Patient-specific seizure onset detection

Ali Shoeb,[a,*] Herman Edwards,[b] Jack Connolly,[b] Blaise Bourgeois,[b,c,d]
S. Ted Treves,[d,e] and John Guttag[f,*]

[a] *Harvard–MIT Health Sciences and Technology, Medical Engineering and Medical Physics Program, Cambridge, MA, USA*
[b] *Division of Epilepsy and Clinical Neurophysiology Laboratory, Children's Hospital Boston, Boston, MA, USA*
[c] *Children's Hospital Boston, Boston, MA, USA*
[d] *Harvard Medical School, Boston, MA, USA*
[e] *Division of Nuclear Medicine, Children's Hospital Boston, Boston, MA, USA*
[f] *Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge, MA, USA*

Abstract

This article presents an automated, patient-specific method for the detection of epileptic seizure onset from noninvasive electroencephalography. We adopt a patient-specific approach to exploit the consistency of an individual patient's seizure and nonseizure electroencephalograms. Our method uses a wavelet decomposition to construct a feature vector that captures the morphology and spatial distribution of an electroencephalographic epoch, and then determines whether that vector is representative of a patient's seizure or nonseizure electroencephalogram using the support vector machine classification algorithm. Our completely automated method was tested on noninvasive electroencephalograms from 36 pediatric subjects suffering from a variety of seizure types. It detected 131 of 139 seizure events within 8.0 ± 3.2 seconds of electrographic onset, and declared 15 false detections in 60 hours of clinical electroencephalography. Our patient-specific method can be used to initiate delay-sensitive clinical procedures following seizure onset, for example, the injection of a functional imaging radiotracer.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Machine learning; Seizure detection

1. Introduction

Approximately 1% of the world's population exhibits symptoms of epilepsy [1], a serious disorder of the central nervous system that predisposes those affected to recurrent seizures. A seizure is a sudden breakdown of the neuronal activity of the brain that is clinically manifested by an involuntary alteration in behavior, movement, sensation, or consciousness. These clinical behaviors are preceded and then accompanied by electroencephalographic alterations that include discharges of monomorphic (single-frequency) waveforms; polymorphic (multifrequency) waveforms; spike and sharp wave complexes; or periods of reduced electrocerebral activity [2,3].

More than 20% [4] of epilepsy patients suffer from seizures that are refractory to medication. For some of these patients a cerebral resection is an option if the brain region giving rise to seizure activity, the *epileptogenic focus*, can be identified. Single-photon emission computed tomography (SPECT) can aid in the localization of an epileptogenic focus [5] by revealing focal increased perfusion in a region of the brain associated with epileptogenic activity. For the ictal hyperperfusion to be as localized as possible, it is essential that injection of the radiotracer occur within a few seconds of seizure onset. Delayed tracer injections often result in visualization of other foci of hyperperfusion away from the primary focus which may be related to post ictal secondary activation.

* Corresponding authors.
E-mail addresses: ashoeb@mit.edu (A. Shoeb), guttag@lcs.mit.edu (J. Guttag).

One approach to minimizing the delay between seizure onset and injection of the radiotracer requires that an experienced electroencephalographer continuously monitor a subject's electroencephalogram, and request injection of the radiotracer soon after noting the earliest *electrographic* signs of a seizure. Doing this is costly, difficult, and mentally taxing and is therefore rare. In practice, a nurse injects a dose of the radiotracer after a caregiver near the patient observes *clinical* manifestations of a seizure [6]. This often results in long delays because of the subtlety of early clinical signs and the potential distance of the nurse from the patient. In our experience, injections are started 30–55 seconds after the onset of clinical indications, which often leads to poor localization of the epileptogenic focus. Our automated seizure onset detector provides assistance that could be used to consistently minimize the delay between the onset of electroencephalographic alterations and injection of the radiotracer. The detector may alert staff to the seizure's onset or activate a drug infusion pump that delivers the radiotracer.

Early work in seizure detection [7,8] produced algorithms that were meant to recognize seizure events, but not necessarily their onset, in any subject. The variability of EEGs among patients limited these algorithms to detecting the most common seizure patterns and contributed to their high false-alarm rate. We developed a *patient-specific* method to exploit the consistency of seizure and nonseizure EEGs within patients. This consistency also motivated our treatment of patient-specific seizure detection as a binary classification problem. In such problems, a classifier determines to which of two classes an observation most likely belongs based on a comparison of its features with the learned features of training examples from each of the two classes. In our case, the observation is an electroencephalographic signal; its features include the *morphology* and *spatial distribution* of waveforms on the scalp; and it is classified as an instance of seizure or nonseizure EEG based on training examples from these classes. We include as part of the nonseizure class examples of baseline, artifact, and hallmark activity from different states of consciousness so that the detector can *learn* to recognize these activities. This is in contrast to the more usual approach of actively removing or rejecting these signals using linear filters, adaptive filters, or blind source separation algorithms [9–11].

We capture the morphology of EEG waveforms by measuring their *energy at different time scales* using a multiresolution wavelet decomposition, and we encode their spatial distribution by their placement within the vector of features passed to the classifier, a support vector machine [12–14]. This particular classification algorithm was chosen because it is well suited for classifying nonlinearly separable, high-dimensional feature vectors, even when trained on imbalanced training sets.

This is crucial since our detector is expected to recognize waveform morphologies with a certain spatial distribution given an abundance of nonseizure examples and a far smaller number of seizure training examples.

In Section 2 we present the details of our detector. This is followed by a detection example in Section 3, a review of our testing methodology and results in Sections 4 and 5, a discussion of an alternate detection architecture in Section 6, related work in Section 7, and a summary and conclusions in Section 8.

2. Detection algorithm

2.1. Overview

The block diagram in Fig. 1 presents the architecture of the patient-specific seizure detector. The detector passes 2-second epochs from each of 21 bipolar EEG channels through a feature extractor to compute features characterizing the morphology of each channel's waveform. The features extracted from all the channels are grouped into one large feature vector to capture spatial correlations between the channels. This feature vector is then assigned to either the seizure or nonseizure class using a support vector machine trained on previously acquired feature vectors representing patient-specific examples of seizure and nonseizure EEGs. Seizure onset is declared only when three consecutive 2-second epochs are classified as members of the seizure class. Requiring seizure activity to last for 6 seconds prior to declaring a seizure event helps avoid false detections due to short-time, seizure-like activity, which is commonly observed between actual seizures.

2.2. Feature extraction

In any application of machine learning, deciding how to extract salient features from the input data is a critical step. In this application, the salient features of the data are the morphology of EEG waveforms and their spatial localization. We represent the morphology of EEG waveforms by their allocation of energy within different time scales. For instance, a spike-and-slow-wave pattern is represented by a simultaneous allocation of energy within a short-time-scale (high-frequency) "spike" signal and a long-time-scale (low-frequency) "wave" signal as illustrated in Fig. 2. In the actual algorithm, we used the energy in four time scales to describe the morphology of the EEG signal. The spatial localization of the EEG waveforms is encoded in their location within the feature vector presented to the support vector machine classifier. Details regarding our use of a wavelet decomposition to extract activity at four time scales are described in Appendix A.
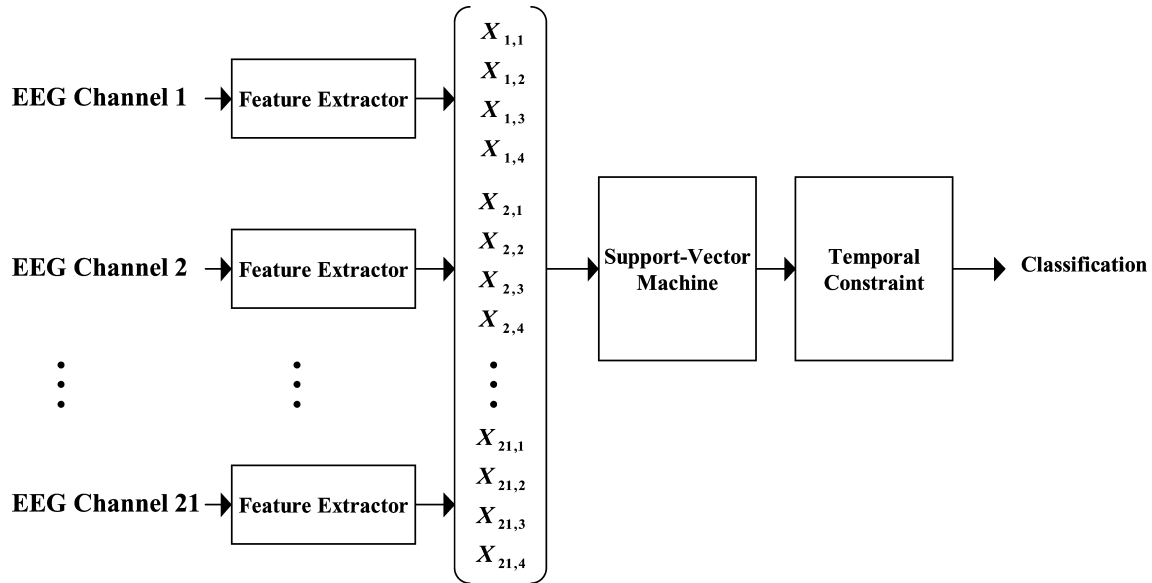
Fig. 1. Detector architecture. Four features representing waveform morphology on each of 21 channels are extracted independently and then assembled into a large feature vector that captures spatial correlations between channels. A support vector machine determines the class membership of the feature vector. The onset of a seizure is declared only when seizure activity is detected for a duration of 6 seconds.
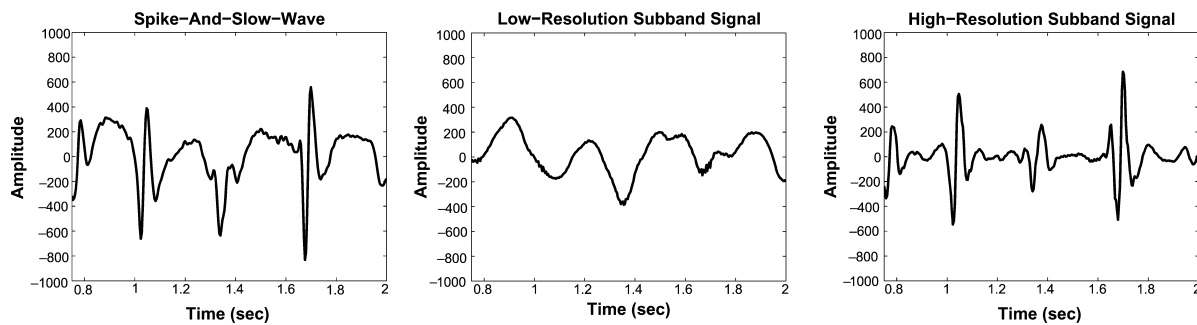

Fig. 2. Decomposition of spike-and-slow-wave signal. Left: Original spike-and-slow-wave pattern. Middle: Signal capturing long-time scale "wave" component. Right: Signal capturing short-time-scale "spike" component.

2.3. Classification

In the classification stage of the detector, a newly observed feature vector is assigned to the seizure or nonseizure class using a support vector machine trained on feature vectors representing seizure and nonseizure EEGs. When used with a linear *kernel*, support vector machines determine the class membership of an observed feature vector based on which side of a separating hyperplane the observation lies. The separating hyperplane is defined to be maximally distant from the boundary cases of each class. These boundary cases are called *support vectors*, and they carry the information relevant to solving the classification problem. If the classes cannot be well separated by a hyperplane, as is the case in our application, support vector machines can be used with more complex kernels to determine nonlinear decision boundaries.

Support vector machine classifiers are different from neural networks, which have been used extensively for the purpose of seizure detection. Neural networks determine a decision boundary that minimizes the misclassification of samples in the training set, while support vector machines maximize the margin between the decision boundary and the most similar cases from each class. The latter approach results in better performance of the classifier on unseen samples. Furthermore, support vector machines do not require a balanced number of seizure and nonseizure training examples, as is the case with neural networks. This is crucial given the much greater number of nonseizure training examples available.

As an example, Fig. 3 illustrates linear and nonlinear decision boundaries determined by a support vector machine trained on two-dimensional projections of 4 dimensional seizure and nonseizure feature vectors from

Linear Decision Boundary   **Radial−Basis Nonlinear Decision Boundary**
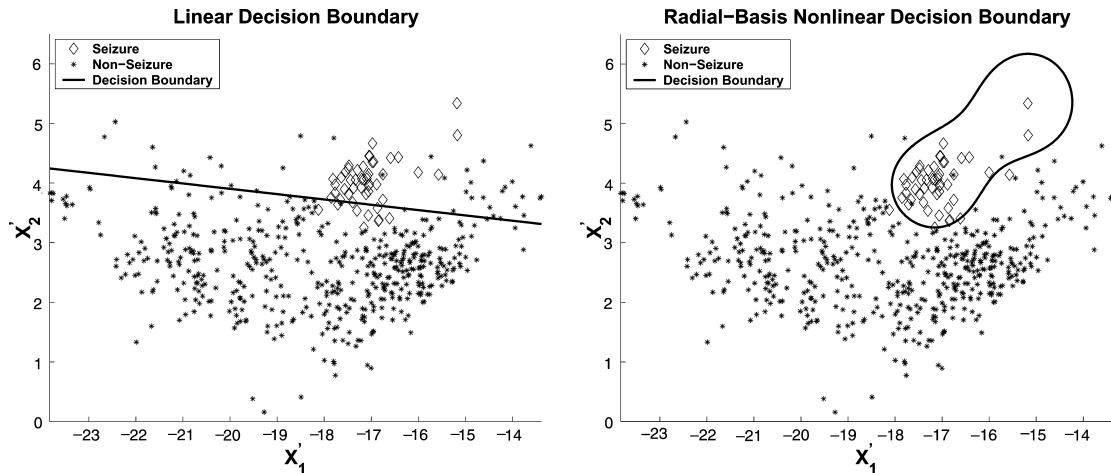
Fig. 3. Support vector machine linear and nonlinear decision boundaries. The linear decision boundary does not separate the seizure and nonseizure feature vectors as effectively as the nonlinear decision boundary.

a single EEG channel. These two-dimensional projections are used only for purposes of exposition. The nonlinear decision boundary determined using a radial-basis kernel better separates the training examples than the linear decision boundary. The nonlinear boundary is more effective as it encloses the seizure feature vectors, which are surrounded circumferentially by nonseizure feature vectors.

More generally, the ability of a support vector machine to discriminate between two classes is influenced by their separability, the parameters of the chosen kernel, and the class-specific penalty for determining a decision boundary that misclassifies a percentage of training examples. In the case of the radial-basis kernel, decreasing its sole parameter σ translates into increasingly sophisticated boundaries that correctly classify a higher percentage of training examples. Similarly, increasing the penalty for misclassifying the training examples of a given class favors the determination of a decision boundary that correctly classifies those examples. Extreme choices for both of these variables increase the risk of overfitting, that is, creating a classifier that performs well on the training set, but poorly on an unseen test set. The risk of overfitting can be gauged by the percentage of training examples considered as support vectors; the greater the percentage, the higher the risk of overfitting. By studying our detector's performance on a fraction of the data, we found that its best performance was achieved using equal class-specific penalties for the seizure and nonseizure classes, and setting $\sigma = 1$.

3. Detection example

Consider detecting the electrographic onset of the seizure in Fig. 4. This seizure's onset is characterized by a paroxysmal 4-Hz burst of monomorphic waves local-

izing to the right frontocentral channels (FP_2-F_4; F_4-C_4); the right frontal and posterior channels (FP_2-F_8; F_8-T_8; T_8-P_8; P_8-O_2), and the central channel (F_Z-C_Z).

The first step in the detection process is to train the detector on two to four previous seizure onsets and on nonseizure EEG separating these occurrences. Multiple training seizures are necessary so that the detector can reliably discover the channels that are consistently active following the electrographic onset. Fig. 5 shows one of the training seizures presented to the detector; the training seizure is very similar to the one we hope to detect.

When the trained detector was used, a seizure event was declared 7 seconds following the electrographic onset, as shown in Fig. 6. This means that the detector could potentially initiate injection of a SPECT radiotracer 7 seconds following the seizure's electrographic onset.

The actual radiotracer injection began 56 seconds after the electrographic onset shown in Fig. 4. At this point in time, both the morphology and localization of seizure activity had changed, as shown in Fig. 7. Seizure activity localized most prominently to EEG channels on the right side of the head at seizure onset, but at the time of injection seizure activity localized most prominently to the left side of the head. We suspect that reviewing the images associated with the 7- and 56-second radiotracer injection delays would lead to different conclusions about the localization of the seizure focus.

4. Testing methodology

The performance of the seizure onset detector was evaluated on 36 de-identified pediatric test subjects. A set including two to five bipolar EEG recordings sampled at 256 Hz were available for each subject. Each
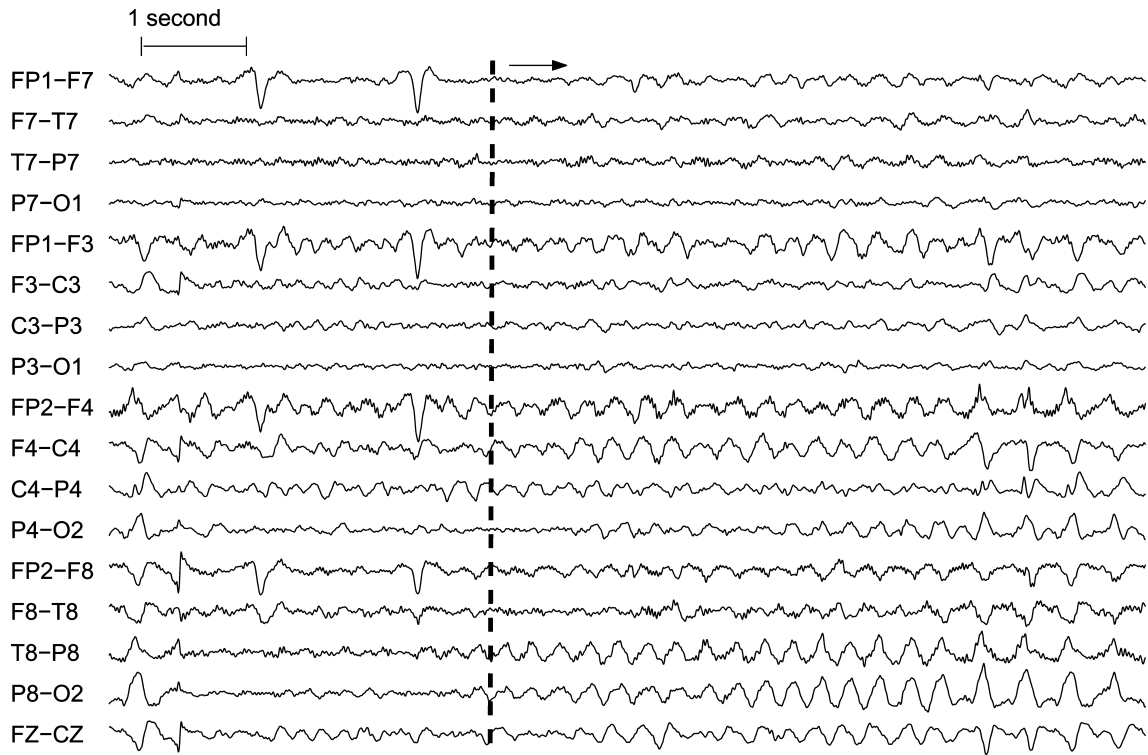
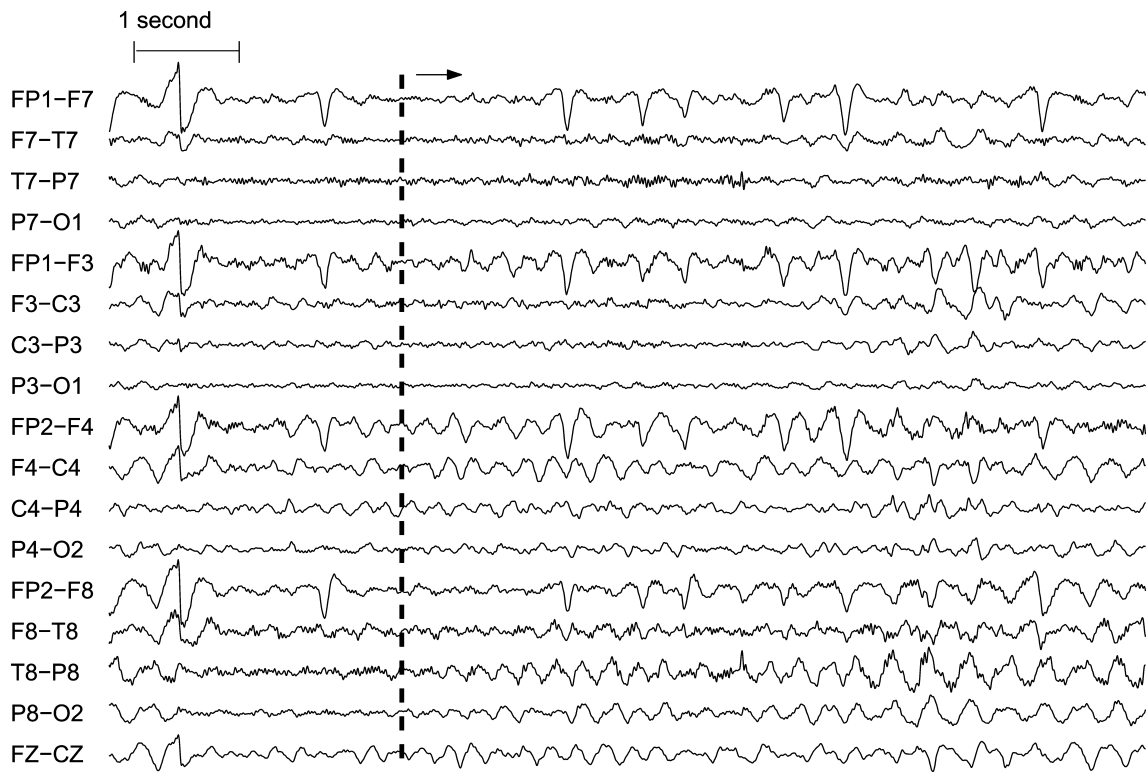Fig. 4. Electrographic onset of test seizure following dashed line.

Fig. 5. Electrographic onset of training seizure following dashed line.

recording contained a seizure event with a labeled onset that was attached by an experienced electroencephalographer, who worked backward from the observed clinical onset to find the electrographic onset. No constraints regarding the types of seizure onsets were imposed; the data set contains focal, lateral, and
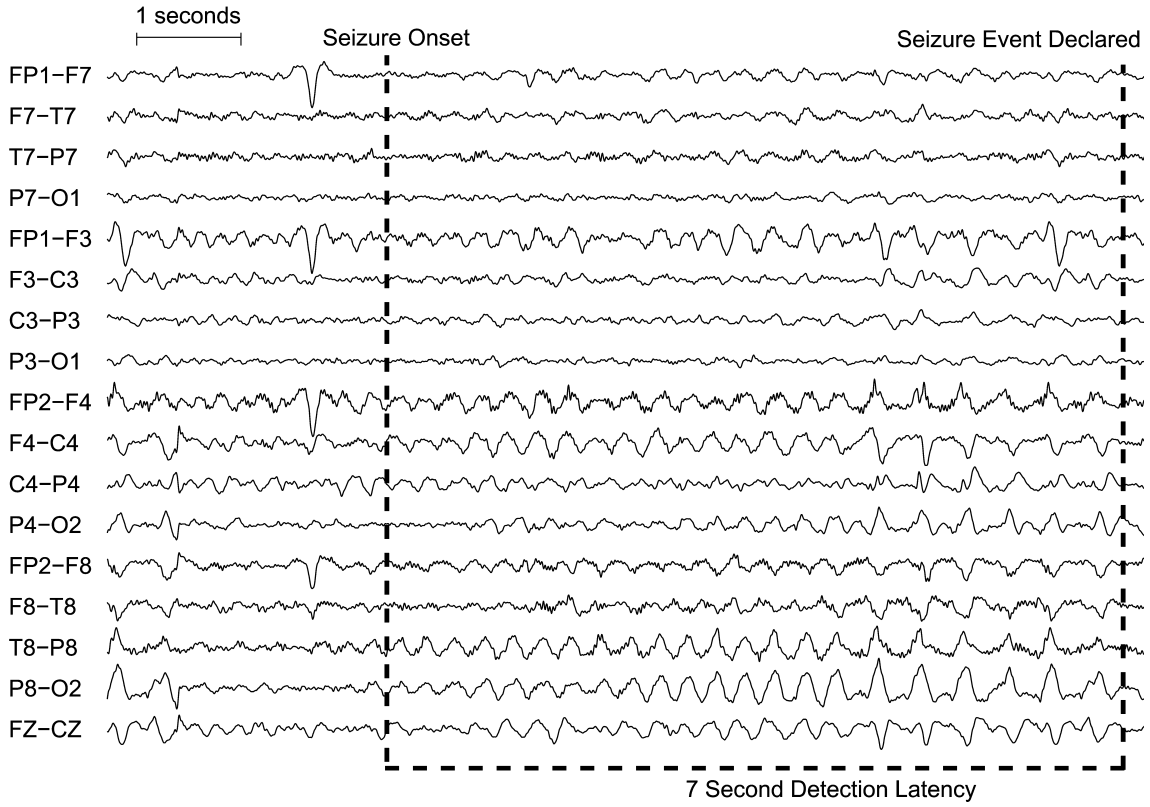
Fig. 6. Detection of seizure onset. Seizure event is declared 7 seconds after the actual seizure onset.

Fig. 7. Injection of radiotracer. Hospital staff begin injection of radiotracer at the dashed line, 56 seconds after electrographic seizure onset.

generalized seizure onsets. Furthermore, the recordings were made in a routine clinical environment, so non-seizure activity and artifacts such as head/body movement, chewing, blinking, early stages of sleep, and electrode pops/movement were present. A set of recordings lasted on average 35 minutes for 30 subjects; 2 hours for 4 subjects; and 12 hours for 2 other subjects. Taken together the recordings account for 60 hours of EEG recordings and 139 seizures.

Our first set of experiments involved following a leave-one-out cross-validation testing scheme for each subject. In particular, the detector was given a training set that included the seizure and nonseizure epochs from all but one of the subject's recordings, and was subsequently used to attempt to detect the seizure in the excluded recording. This was repeated until each recording from the subject was excluded once. For each subject we report:

- *Mean detection latency*: the average delay between electrographic seizure onsets marked by the electroencephalographer and algorithmic seizure event declaration.
- *True detections*: the number of test seizures declared as seizure events.
- *False detections*: the number of false positives declared during analysis of nonseizure EEGs.

In another set of experiments we tracked changes in the detector's performance when trained on one, two, and three recordings from a given patient. Specifically, for each subject with four EEG records (21 of the 36 subjects), we attempted to detect each seizure using a randomly selected training set of size one, two, and then three recordings. We noted whether or not each seizure was detected and with what latency for each size training set. We report the mean latency (averaged across all patients) and the percentage of all seizures detected for each training condition. The goal of this set of experiments was to gain some insight into how much training is needed to build an effective detector.

Finally, we recorded the performance of our detector when used in a non-patient-specific mode. In the non-patient-specific mode the detector was trained on seizure and nonseizure EEG epochs from a collection of 10 randomly selected subjects that excludes the test subject, and was subsequently used to attempt to detect the seizures of the excluded test subject. For each test subject we report the mean detection latency, true detections, and false detections as defined above.

5. Experimental results

Fig. 8 shows the average detection latency for each test subject; latencies were calculated using the leave-one-out cross-validation scheme. For most subjects, the mean detection latency is less than 10 seconds. The detection latency for a few subjects is less than 6 seconds because of recognition of seizure activity in a 2-second epoch that precedes, or overlaps, the onset of seizure activity marked by the expert. The detector performed poorly on two subjects. An artifact masking seizure onset activity on a number of channels resulted in poor performance on subject 33. The inconsistent seizure onsets of subject 23 resulted in failure of the algorithm to detect any seizures; this is indicated by the absence of a bar for subject 23 and is discussed further below.

The preceding results can be better appreciated if the delay incurred by starting injection of the radiotracer using existing protocols is compared with the delay that would be incurred provided injections were triggered using our method. The latter injection delay corresponds to our detector's seizure onset detection latency. Fig. 9 illustrates this comparison for a number of subjects, and shows that the detector, on average, would begin injection of the radiotracer more than 20 seconds before the actual injection.

Fig. 10 shows the number of false detections declared for each test subject under the leave-one-out cross-validation study. The detector performed well on most patients; there were no false detections for 26 of the 36 subjects. The false detections of subjects 8, 9, 29, and 36 were caused by nonphysiological artifacts. All other false detections were caused by periodic discharges that exceeded 6 seconds in duration, and closely resembled seizure onset of the subject in every aspect except duration. They may, in fact, have been "miniseizures." These false detections can be avoided by forcing the detector to declare a seizure event only in the presence of seizure activity for 8 or 10 seconds. However, decreasing false detections in this fashion will increase the detector's average latency.

Fig. 11 shows the true detections declared on each test subject under the leave-one-out cross-validation tests. No seizures were detected for subject 23 because of inconsistencies in the spatial distribution of seizure activity present in each of the three recordings. In particular, one of the seizures involves activity that is most prominent on the frontocentral channels; the second involves activity that is most prominent on the central channels; and the third involves activity that is most prominent on central and posterior channels.

Fig. 12 illustrates the improvement in our detector's average detection latency and true detection rate as a function of the number EEG training recordings observed; a training recording includes a single seizure event as well as nonseizure activity from a given subject. The figure shows that a detector trained on one recording from a test subject is capable, on average, of detecting 91% of that subject's future seizures with a mean latency of 9.5 ± 5.0 seconds. When an additional
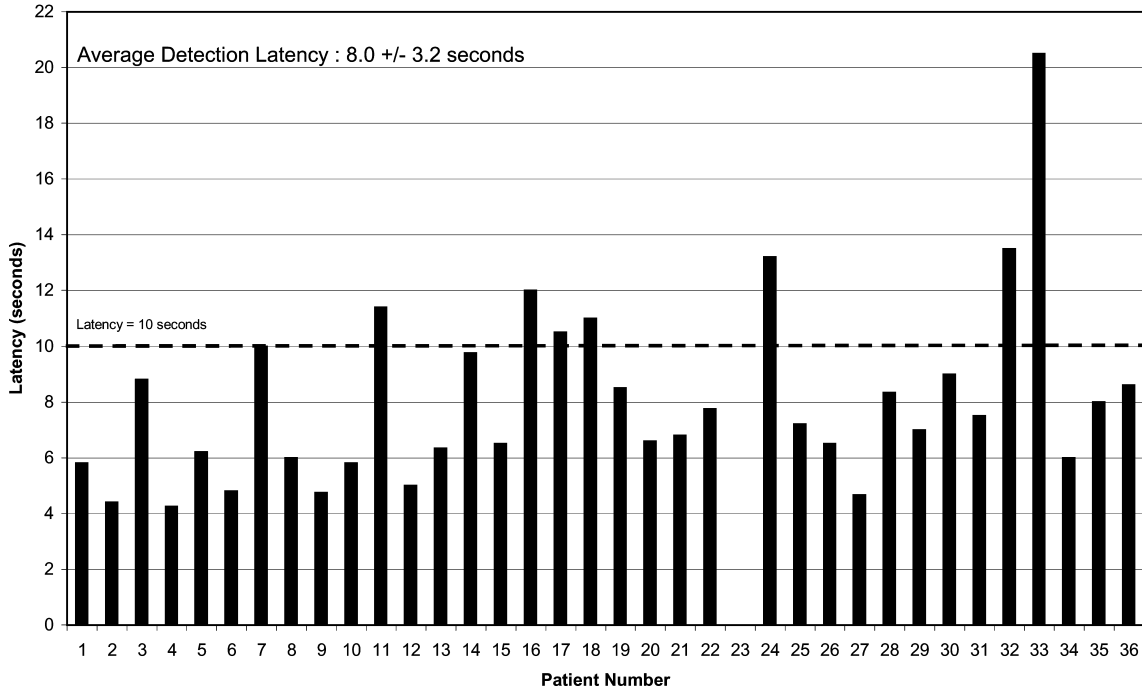
Fig. 8. Patient-specific seizure onset detector latency. Each bar indicates the average latency in detecting a subject's seizure onset. The detection latency for the majority of subjects is below 10 seconds.
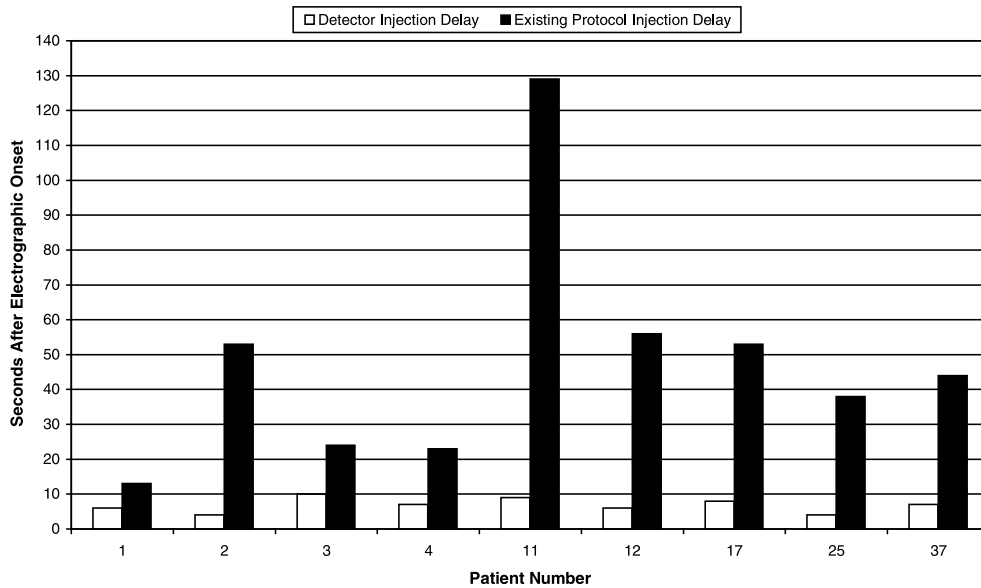

Fig. 9. Comparison of radiotracer injection delays. Black bars represent the time elapsed between electrographic onset of a seizure and injection of a radiotracer using existing protocols. White bars represent the time that would elapse if injections were triggered by the patient-specific seizure detector; this period corresponds to the detector's seizure onset detection latency. Subject 37 was only part of the study comparing radiotracer injection delays.

training recording is observed, the detector identifies 96% of the subject's future seizures with a latency of 7.6 ± 2.4 seconds. Observing a third recording only slightly improves overall performance. In particular, a detector trained on three recordings detects on average 97% of a subject's future seizures with a mean latency of 7.1 ± 1.9 seconds.

Fig. 13 contrasts the performance of our detector when used in patient-specific and generic modes. Performance tests were conducted for patients numbered 1–34, but results from patients 4 and 29–34 are omitted from the figure due to exceedingly poor performance by the generic detector. In particular, the generic detector declared an excess of 50 false detections on the EEG
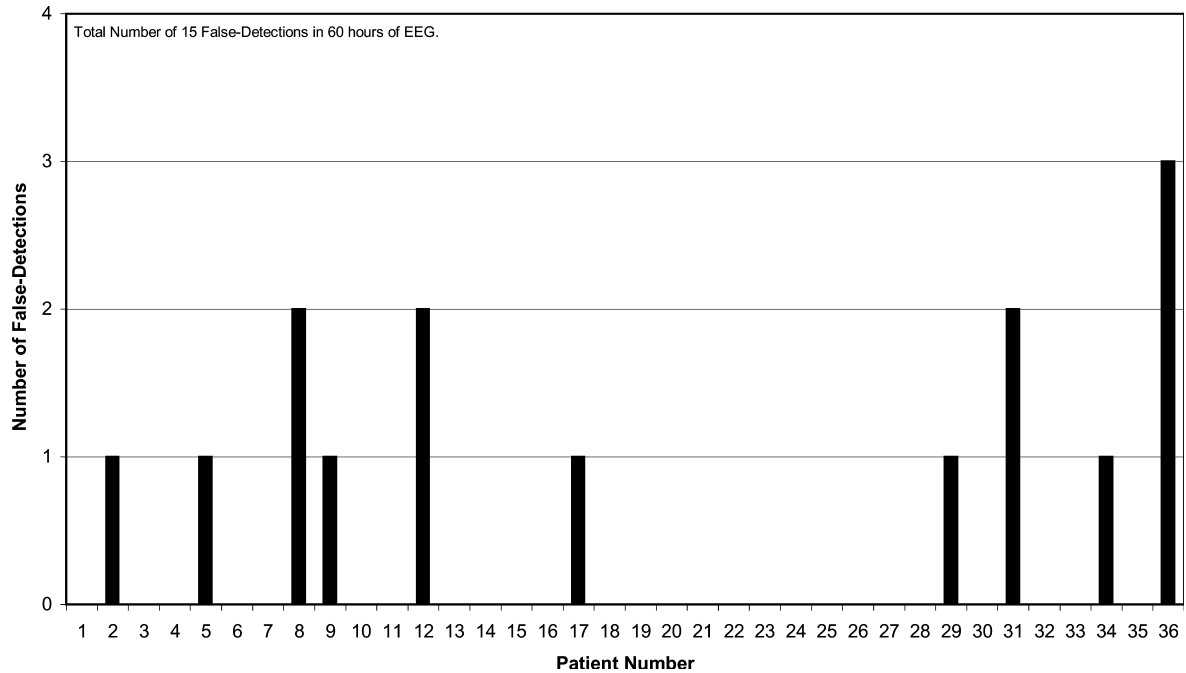
Fig. 10. Patient-specific seizure onset detector false detections: Each bar indicates the number of false detections declared while processing a subject's EEG recordings. The majority of false detections declared by the proposed method are due to short-time, seizure-like discharges commonly observed between a subject's actual seizure events.
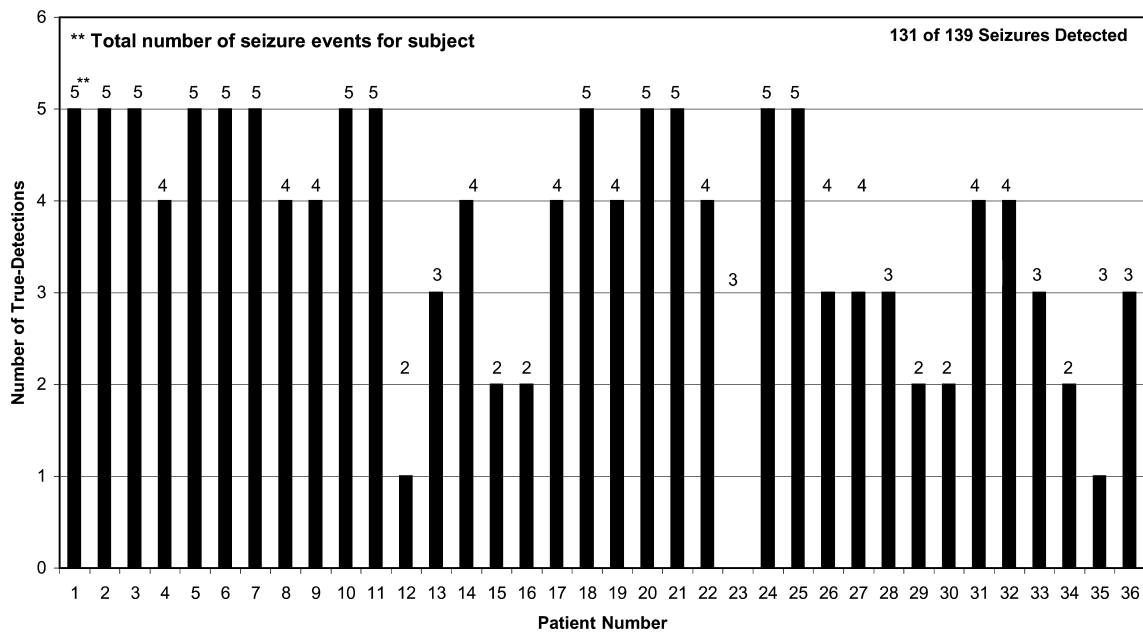

Fig. 11. Patient-specific seizure onset detector true detections. Each bar indicates the number of seizures detected per subject. The total number of seizures available for a subject is indicated over each bar.

recordings of these subjects due to the similarity of seizure EEG from the training subjects and nonseizure EEGs from the test subjects. These results suggest that a generic detector always *runs the risk of declaring many false detections due to an a priori unknown similarity between a training seizure and testing nonseizure EEG.*

To verify that poor performance resulted from the seizure EEG of training subjects matching the nonseizure EEG of the test patients (4,29–34), we conducted the following experiment: We recorded the number of false detections for each of patients 4 and 29–34 when the detector was trained on a single training subject at a
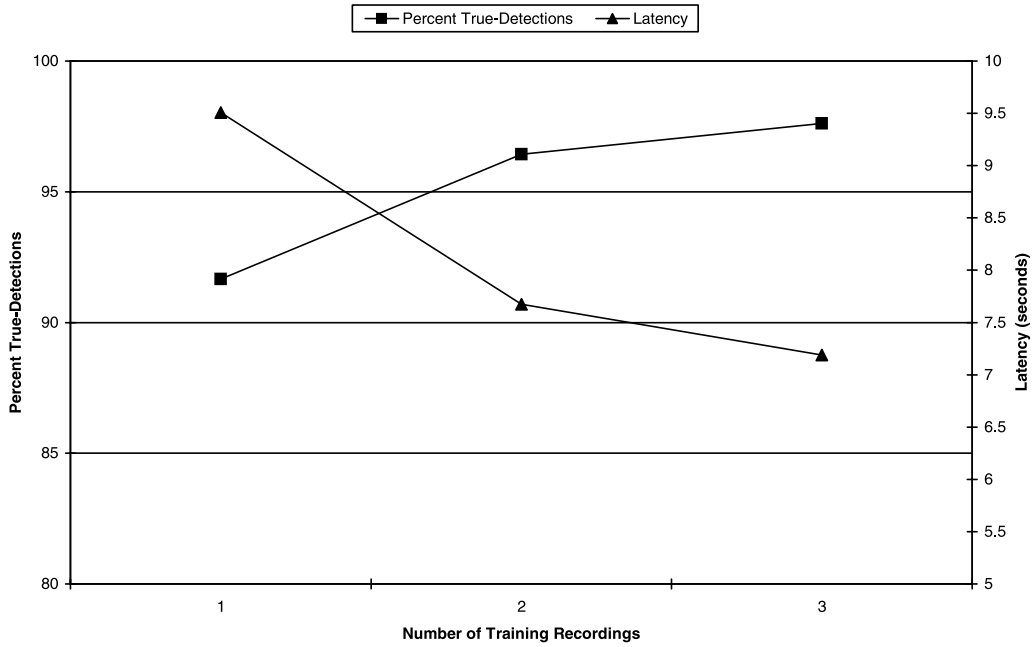
Fig. 12. Effect of number of training records on patient-specific detector's performance. There is a greater improvement in performance following the first increment in the number of training records.

time, and then noted the training subjects resulting in 15 or more false detections. Next, we constructed a training set for the test patients that excludes the problematic training subjects, and verified the ability of the generic detector to process the recordings with fewer false detections. Specifically, the new training sets allowed the detector to process each of the recordings from patients 4 and 29–34 with only 2–10 false detections.
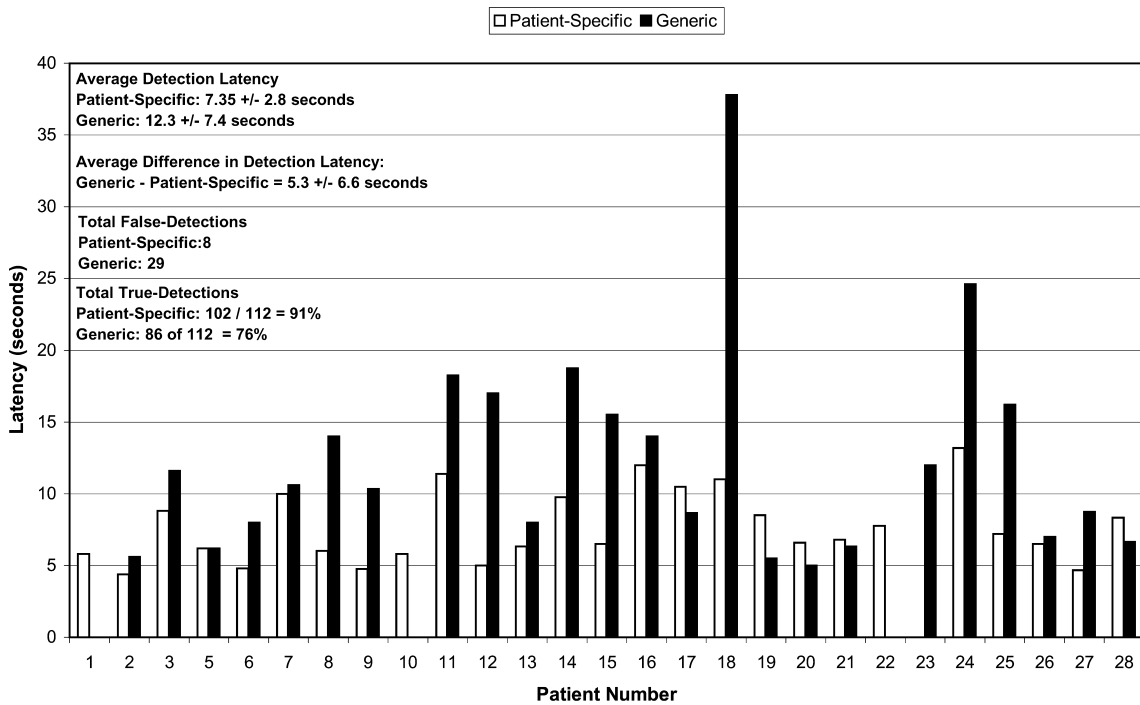

Fig. 13. Comparison of patient-specific and generic seizure detection. Each bar indicates the average latency in detecting a subject's seizure onset using both a patient-specific detector and a generic detector. On average, the patient-specific detector detects more seizures sooner and with fewer false detections. Furthermore, the generic detector runs the risk of declaring many false detections due to possible matches between the training seizure and testing nonseizure EEG. Note that patient-specific results are the same as those in Fig. 10.

On the remaining test subjects of Fig. 13, the generic detector exhibited a smaller number of true detections and greater number of false detections relative to the patient-specific detector. This is clear when comparing the generic detector's 76% seizure detection rate and 29 false detections with the 91% detection rate and 8 false detections of the patient-specific detector. Furthermore, the generic detector declared a seizure event on average 5.3 ± 6.6 seconds after the patient-specific detector. Finally, subject 27 is noteworthy, because the patient-specific detector failed to detect a test seizure due to its dissimilarity to other training seizures from the same subject; in contrast, the generic detector successfully detected the test seizure because it resembled a seizure contributed by one of the many training subjects. A similar phenomenon may explain why the generic detector is able to identify one of three seizures from subject 23, while the patient-specific detector found none.

6. Alternate detection architecture

In addition to the detector architecture shown in Fig. 1, we also studied the performance of a patient-specific detector with the architecture shown in Fig. 14. Under this alternate architecture, 2-second epochs from each of 21 EEG channels are passed through the wavelet-based feature extractor to compute features characterizing the morphology of each channel's waveform. The four features extracted from each derivation are then assembled into a distinct feature vector and assigned to the seizure or nonseizure class independently of other channels. This is accomplished using a support vector machine trained only on seizure and nonseizure EEGs from the feature vector's source channel. A final decision regarding the onset of a seizure is declared after all classifications are examined in the context of temporal and patient-specific spatial localization constraints. These constraints amount to requiring that seizure activity be present for 6 seconds on half of the EEG channels in one of the groups in Fig. 15. For each subject we automatically select the group whose channels exhibit the greatest difference between training seizure and nonseizure EEGs. The separability of seizure and nonseizure EEGs on a given channel was quantified by noting the percentage of training samples correctly classified by the associated support vector machine. We refer to the architecture in Fig. 14 as the *spatially independent processing* (SIP) architecture, because channels are classified independently, and to the architecture in Fig. 1 as the *spatially dependent processing* (SDP) architecture, because the interdependence of channels is exploited during classification.

The SIP and SDP architectures differ primarily in the stage at which patient-specific constraints on the spatial localization of seizure waveforms are captured or enforced. In the case of the SIP architecture, localization constraints are imposed using an explicit, transparent rule in the final element of the detector. This allows for the independent classification of activity on each channel in a low-dimensional feature space. In contrast, the SDP architecture expresses spatial constraints through the elements of a large feature vector summarizing interrelations between channels. While this obviates the need to explicitly enforce localization constraints, it hides which channels are necessary for detection and results in classification taking place in a higher-dimensional feature space.
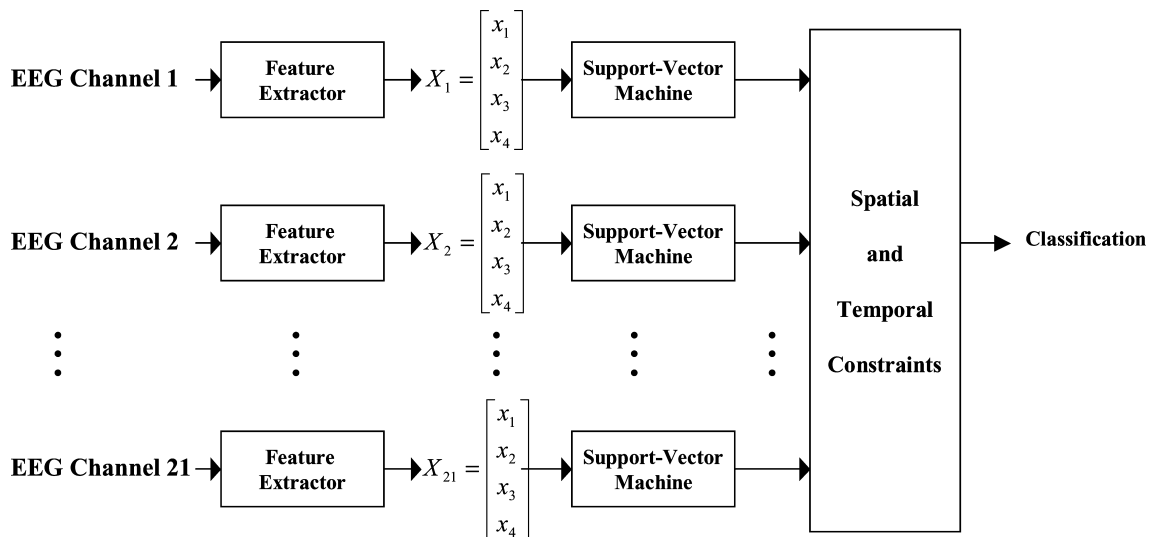

Fig. 14. Alternate detector architecture. Four features representing waveform morphology on each of 21 channels are extracted and classified independently using support vector machines. Seizure onset is declared after the classifications of all support vector machines are examined in the context of temporal and patient-specific localization constraints.
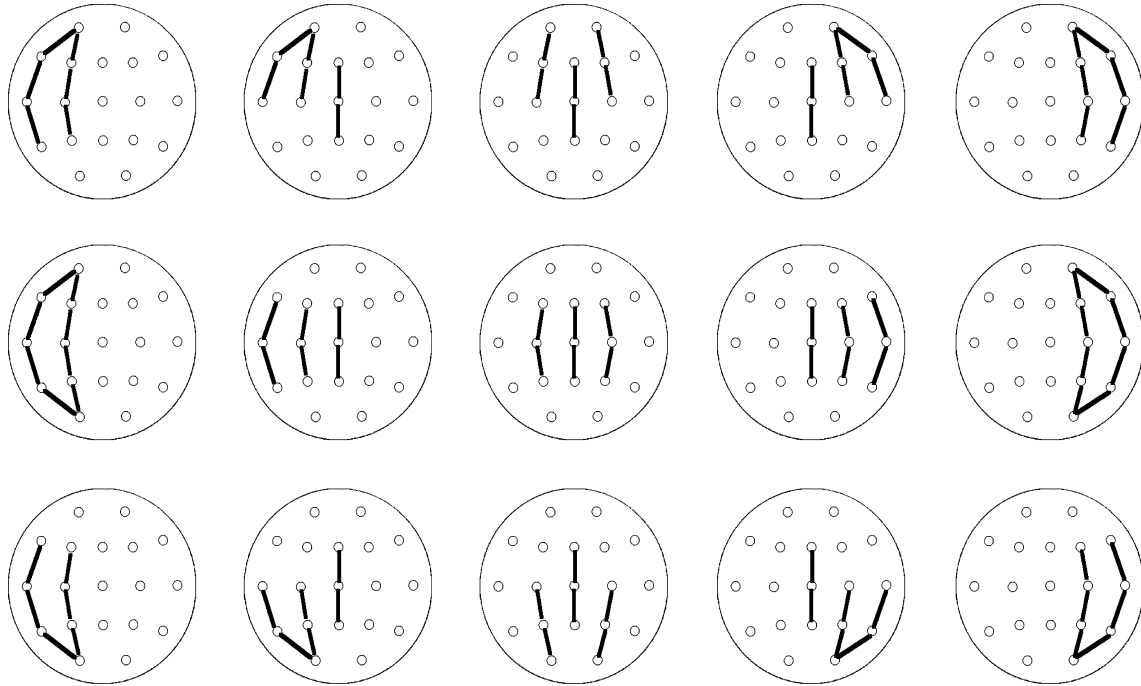
Fig. 15. Groups of EEG channels. Each circle represent an EEG electrode. A line between two electrodes represents an EEG channel whose signal is formed by taking the difference between each electrode's potential.

Fig. 16 compares the performance of the SDP and SIP architectures; performance was assessed using the leave-one-out cross-validation testing scheme. The SDP architecture exhibits a smaller average detection latency and a larger number of true detections relative to the SIP architecture, but a slightly greater number of false detections. The smaller average detection latency of the SDP architecture suggests that the support vector machine, to some extent, was handicapped by the smaller feature vectors in the SIP architecture and is more effective when allowed to freely exploit the interrelations of elements within larger feature vectors.

7. Related work

Seizure detection algorithms developed can be categorized as follows: patient-specific and patient-independent methods that rely on training a classifier to recognize features of seizure and nonseizure EEGs, and patient-independent methods that rely on satisfying a criterion specific to seizure evolution. In this section we review work from each of these categories.

Qu and Gotman [15] exploited patient specificity in detecting seizure onsets from noninvasive EEG. Their method required that an expert specify EEG channels likely to exhibit seizure activity, and used one nearest-neighbor classifier to sequentially classify time and frequency domain features from each channel indepen-dently. This is in contrast with our method, which makes a classification within the context of activity present on all channels using a support vector machine. Qu and Gotman reported a 100% onset detection rate with an average detection delay of 9.35 seconds and a false-positive rate of 0.02 alarm/hour on a data set that included 12 patients and 47 seizures.

Patient specificity was also the cornerstone of the detector described by D'Alessandro et al. [16]. Their method automatically selected both the optimal features and the channels for a given patient. The features include lower-level features from the time domain, frequency domain, wavelet domain, and fractal dimension; and higher-level features that capture the statistical properties of those at lower levels. The method was designed to be used with invasive EEG, and was tested on recordings from only four patients. The algorithm correctly predicted the onset of 62.5% of seizures at a false prediction rate of 0.27 false prediction/hour. We have focused our efforts on designing a robust algorithm for the detection of seizure onsets in noninvasive electroencephalography, which is far more susceptible to artifact contamination than invasive electroencephalography. Our approach to dealing with artifacts is based on training our detector to recognize artifacts, rather than actively remove them using standard signal processing techniques.

Hively and Protopopescu [17] designed a patient-specific algorithm that predicts seizure onsets by inferring from nonlinear dynamical indicators of electroen-
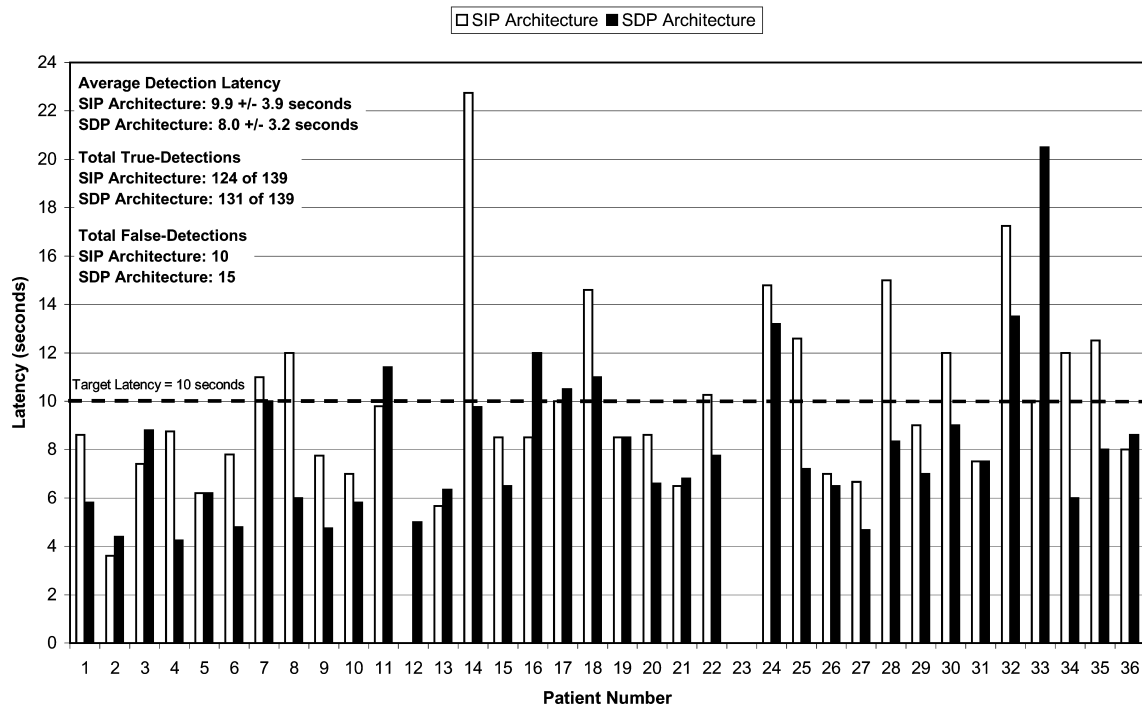
Fig. 16. Comparison of SIP and SDP architectures. Each bar indicates the average latency in detecting a subject's seizure onset using both the SIP and SDP detection architectures. The SDP architecture detects more seizures with a smaller average latency. Both architectures declare a comparable number of false detections. Note that the SDP detection architecture results are the same as those in Fig. 10.

cephalographic evolution the eventual transition of the brain from an interictal to an ictal state. The nonlinear dynamical indicators used by Hively and Protopopescu are called phase-space dissimilarity measures. The algorithm successfully predicted the onset of 87.5% of seizures at some point in a window that precedes the electrographic onset by 60 minutes, and exhibited a false prediction rate of 0.021 false prediction/hour on a dataset of 260 hours of EEG recordings from 41 patients. The long prediction time of this algorithm probably makes it unsuitable for the applications for which our approach was developed.

Patient-independent seizure detection algorithms that rely on machine learning include that of Gabor et al. [18]. They developed a detector that uses a self-organizing map (SOM) neural network to classify feature vectors constructed by transforming spectrograms of 4-second, channel-averaged EEG epochs using a two-dimensional FFT. The channel-averaged EEG epochs are first processed by a wavelet transform that emphasizes time scales prominent in the training seizure examples. This methodology differs from ours in that channel-averaged epochs are classified, as opposed to the classification of a feature vector that preserves the spatial distribution of channels. The algorithm detected 52 of 62 seizures and declared 0.71 ± 0.79 false positive per hour. The authors did not report on how long after electrographic seizure onset the events were declared.

Patient-independent seizure detectors that rely on the properties of seizure evolution include those developed by Schindler et al. and Iasemidis et al. Schindler et al. [19] developed a neuronal cell model, and used it to estimate neural cell spiking rates in different regions of the scalp from noninvasive EEGs. They demonstrated that a persistent elevation in the model's spiking rate could be used to detect the onset of a seizure event. As the model's spiking rate depends on the derivative of the EEG signal exceeding a threshold, it is not clear whether or not this methodology can be used to detect seizures characterized by high-amplitude, slow waves with small-magnitude time derivatives. As our works shows, these seizures are readily detected by a wavelet decomposition that captures long-time scale activity. In a later publication, Schindler et al. [20] demonstrated on nine patients the ability of their method to predict seizure onset 83 ± 91 minutes before its occurrence. This method may be useful for detecting seizure events preceded by an increase in the frequency of spikes, but not those lacking such an increase. This emphasizes the need for a patient-specific seizure detection methodology that does not impose specific requirements on the evolution of seizures.

Iasemidis et al. [21] developed an algorithm that predicts seizure onsets by characterizing the level of entrainment observed in spatially related channels using the Lyapunov exponent nonlinear dynamical indicator. Specifically, the algorithm relies on the observation that long before seizure onset, channels surrounding the

epileptogenic focus behave independently, but as the brain gradually transitions toward the ictal state, the degree of entrainment of these channels increases. When the algorithm was tested on intracranial recordings from five patients using fixed prediction parameter settings, it correctly predicted 82.5% of seizures with an average prediction time of 71.7 minutes and an average false prediction rate of 0.16 false prediction/hour. Whether or not this entrainment is evident in noninvasive electroencephalography, and observable in all seizure types, is not clear.

8. Summary and conclusions

We presented a patient-specific method that detects the onset of epileptic seizures in noninvasive electroencephalography and that can be used to initiate time-sensitive clinical procedures such as the injection of an imaging radiotracer immediately following seizure onset. Our method exploits the conservation of waveform morphology and spatial distribution of seizure and nonseizure activity for a given patient. It is designed to work on a variety of seizure types and was tested on focal, lateral, and generalized seizure onsets.

A multiresolution wavelet decomposition is used to capture EEG waveform morphology. The decomposition and spatial distribution of the waveforms are encoded in feature vectors passed to a support vector machine. The machine, which is trained on nonseizure activity and two or more seizures, is then used to determine whether the feature vector is representative of the patient's seizure activity. Nonseizure activity includes baseline, artifact, and hallmark activity from different states of consciousness.

Our method was tested on noninvasive EEGs from 36 pediatric subjects suffering from a variety of seizure types. It exhibited an average latency of 8.0 ± 3.2 seconds while identifying 131 of 139 seizure events and declared 15 false detections in 60 hours of clinical EEG recordings. We believe our detector can be used to consistently ensure prompt infusion of an imaging radiotracer following seizure onset in a clinical setting, and consequently enhance noninvasive localization of epileptogenic foci using functional imaging modalities.

Acknowledgments

Appendix A. Wavelet decomposition

A multiresolution wavelet decomposition extracts *subband signals* containing activity at different time scales by passing the EEG signal through an iterated filterbank structure like that shown in Fig. 17. The time scale or frequency of activity resolved by a particular subband signal is determined by the iteration level producing it and the choice of analysis filters $H_1(z)$ and $H_0(z)$. The time scale resolved by a subband signal increases with its iteration level, which is equivalent to a decrease in the frequency of the resolved activity.

In the case of the detection algorithm, $H_1(z)$ and $H_0(z)$ were chosen to be the filters associated with the
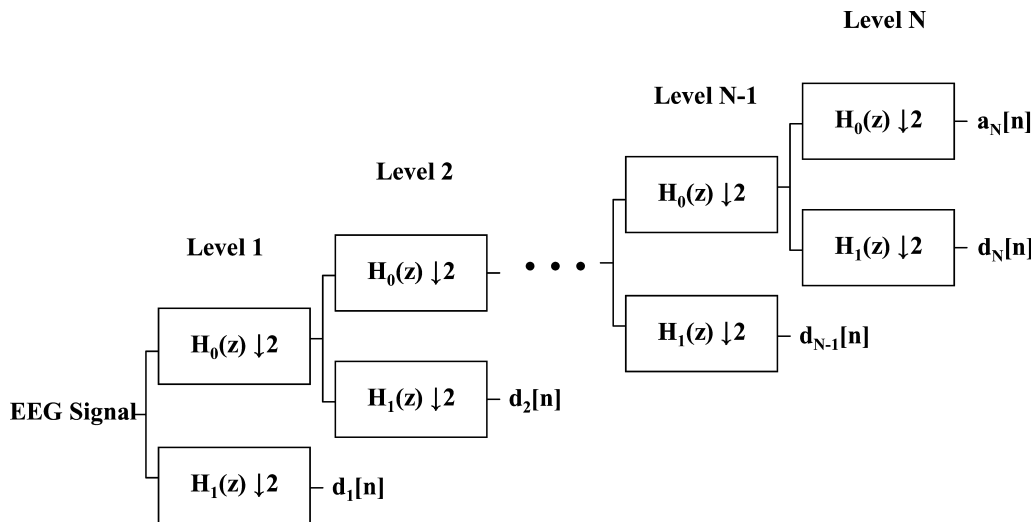

Fig. 17. Multilevel wavelet decomposition filter bank. The time scale of activity captured within each subband signal depends on the choice of analysis filters $H_1(z)$ and $H_0(z)$ and the iteration level producing it. Generally, the subband signals produced by higher iteration levels capture long time-scale activity, while those produced by lower iteration levels capture shorter time-scale activity.
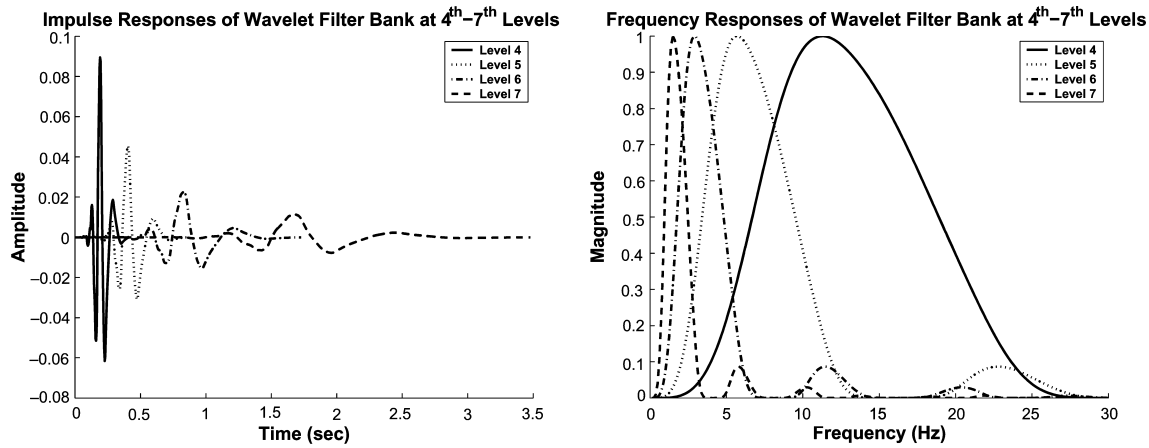
Fig. 18. Effective impulse and frequency responses of wavelet filter bank. Lowest-level subband signal is associated with the shortest time-scale impulse response and widest bandwidth frequency response. Highest-level subband signal is associated with the longest time-scale impulse response and narrowest bandwidth frequency response.

fourth member of the Daubechies wavelet family [22], as those filters exhibit a maximally flat response in their passband as well as little spectral leakage in their stopbands. Furthermore, only the subband signals $\{d_4[n]\ d_5[n]\ d_6[n]\ d_7[n]\}$ are computed because collectively they represent activity at time scales corresponding to frequencies from 0.5 to 25 Hz, which is a frequency that captures seizure onsets of various electrographic manifestations [7]. The remaining subband signals primarily resolve activity of no clinical relevance to this application. In particular, the subband signal $a_7[n]$ captures slow baseline variations like those caused by sweating, while the subband signals $\{d_1[n]\ d_2[n]\ d_3[n]\}$ capture high-frequency artifacts similar to those resulting from muscular contractions. The detector learns the morphology and spatial distribution of artifacts with energy in the subband signals $\{d_4[n]\ d_5[n]\ d_6[n]\ d_7[n]\}$ by their inclusion in the training examples.

The time scales and frequencies captured within each of the subband signals $\{d_4[n]\ d_5[n]\ d_6[n]\ d_7[n]\}$ can be appreciated by examining the impulse and frequency response of the cascade of filters producing each, as shown in Fig. 18. The lowest-level subband signal $d_4[n]$ is associated with the shortest-time-scale impulse response and widest bandwidth frequency response; while the highest-level subband signal $d_7[n]$ is associated with the longest-time-scale impulse response and narrowest bandwidth frequency response.

The subband signals $\{d_4[n]\ d_5[n]\ d_6[n]\ d_7[n]\}$ are not used directly as entries of the feature vector. Using such a direct representation of the EEG waveform is too sensitive to noise and slight variations in morphology. Instead, the energy in each of the subband signals $\{d_4[n]\ d_5[n]\ d_6[n]\ d_7[n]\}$ is used. An explicit representation of the four features computed for each EEG channel $i = 1 \ldots 21$ is:

$$\begin{bmatrix} X_{i,1} \\ X_{i,2} \\ X_{i,3} \\ X_{i,4} \end{bmatrix} = \begin{bmatrix} \log(\sum_n |d_{i,4}(n)|) \\ \log(\sum_n |d_{i,5}(n)|) \\ \log(\sum_n |d_{i,6}(n)|) \\ \log(\sum_n |d_{i,7}(n)|) \end{bmatrix}.$$

References

[1] Leppik I. Contemporary diagnosis and management of the patient with epilepsy. Newtown, PA: Handbooks in Health Care; 2000.

[2] Tyner FS, Knott JR, Mayer WB. In: Fundamentals of EEG technology: basic concepts and methods, vol. 1. Lippincott Wiliams & Wilkins; 1983.

[3] Klass D, Daly D. Current practice of clinical electroencephalography. New York: Raven Press; 1979.

[4] Camfield C. Will I always have epilepsy? [online]. 2004 [cited February 2004]. Available from: URL: http://www.epilepsy.com/101/ep101_duration.html.

[5] Packard AB, Roach PJ, Davis RT, et al. Ictal and interictal 99mTc-bicisate brain SPECT in children with refractory epilepsy. J Nucl Med 1996;37:1101–6.

[6] Davis RT, Treves ST, Packard AB, Farley JB, Amoling RK, Ulanski JS. Ictal perfusion brain SPECT in pediatric patients with intractable epilepsy: a multidisciplinary approach. J Nucl Med Technol 1996;24:219–22.

[7] Gotman J. Automatic recognition of epileptic seizures in the EEG. Electroencephalogr Clin Neurophysiol 1982;54:530–40.

[8] Webber WRS, Lesser RP, Richardson RT, Wilson K. An approach to seizure detection using an artificial neural network. Electroencephalogr Clin Neurophysiol 1996;98:250–72.

[9] Gotman J, Ives JR, Gloor P. Frequency content of EEG and EMG at seizure onset: possibility of removal of EMG artifact by digital filtering. Electroencephalogr Clin Neurophysiol 1981;52:626–39.

[10] James CJ, Hagan MT, Jones RD, Bones PJ, Carroll GJ. Multireference adaptive noise canceling applied to the EEG. IEEE Trans Biomed Eng 1997;44:775–9.

[11] Jung T, Makeig S, Humphries C, et al. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000;37:163–78.

[12] Cristianini N, Shawe-Taylor J. Support vector machines and other kernel-based learning methods. Cambridge, UK: Cambridge Univ. Press; 2000.

[13] Joachims T. SVMLight Support Vector Machine [online]. 2003. Available from: URL: http://svmlight.joachims.org.

[14] Schwaighofer A. Matlab Interface to SVMLight [online]. 2003. Available from: URL: http://www.cis.tugraz.at/igi/aschwaig/software.html.

[15] Qu H, Gotman J. A patient-specific algorithm for the detection of seizure-onset in long-term EEG monitoring: possible use as a warning device. IEEE Trans Biomed Eng 1997;44:115–22.

[16] D'Alessandro M, Esteller R, Vachtsevanos G, Hinson A, Echauz J, Litt B. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. IEEE Trans Biomed Eng 2003;50:603–15.

[17] Hively L, Protopopescu A. Channel-consistent forewarning of epileptic events from scalp EEG. IEEE Trans Biomed Eng 2003;50:584–92.

[18] Gabor A, Leach R, Dowla F. Automated seizure detection using self-organizing neural network. Electroencephalogr Clin Neurophysiol 1996;99:257–66.

[19] Schindler K, Wiest R, Kollar M, Donati F. Using simulated neuronal cell models for detection of epileptic seizures in foramen ovale and scalp EEG. Clin Neurophysiol 2001;112:1006–17.

[20] Schindler K, Wiest R, Kollar M, Donati F. EEG analysis with simulated neuronal cell models helps to detect pre-seizure changes. Clin Neurophysiol 2002;113:604–14.

[21] Iasemidis L, Shiau D, Chaovalitwongse W, et al. Adaptive epileptic seizure prediction system. IEEE Trans Biomed Eng 2003;50:616–25.

[22] Strang G, Nguyen T. Wavelets and filter banks. Wellesley–Cambridge Press; 1997.