is also possible that our results may identify important regions whose recruitment may be necessary prior to seizure onset. Patient D, in whom the "best" channel was on the same side but in a different region than the epileptic focus, may have different physiology, due to the location of the epileptic focus.

To date, a common set of features across individuals that gives adequate performance has not been found. To identify such an optimal feature set, an exhaustive search should be conducted; however, if a suboptimal solution can be obtained from a reasonable feature set that yields prediction sufficient for clinical application, such an exhaustive search may not be required. Due to the heterogeneity of epilepsy, we believe that it is likely that patient specific features will be found to be more useful for seizure prediction over a large subset of patients. While the feature vector selected for each patient in this study is the best possible solution among the possible solutions presented by the general feature set, based on the application of classifier based performance metrics, it is unlikely that our findings represent an optimal solution.

These results demonstrate the utility of a hybrid genetic and classifier based feature selection process for selecting a reasonable set of features for predicting seizures. These results are promising, but may still far short of the performance required for an implantable device. The next step in this research is to perform this analysis over a very large number of patients with sufficient training and test data to determine if there is a tendency for particular features and feature vectors to be selected as optimal. A large number of similar results could provide information to limit the number of preselected features in the search space, thereby providing more opportunity for examining potential objective functions and GA techniques. Further iterations of the technique may provide methods for "fusing", or deriving other types of artificial features that take advantage of the best predictive aspects of each individual feature. In addition, looking to other areas of neuro or computational science for other features to be placed into the search space may be of great utility. These features might also include some promising features from nonlinear dynamics found useful for prediction by other groups, as improvements in their computational efficiency allows.

The results presented here emphasize the "short-term" prediction horizon, minutes before the UEO of the seizure. The features selected by the GA were different for each patient analyzed, suggesting that a patient specific system may be necessary for accurate seizure prediction. The results were further refined by applying a PNN classifier to obtain a true assessment of performance. Using the PNN after the GA to assess performance provided a direct assessment of classification errors. In all four patients, classifier based performance did not select the feature with the highest theoretical performance measure. Furthermore, the feed forward approach did not result in a feature vector representing multiple channels. Only one derived feature was selected to effect prediction in each case. These results may be due, in part, to prediction being a more complex and ambitious goal than detection, or as a product of the method. More research needs to be performed to determine exactly how much of the multilevel feature extraction may be required for this application. This paper provides a good starting point for future research in this area. It is clear that channels other than the focus channel could predict seizures on a patient specific basis. What is not clear, but worthy of further investigation, is methods other than unilateral combinations for selecting multiple features and channels. Reducing the feature space as described in this study provided somewhat redundant features, but no complementarity of features and channels to predict seizures. Another disadvantage of reducing the feature space is that computational constraints may have limited the selection of optimal features.

At a minimum, this paper should help pave the way for improving currently available technology. As features researched in this paper were chosen with an eye toward real-time implementation, even this early implementation of a multichannel, multifeature method has the potential to be applied in prototype implantable devices for treating epilepsy. As this work matures, there are many important issues to consider, such as longer prediction horizons, longer duration and more continuous training records, and other approaches to combining features, to name a few. We currently are investigating other methods for combining channels and features to address multivariate feature prediction and alternative classifier based performance metrics for the GA objective function. Ultimately, an automated approach to feature and electrode contact selection is envisioned.

The Vagal Nerve Stimulator, currently the only commercially available technology used to control epileptic seizures, is an open loop system that provides adjunctive therapy by applying stimulation on average for 30 s every 5 min, achieving less than 41% efficacy. If a system could be developed to predict over 60% of seizures with performance at the level reported in this paper, significant improvement would be realized.

ACKNOWLEDGMENT

The authors would like to express their gratitude to L. Finkel, Director of Neuroengineering at the University of Pennsylvania, for his support.

REFERENCES

[1] *Epilepsy Facts and Figures*. Landover, MD: Epilepsy Foundation of America, 1999.

[2] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, and G. Vachtsevanos, "Epileptic seizures may begin hours in advance of clinical onset: A report of five patients," *Neuron*, vol. 30, pp. 51–64, 2001.

[3] L. Iasemidis and J. C. Sackellares, "Seizure warning and prediction," US Patent 6 304 775, 2001.

[4] R. Esteller *et al.*, "Adaptive method and apparatus for forecasting and controlling neurological disturbances under a multi-level control," US Patent filed, 2000.

[5] J. Echauz *et al.*, "Unified probabilistic framework for predicting and detecting seizure onsets in the brain and multitherapeutic device," US Patent filed, 2000.

[6] B. Litt, G. Vachtsevanos, J. Echauz, and R. Esteller, "Method and apparatus for predicting the onset of seizures based on features derived from signals indicative of brain activity," U.S. Patent filed, 1998.

[7] J. Dorfmeister, M. Frei, D. Lerner, I. Osorio, and J. Ralston, "System for the prediction, rapid detection, warning, prevention, or control of changes in activity states in the brain of a subject," U.S. Patent 5 995 868, 1997.

[8] R. Fischell, D. Fischell, and A. Upton, "System for treatment of neurological disorders," U.S. Patent 6 016 449, 2000.

[9] L. Hively, N. Clapp, C. S. Daw, and W. Lawkins, "Epileptic seizure prediction by nonlinear methods," U.S. Patent 5 857 978, 1996.

[10] S. Liss, "Apparatus for monitoring and counteracting excess brain electrical energy to prevent epileptic seizures and the like," U.S. Patent 3 850 161, 1973.

[11] D. Olsen, R. Lesser, J. Harris, R. Webber, and J. Cristion, "Automatic detection of seizures using electroencephalographic signals," U.S. Patent 5 311 876, 1994.

[12] M. Rise, "Implantable seizure warning system," U.S. Patent 6 018 682, 1998.

[13] S. Ward and M. Rise, "Techniques for treating epilepsy by brain stimulation and drug infusion," U.S. Patent 5 713 923, 1996.

[14] S. Viglione, V. Ordon, W. Martin, and C. Kesler, "Epileptic seizure warning system," U.S. Patent 3 863 625, 1973.

[15] Zabara, "Neurocybernetic prosthesis," U.S. Patent 4 702 254, 1985.

[16] R. Fischell, D. Fischell, and A. Upton, "Responsive implantable system for the treatment of neurological disorders," US Patent 6 134 474, 2000.

[17] V. Piccone, J. Piccone, L. Piccone, R. LeVeen, and E. L. Veen, "Implantable epilepsy monitor apparatus," US Patent 4 566 464, 1981.

[18] H. Fernández and G. Pardue, "Seizure suppression device," U.S. Patent 3 993 046, 1976.

[19] C. S. Schachter and D. Schmidt, *Vagus Nerve Stimulation*, 2001.

[20] A. Petrosian and R. Homan, "The analysis of EEG texture content for seizure prediction," presented at the *IEEE EMBS 16th Annu. Int. Conf.*, Baltimore, MD, 1994.

[21] A. Petrosian, "Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns," in *8th IEEE Symp. Computer-Based Medical Systems*, Lubbock, TX, 1995.

[22] A. Petrosian, R. Homan, S. Pemmaraju, and S. Mitra, "Wavelet based texture analysis of EEG signal for prediction of epileptic seizure," in SPIE Conf. Wavelet Applications in Signal and Image Processing San Diego, CA, 1995.

[23] L. Iasemidis, H. Zaveri, J. Sackellares, W. Williams, and T. Hood, "Nonlinear dynamics of ECoG data in temporal lobe epilepsy," *Electroencephalogr. Clin. Neurophysiol.*, vol. 5, p. 339, 1988.

[24] L. D. Iasemidis and J. C. Sackellares, "The evolution with time of the spatial distribution of the largest lyapunov exponent on the human epileptic cortex," in *Measuring Chaos in the Human Brain*, D. Duke and W. Pritchard, Eds, Singapore: World Scientific, 1991.

[25] L. D. Iasemidis, L. D. Olson, R. S. Savit, and J. C. Sackellares, "Time dependencies in the occurrences of epileptic seizures: A nonlinear approach," *Epilepsy Res.*, vol. 17, pp. 81–94, 1994.

[26] L. Iasemidis, K. Pappas, R. Gilmore, S. Roper, and J. Sackellares, "Preictal entrainment of a critical cortical mass is a necessary condition for seizure occurrence," *Epilepsia*, vol. 37, 1996.

[27] L. Iasemidis, R. Gilmore, S. Roper, and J. Sackellares, "Dynamical interaction of the epileptogenic focus with extrafocal sites in temporal lobe epilepsy," *Ann. Neurol.*, vol. 42, p. 429, 1997.

[28] K. Lehnertz and C. Elger, "Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity," *Phys. Rev. Lett.*, vol. 80, pp. 5019–5022, 1998.

[29] K. Lehnertz, G. Widman, and C. Elger, "Spatiotemporal neuronal complexity loss: Condensed information content of long-term intracranial EEG recordings," *Epilepsia*, vol. 38, p. 63, 1997.

[30] K. Lehnertz, R. Andrzejak, F. Mormann, J. Wegner, G. Widman, P. David, and C. Elger, "Linear and nonlinear analysis techniques for anticipating epileptic seizures," *Epilepsia*, vol. 40, p. 71, 1999.

[31] K. Lehnertz, "Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy—An overview," *Int. J. Psychophysiol.*, vol. 34, pp. 45–52, 1999.

[32] K. Lehnertz *et al.*, "Nonlinear EEG analysis in epilepsy," *J. Clin. Neurophysiol.*, vol. 18, pp. 209–222, 2001.

[33] R. Mormann, K. Lehnertz, P. David, and C. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D*, pp. 358–369, 2000.

[34] M. Le Van Quyen *et al.*, "Characterizing neurodynamic changes before seizures," *J. Clin. Neurophysiol.*, vol. 18, pp. 191–208, 2001.

[35] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clémenceau, and F. Varela, "Spatio-temporal characterizations on nonlinear changes in intracranial activities prior to human temporal lobe seizures," *Eur. J. Neurosci.*, vol. 12, pp. 2124–2134, 2000.

[36] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizures in real time by a nonlinear analysis of similarity between EEG recordings," *NeuroReport*, vol. 10, pp. 2149–2155, 1999.

[37] J. S. Ebersole, "The last word," *J. Clin. Neurophysiol.*, vol. 18, pp. 299–300, 2001.

[38] D. H. Kil and F. B. Shin, *Pattern Recognition and Prediction With Applications to Signal Characterization*. Woodbury, NY: AIP, 1996.

[39] R. Esteller, G. Vachtsevanos, J. Echauz, and B. Litt, "A comparison of fractal dimension algorithms using synthetic and experimental data," presented at the *IEEE Int. Symp. Circuits and Systems*, Orlando, FL, 1999.

[40] R. Esteller, J. Echauz, T. Cheng, B. Litt, and B. Pless, "Line length: An efficient feature for seizure onset detection," *Proc 23rd Int. Conf. IEEE Engineering Medicine Biology Soc.*, vol. 2, pp. 1707–1710, 2001.

[41] R. Esteller, "Detection of seizure onset in epileptic patients from intracranial EEG signals," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, 2000.

[42] H. Zaveri, W. Williams, and J. Sackellares, "Energy based detectors of seizures," *Proc. 15th Annu. Int. Conf. IEEE Engineering Medicine Biology Soc.*, pp. 363–364, 1993.

[43] R. Agarwal and J. Gotman, "Adaptive segmentation of electroencephalographic data using a nonlinear energy operator," *Proc. 1999 IEEE Int. Symp. Circuits and Systems*, vol. 4, pp. 199–202, 1999.

[44] J. F. Kaiser, "On a simple algorithm to calculate the "energy" of a signal," in *Proc1990 Int. Conf. Acoustis, Speech, Signal Processing*, vol. 1, 1990, pp. 381–384.

[45] P. Maragos, J. Kaiser, and T. Quatieri, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Signal Processing*, vol. 41, pp. 1532–1550, Apr. 1993.

[46] M. D'Alessandro, G. Vachtsevanos, R. Esteller, J. Echauz, A. Koblasz, and B. Litt, "Spectral entropy and neuronal involvement in patients with mesial temporal lobe epilepsy," presented at the *METMBS2000*, Las Vegas, NV, 2000.

[47] L. Iasemidis, A. Barreto, R. Gilmore, B. Uthman, S. Roper, and J. Sackellares, "Spatiotemporal evolution of dynamical measures precedes onset of mesial temporal lobe seizures," *Epilepsia*, vol. 35, p. 133, 1994.

[48] R. Q. Quiroga, *Quantitative analysis of EEG signals time frequency methods and chaos theory*. Lubeck, Germany: Inst. Physiol., Med. Univ. Lubeck, 1998.

[49] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda, and M. Hirano, "Quantification of EEG irregularity by use of the entropy of the power spectrum," *Electroencephalogr. Clin. Neurophysiol.*, vol. 79, pp. 204–210, 1991.

[50] A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, and D. Wunsch, "Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG," *J. Neurocomput.*, vol. 30, pp. 201–218, 2000.

[51] S. Blanco, H. Garcia, R. Q. Quiroga, L. Romanelli, and O. A. Rosso, "Stationarity of the EEG series," *IEEE Eng. Med. Biol. Mag.*, vol. 14, pp. 395–399, 1995.

[52] E. Chang, R. Lippmann, and D. Tong, "Using genetic algorithms to select and create features for pattern classification," in *IEEE Int. Joint Conf. Neural Networks*, vol. 3, 1990, pp. 474–752.

[53] C. C. White, III, "A survey of the integration of decision analysis and expert systems for decision support," *IEEE Trans Syst., Man, Cybern.* , vol. 20, pp. 358–364, Mar./Apr. 1990.

[54] M. D'Alessandro, "The utility of intracranial eeg feature and channel synergy for evaluating the spatial and temporal behavior of seizure precursors," Ph.D.dissertation, Georgia Inst. Technol., Atlanta, 2001.

[55] M. D'Alessandro, G. Vachtsevanos, R. Esteller, J. Echauz, L. Buell, C. Bowen, R. Shor, and B. Litt, "Spectral entropy and ictogenesis in mesial temporal lobe epilepsy," *Epilepsia*, vol. 40, p. 174, 1999.

Maryann D'Alessandro (S'95–M'02) received the B.S.E.E. and M.S.E.E. degrees from the Florida Institute of Technology, Melbourne, and Fairleigh Dickinson University, Teaneck, NJ, in 1988 and 1992, respectively. She received the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2001, after more than ten years of active military service in the U.S. Army in various research and project engineering positions at the Communications Electronics Command (CECOM) and the Army Research Labs (ARL).

Currently, she works as a Postdoctoral Researcher at the University of Pennsylvania, Philadelphia, refining her signal processing approaches to seizure prediction and developing new signal analysis techniques. Her research interests are in the application of signal processing techniques to address epileptic seizure prediction, as well as other disorders such as Parkinson's disease, stroke, and autism.

Rosana Esteller (S'98–M'00) received the B.S.E.E. and M.S.E.E. degrees from the Simón Bolívar University, Caracas, Venezuela, in 1986 and 1994, respectively, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2000.

Currently, she works as a Research Scientist at NeuroPace, Inc., Mountainview, CA, developing an implantable device to treat epilepsy and other disorders. Her research interests are in applications of signal processing techniques from linear and nonlinear dynamics for detection, prediction, classification, pattern recognition, modeling, and estimation of signals and systems. Particular interest in applications to biomedical signals and intracranial EEG signals from epileptic patients.

Dr. Esteller is a member of Sigma Xi.

George Vachtsevanos (S'62–M'63–SM'89) received the B.E.E. degree from the City College of New York, the M.E.E. degree from New York University, and the Ph.D. degree in electrical engineering from the City University of New York, in 1962, 1963, and 1970, respectively.

He is currently a Professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology, Atlanta, where he directs the Intelligent Control Systems Laboratory. His research interests include intelligent systems, diagnostics and prognostics, robotics and manufacturing systems. He has published in the areas of control systems, power systems, bioengineering, and diagnostics/prognostics. Dr. Vachtsevanos is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

Arthur Hinson (S'00–M'01) was born in Los Angeles, CA. He received the B.A.E. and M.S. E.C.E. degrees in the areas of Control Systems Engineering and Electromagnetics from the Georgia Institute of Technology, Atlanta, in 1999 and 2002, respectively.

He is currently working for the Department of the Navy in the Advanced Concept Development Group of the Naval Surface Warfare Center in Dahlgren, VA.

Javier Echauz (S'93–M'95) was born in San Juan, Puerto Rico. He received the B.S.E.E. degree from the University of Puerto Rico – Mayagüez (UPRM) in 1988, and the M.S.E.E. and Ph.D. degrees in the areas of computational intelligence, control systems, nonlinear dynamics, and quantitative EEG analysis from Georgia Institute of Technology, Atlanta, in 1989 and 1995, respectively.

Before joining IntelliMedix, Inc., Atlanta, GA, he was Associate Professor and Associate Director of the Electrical and Computer Engineering Department at UPRM. His research interests over the past five years have centered around intelligent devices for epilepsy. He is currently Senior Research Scientist at NeuroPace, Inc., Mountainview, CA

Brian Litt (A'88–S'89–M'91) received the A.B. degree in engineering and applied science from Harvard University, Cambridge, MA, in 1982 and the M.D. degree from Johns Hopkins University, Baltimore, MD, in 1986. He did a residency in neurology and at Johns Hopkins Hospital 1988–1991.

He was a member of the neurology faculty at Johns Hopkins Hospital from 1991–1996 and of the neurology/biomedical engineering faculty at Emory University/Georgia Institute of Technology, Atlanta, from 1997–1999. He is currently an Assistant Professor of Neurology and Bioengineering, and Director of the EEG Laboratory at the Hospital of the University of Pennsylvania, Philadelphia. His research bridges his work as a clinical neurologist specializing in epilepsy with neuroscience and bioengineering. His current focuses include automated implantable devices for the treatment of epilepsy; seizure prediction: computational modeling and mechanisms underlying seizure generation, localization of seizures in extra-temporal epilepsy, and minimally invasive tools for acquisition and display of high-fidelity electrophysiologic recording. Dr. Litt's work is supported by grants from the Whitaker Foundation, American Epilepsy Society, Epilepsy Foundation, Dana Foundation and National Institutes of Health grants RO1NS041 811-01 and RO1MN062 298-02.