Patient-Specific Seizure Onset Detection

by

Ali Hossam Shoeb

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science
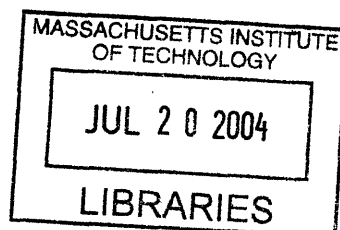
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2003 [September 2003]

Author ....
Department of Electrical Engineering and Computer Science
August 29, 2003

Certified by ....
John V. Guttag
Professor, Computer Science and Engineering
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Patient-Specific Seizure Onset Detection

by

Ali Hossam Shoeb

Abstract

Approximately one percent of the world's population exhibits symptoms of epilepsy, a serious disorder of the central nervous system that predisposes those affected to experiencing recurrent *seizures*. The risk of injury associated with epileptic seizures might be mitigated by the use of a device that can reliably detect or predict the onset of seizure episodes and then warn caregivers of the event. In a hospital this device could also be used to initiate time-sensitive clinical procedures necessary for characterizing epileptic syndromes. This thesis discusses the design of a real-time, patient-specific method that can be used to detect the onset of epileptic seizures in non-invasive EEG, and then initiate time-sensitive clinical procedures like *ictal SPECT*.

We adopt a patient-specific approach because of the clinically observed consistency of seizure and non-seizure EEG characteristics within patients, and their great heterogeneity across patients. We also treat patient-specific seizure onset detection as a binary classification problem. Our observation is a multi-channel EEG signal; its features include amplitude, fundamental frequency, morphology, and spatial localization on the scalp; and it is classified as an instance of non-seizure or seizure EEG based on the learned features of training examples from a single patient.

We use a multi-level wavelet decomposition to extract features that capture the amplitude, fundamental frequency, and morphology of EEG waveforms. These features are then classified using a support vector machine or maximum-likelihood classifier trained on a patient's seizure and non-seizure EEG; non-seizure EEG includes normal and artifact contaminated EEG from various states of consciousness. The outcome of the classification is examined in the context of automatically extracted spatial and temporal constraints before the onset of seizure activity is declared.

During validation tests our method exhibited an average latency of 8.0 ± 3.2 seconds while correctly identifying 131 of 139 seizure events from thirty-six, de-identified test subjects; and only 11 false-detections over 49 hours of randomly selected non-seizure EEG from these subjects. The validation tests also highlight the high learning rate of the detector; a property that allows it to exhibit excellent performance even when trained on as few as two seizure events from the test subject.

3

We also demonstrate through a comparative study that our patient-specific detector outperforms a nonpatient-specific, or generic detector in terms of a lower average detection latency; a lower total number of false-detections; and a higher total number of true-detections. Our study also underscores the likely event of a generic detector performing very poorly when the seizure EEG of a subject in its training set matches the non-seizure EEG of the test subject.

Thesis Supervisor: John V. Guttag
Title: Professor, Computer Science and Engineering

Acknowledgments

I would like to graciously thank:

- The Center for Integration of Medicine & Innovative Technology (CIMIT) for funding this research.

- Dr. Blaise Bourgeois and Dr. Ted Treves for clinically guiding this work.

- Mr. Jack Connolly for teaching me about EEG data formats and recording technologies.

- Mr. Herman Edwards for taking the time to mentor me in the analysis and interpretation of EEG signals.

- Ms. Dorothy Curtis for listening and commenting on all my good and bad ideas.

- Prof. John Guttag for supervising my research with great care, and for always encouraging me to excel.

- My mother, father, and sister for unconditional love, support, and encouragement.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Approximately one percent of the world's population exhibits symptoms of epilepsy, a serious disorder of the central nervous system that predisposes those affected to experiencing recurrent *seizures*. The underlying genetic and molecular mechanisms that give rise to epilepsy are not clearly understood, but the disorder is most common among people in whom the brain has been compromised by some sort of disturbance. In children and young adults, genetic disorders, congenital abnormalities, and birth trauma affecting the brain are most often blamed for the onset of epilepsy; while in middle-aged adults and the elderly, strokes, tumors, and cerebrovascular disease are more frequent explanations [12].

A seizure is a sudden breakdown of the neuronal activity of the brain that precipitates an involuntary alteration in behavior, movement, sensation, or consciousness. Seizures are triggered by a combination of physiological and environmental factors, and can occur sporadically over the course of a week or frequently over the course of a day. The confusion, loss of consciousness, and lack of muscle control that accompanies certain types of seizures can lead to serious injuries that include fractures, head injuries, and burns. These injuries, rather than the seizure events, account for a significant component of the risk associated with epilepsy [28].

The risk of injury associated with epilepsy might be mitigated by the use of a device that can quickly and reliably detect or predict the onset of seizure episodes. For instance the device can abort or prevent the onset of seizures by selective stimulation of brain regions or infusion of anti-epileptic drugs; alert a nearby or remote caregiver of the event; or stimulate patients following the conclusion of a seizure to restart breathing impaired by the seizure episode [28].

Within a hospital a device that detects or predicts seizures would also have numerous applications. The device can potentially be used to identify epileptic events in long-term, non-invasive recordings of brain electrical activity known as electroencephalographic (EEG) recordings, so that the amount of data that must be visually inspected to make a diagnosis is reduced [7]; it may also be used to initiate time-sensitive clinical procedures necessary for the accurate characterization of epileptic syndromes. This final application is the primary motivator for our investigation of quick, reliable and automated seizure onset detection.

The clinical procedure motivating our work is called an *ictal SPECT*. Its aim is to accurately localize each patient's *epileptogenic focus*, the brain region that gives rise to seizure activity. An ictal SPECT localizes the epileptogenic focus by highlighting its altered cerebral blood flow at the onset of a seizure. The accuracy of the localization critically depends on minimizing the delay between starting the ictal SPECT procedure and the *electrographic onset* of a seizure, which is the beginning of sustained abnormal brain electrical activity. This is in contrast to the *clinical onset* of a seizure, which coincides with the beginning of observable abnormal motor or sensory behavior and typically lags the electrographic onset by several seconds.

In today's hospital environment minimizing this delay would require an experienced electroencephalographer to continuously monitor and detect seizure activity in streaming EEG over extended periods of time; such a task is costly, difficult, and mentally taxing and is therefore rarely done. The adopted alternative is to initiate ictal SPECTs following the *clinical onset* of a seizure, and to accept an inherent delay of several seconds. Unfortunately, this delay is very often greatly prolonged as a result of the subtlety of clinical onsets and the physical proximity of nurses carrying out

16

the ictal SPECT procedure. On average ictal SPECTs are started 30-45 seconds [5] after the clinical onset of a seizure, which usually leads to very poor localization of the epileptogenic focus and downgrading of the clinical value of the procedure.

With an automated seizure onset detector, the delay between electrographic onsets and initiation of ictal SPECTs can be consistently minimized without the costly resort to continuous monitoring by an experienced electroencephalographer. This consistency will bolster confidence and increase reliance on ictal SPECTs as a tool for seizure focus localization.

1.2 Goal

The goal of this research is to design, implement, and evaluate a real-time method that can be used to detect electrographic seizure onsets in non-invasive EEG for the purpose of initiating time-sensitive clinical procedures such as ictal SPECT. More generally, the algorithm can serve as the heart of a seizure detection device designed to support epilepsy patients outside of the hospital by alerting nearby or remote caregivers of seizure events; or aborting the onset of seizures via selective stimulation of brain regions or infusion of anti-epileptic drugs.

The detector's performance requirements are defined by the nature of clinical settings and the ictal SPECT procedure. The hectic clinical environment necessitates that the detector require minimal direction from experienced hospital staff. The detector should also initiate all necessary protocols within a maximum of ten seconds of electrographic onset to improve upon the current ability to localize epileptogenic foci using ictal SPECTs. Finally, the detector should exhibit a low false-positive rate, or be easily adjusted to tradeoff more false-negatives for fewer false-positives. This tradeoff is acceptable since incorrect detections lead to costly repeats of the procedure, but misses result in performing ictal SPECTs using existing hospital protocols.

1.3 Approach

We chose to design a patient-specific seizure onset detector because of the clinically observed consistency of seizure and non-seizure EEG characteristics within patients, and their great heterogeneity across patients. Furthermore, we decided to treat patient-specific seizure onset detection as a binary classification problem. In such a problem one 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 a multi-channel EEG signal; its features include amplitude, fundamental frequency, morphology, and spatial localization on the scalp; and it is classified as an instance of non-seizure or seizure EEG based on the learned features of training examples from a single patient. The process of measuring the features of an observation is known as *feature extraction*, and the computational element responsible for determining its class membership is known as a *classifier*.

Training:
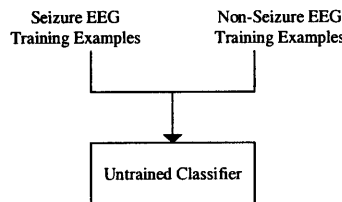
Detection:
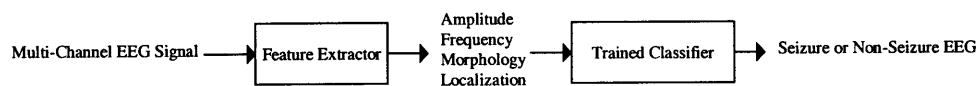
Figure 1-1: Seizure Onset Detection Approach

Figure 1-1 illustrates the problem of patient-specific seizure onset detection in the binary classification framework. The first step in this framework involves a classifier

learning to differentiate between the features of a patient's seizure and non-seizure EEG observations based on training examples. The next step involves using the trained classifier to classify new observations of EEG from the same patient as seizure or non-seizure.

In this thesis we used a multi-level wavelet decomposition to extract features sensitive to the amplitude, fundamental frequency, and morphology of the input EEG signal. We also developed automatic methods to extract spatial localization constraints; and experimented with both support vector machine and maximum-likelihood classifiers.

1.4   Thesis Contributions and Organization

Our research into automatic seizure onset detection makes the following contributions:

- **Provides an Effective Clinical Tool**

  Our detector exhibits the properties of an effective clinical tool. In particular the detector is simple to operate; broadly applicable; and has a high sensitivity and specificity.

  - **Simple To Operate:** To use the detector an electroencephalographer only needs to mark seizure onsets in EEG training records that are approximately thirty-minutes in duration; the number of records maybe as small as one and need not exceed three. The detector automatically combines sections of the training records not marked as seizure with both artifact and generic EEG from various states of consciousness to form a representation of non-seizure activity, and to extract channels to which the seizure onset localizes.

  - **Broadly Applicable:** The detector's use of multi-scale, wavelet-based features allows it to detect seizure onsets with diverse electrographic manifestations. These manifestations include bursts of sharp waves, spike-and-

slow-wave complexes, polymorphic waves, and rhythmic hypersynchrony of variable amplitude and frequency.

– **Highly Sensitive and Specific:** When tested on the seizures of thirty-six de-indentified test subjects, the detector exhibited an average delay of 8.0± 3.2 seconds in correctly declaring 131 of 139 seizure events. Furthermore, the detector only declared 11 false-detections during 49 hours of randomly selected non-seizure EEG from these subjects.

- **Demonstrates Utility of Patient-Specificity in Seizure Detection.**

We demonstrate through a comparative study the improved performance of a patient-specific seizure detector over a generic detector. The patient-specific detector exhibits a lower average detection latency; a lower total number of false-detections; and a higher total number of true-detections. Our study also underscores the likely event of a generic detector performing very poorly when the seizure EEG of a subject in its training set matches the non-seizure EEG of the test subject. These results argue that future research on seizure onset detection should exploit patient-specificity.

- **Provides a Novel Perspective on Artifact Rejection.**

We include various classes of EEG artifacts as part of the detector's non-seizure training set so that they can be identified and avoided through a learning methodology. This is in contrast to the more common approach of removing artifacts using traditional linear, nonlinear, or adaptive signal processing techniques.

- **Compares Alternative Approaches to Seizure Onset Detection**

We designed and compared the performance of two different detector architectures that differ in how they capture and enforce the spatial localization constraints of seizure and non-seizure EEG. We also compared the efficacy of different classification schemes within each architecture; specifically, we experimented with both support vector machine and a maximum-likelihood classifiers.

20

The thesis begins by presenting background on the disorder of epilepsy and the characteristics of normal, abnormal, seizure, and artifact contaminated EEG in Chapter 2 and Chapter 3 respectively. This is followed by a discussion of previous work in the detection and prediction of seizure events in Chapter 4. Chapter 5 examines in detail the computational stages of the detector, and Chapter 6 delves into the results of performance tests. Chapter 7 demonstrates the high learning rate of our patient-specific detector and compares its performance with that of a nonpatient-specific, or generic detector. Finally, Chapter 8 concludes the thesis and outlines directions for future work.

Chapter 2

Epilepsy

2.1 Epilepsy

For centuries epilepsy was considered a damning curse from the gods or a strange type of insanity. Today epilepsy is known to be a neurological disorder of the central nervous system that predisposes individuals to experiencing recurrent seizures. The underlying genetic and molecular mechanisms that give rise to epilepsy remains unknown, but the disorder is most common among people in whom the brain has been compromised by some sort of disturbance. Specifically, in children and young adults, genetic disorders, congenital abnormalities, and birth trauma affecting the brain are most often blamed for the onset of epileptic symptoms; in middle-aged adults and the elderly, strokes, tumors, and cerebrovascular disease are more frequent explanations[1].

People affected by epilepsy do not suffer from an increasingly worsening disorder and are capable of leading normal career and family lives. At the same time, they cannot engage in activities during which a seizure episode could lead to death; for example, driving an automobile. Furthermore, the side-effects of anti-epileptic drugs; episodes of loss of consciousness and motor control; and the public's misconception of the disorder force patients to deal with challenging clinical and psychosocial issues.

---

[1]Material in this chapter is adapted primarily from [12]
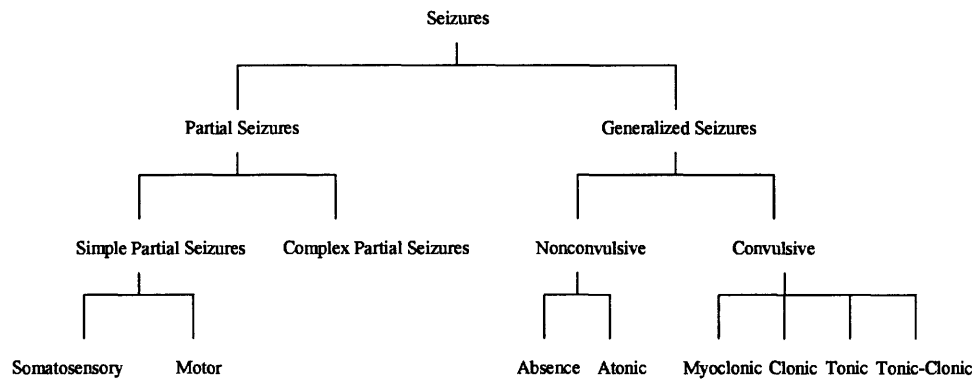
Figure 2-1: Classification of Seizures

2.2 Seizures

A seizure is an involuntary alteration in behavior, movement, sensation, or consciousness resulting from abnormal neuronal activity in the brain. In the case of epilepsy, a malfunctioning region of the brain or the dysfunction of a biochemical mechanism causes the abnormal neuronal activity. This is in contrast to nonepileptic seizures, which are a response to a disturbance external to the central nervous system such as alcohol withdrawal, drug abuse, acute illness, or sleep deprivation.

There are several different types of seizures as shown in Figure 2-1, and the ability to differentiate between them is crucial since each requires a different treatment regiment. The two major seizure types are *parital seizures* and *generalized seizres*. In a partial seizure epileptic activity begins and remains localized in one part of the brain, while in a generalized seizure epileptic activity involves the entire brain from the onset. The sections that follow describe further the clinical and electrographic characteristics of the different seizure types.

2.2.1 Partial Seizures

In a partial seizure epileptic activity begins and remains localized in one part of the brain. Partial seizures that do not affect consciousness are classified as *simple partial seizures*, while those that do are classified as *complex partial seizures*. In the context of epilepsy, impairment or loss of consciousness does not refer to a coma, but rather an individual's lack of understanding and memory of events occurring during seizure episodes.

Simple Partial Seizures

Simple partial seizures do not alter consciousness, but temporarily impair an individual's sensory or motor systems. A simple partial seizure that originates in the somatosensory area of the brain is called a *simple partial sensory seizure*, while one that originates from the motor cortex is called a *simple partial motor seizure*. Individuals typically experience simple partial seizures for less than a minute, and are able to recall events that occurred during the episode.

During simple partial sensory seizures an individual may experience somatosensory, autonomic, or psychic symptoms. Somatosensory symptoms include hallucinations affecting vision, audition, or olfaction; autonomic symptoms include sweating and papillary dilation; and psychic symptoms include sudden sensations of fear, anger, dreamy states, and déjà vu. These clinical manifestations can be very subtle, and are sometimes difficult to distinguish from psychological phenomena. Simple partial motor seizures have clearer clinical manifestations that include rapid muscular jerks and postural movements.

Complex Partial Seizures

Complex partial seizures result in the impairment of consciousness. They are often preceded by auras that include an unusual smell or sensory illusion and are typically accompanied by an automatism such as snapping fingers, picking at clothes, walking aimlessly, mumbling, or lip smacking. After the conclusion of a complex par-

tial seizure, which lasts between 1-3 minutes, individuals will experience a period of confusion lasting several minutes.

2.2.2 Generalized Seizures

In a generalized seizure epileptic activity involves the entire brain from the onset. Generalized seizures whose clinical manifestations include spastic muscle activity are classified as *generalized convulsive seizures*, while those that don't are classified as *generalized nonconvulsive seizures*.

Generalized Convulsive Seizures

The nature of involuntary muscular activity and the individual's state of consciousness during generalized convulsive seizures allows for the further subdivision of this class of seizures into the *myoclonic, clonic, tonic,* and *tonic-clonic* types.

A myoclonic seizure result in unilateral or bilateral rapid alteration of muscular contraction and relaxation, but does not typically alter an individual's state of consciousness. Myoclonic activity is also associated with other neurological disorders, which complicates the classification of this type of seizure.

Clonic seizures exhibit muscular activity similar to that of myoclonic seizures, but with slower cycles of contraction and relaxation. Furthermore, clonic seizures result in the loss of consciousness.

Tonic seizures consist of sudden contraction of truncal and facial muscles accompanied by flexion of upper extremities and extension of lower extremities. These seizures are most common in childhood and may result in serious injuries due to dangerous falls.

Tonic-Clonic seizures combine the clinical manifestation of both the tonic and clonic seizures. These seizures begin without warning with a generalized contraction of muscle groups interrupted by short periods of relaxation. Gradually, these periods become more frequent ultimately leading to rapid muscular contraction and relaxation. Tonic-clonic seizures last between one and two minutes, but individuals may

not regain consciousness until ten to fifteen minutes later and may exhibit symptoms of fatigue for hours or days.

Partial seizures, both simple and complex, that progress to become generalized tonic-clonic seizures are classified as *secondarily generalized seizures.* Sensory or motor auras distinguish between a generalized and secondarily generalized seizures since they are associated only with partial seizures.

Generalized Nonconvulsive Seizures

Absence seizures are generalized nonconclusive seizures that result in the loss of consciousness, eye blinking, staring, and other minor facial movements. These seizures last between a few seconds and a minute and can occur very frequently over the course of a day. Absence seizures are most common in childhood.

Atonic seizures are generalized nonconclusive seizures that do not lead to a loss of consciousness. However, the sudden loss of tone in postural muscles that accompanies atonic seizures leads to dangerous falls that result in serious fractures and injuries to the head.

2.3 Status Epilepticus

Any of the above mentioned types of epileptic seizures may lead to status epilepticus, which is an emergency condition characterized by an epileptic seizure that is so frequently repeated that it is virtually continuous. The condition of status epilepticus can exhibit either convulsive or nonconclusive activity. In nonconclusive status epilepticus an individual appears to be in a coma, while in convulsive status epilepticus an individual experiences repeated generalized tonic-clonic seizures without recovering consciousness.

2.4 Treatment of Epilepsy

Epilepsy affects individuals with variable degrees of severity. Between 70-80% of epilepsy patients suffer from seizures whose severity and frequency can be limited with the use of antiepileptic drugs, each of which essentially limits the capacity of neurons to fire at excessive rates. The correct classification of these patient's seizures is crucial since different seizure types require specific drug regiments. In fact, the use of the wrong antiepileptic drug may exacerbate certain types of seizures.

The remaining 20-30% of epilepsy patients suffer from seizures that are refractory to medication. These patients seek alternative treatment options that include surgery, vagus nerve stimulation, and ketogenic diets.

Surgery

Surgery becomes a viable option for epilepsy patients once a team of epileptologists can accurately identify the region of the brain from which seizures originate. This is accomplished by combining clinical and electrographic evidence from long-term sessions of video and EEG monitoring; anatomical evidence from magnetic resonance imaging; functional evidence from neuropsychological testing; and metabolic evidence from both positron emission tomography (PET) scans and single photon emission tomography (SPECT) scans . The four types of surgery available are removal of a temporal lobe through a temporal lobectomy; removal of cortex through a topectomy; removal of a hemisphere through a hemispherectomy; and separation of the two hemispheres by severing the corpus callosum.

Vagus Nerve Stimulation

Patients that are not surgical candidates may be treated using a vagus nerve stimulator. This implantable, electronic device periodically stimulates the vagus nerve on the left side of the neck. Although the optimal setting for the periodicity and strength of stimulation has not been determined, vagus nerve stimulators can be as effective as antiepileptic drugs in reducing seizure frequency and severity. One of the

28

major applications envisioned for seizure onset detection and prediction algorithms is the modulation of the periodicity and strength of stimulation produced by these devices according to the state of a patient's EEG.

Ketogenic Diet

The ketogenic diet is a high fat, low protein carbohydrate diet that has proven effective in controlling seizures resulting from intractable epilepsy. The diet forces the body to enter *ketosis*, a state in which the brain uses ketones rather than glucose for energy. In this state seizure frequency and severity have been clinically shown to decrease, but the exact mechanism remains unknown.

2.5   Summary

Epilepsy is a neurological disorder of the central nervous system that predisposes individuals to experiencing recurrent seizures. A seizure is an involuntary alteration in behavior, movement, sensation, or consciousness resulting from abnormal neuronal activity in the brain. The two major seizure classes are parital seizures and generalized seizres. In a partial seizure epileptic activity begins and remains localized in one part of the brain, while in a generalized seizure epileptic activity involves the entire brain from the onset.

Between 70-80% of epilepsy patients suffer from seizures whose severity and frequency can be limited with the use of antiepileptic drugs, each of which essentially limits the capacity of neurons to fire at excessive rates. The remaining 20-30% of epilepsy patients suffer from seizures that are refractory to medication, and seek alternative treatment options that include surgery, vagus nerve stimulation, and ketogenic diets.

Chapter 3

Electroencephalogram

This chapter provides background on the recording methodology used to acquire the electroencephalogram (EEG) signal as well as the quantitative variables used to characterize it. Furthermore, this chapter discusses the salient properties of normal, abnormal, seizure, and artifact-contaminated EEG[1].

3.1   Electroencephalogram

The electroencephalogram (EEG) is a non-invasive, multi-electrode recording of time-varying potentials generated by the millions of cortical neurons. The electrodes are distributed symmetrically around the scalp as shown in Figure 3-1 to provide a temporal and spatial summary of brain surface activity; each electrode responds to the aggregate potential generated by many neurons in the area beneath it. EEG activity of clinical relevance is roughly limited to the frequency band 0.5-50 Hz, and that of seizure activity is further limited to the frequency band 0.5-25 Hz. An invasive EEG recording made with electrodes directly in contact with the brain surface is called electrocorticogram (ECoG). ECoGs are not plagued by artifacts and signal attenuation due to the skull as is the case with EEGs, and also provide higher spatial resolution since electrodes responds to the activity of a far smaller number of cortical neurons.

---

[1]The material in this chapter is adapted from [8] and [6]. All images are from [8]
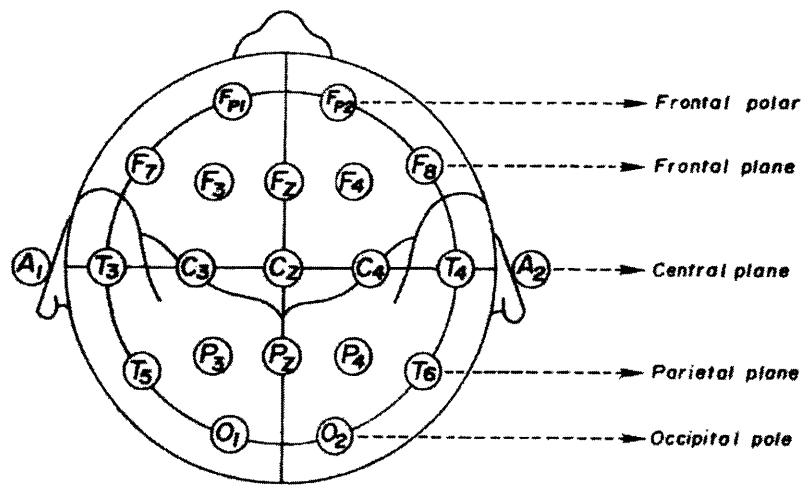
Figure 3-1: 10-20 System of EEG Electrodes Placement

The earliest electroencephalographic recordings were completed and visually characterized in terms of amplitude and frequency content by the Austrian psychiatrist Hans Berger between 1929-1932. Ever since then, the electroencephalogram has been studied and relied on as a clinical tool for the diagnosis of various neurological disorders such as epilepsy.

3.2 Recording EEG

Referential as well as *bipolar* recordings are used for visually reviewing EEG. In a referential recording the potential at each electrode is recorded relative to the potential at either one of the reference electrodes A1 and A2 shown in Figure 3-1. Typically, the electrodes from the left-side of the head are cross-referenced to A2, while those from the right-side of the head are cross-referenced to A1. This scheme ensures that electrodes from each side of the head measure activity relative to a reference that is not greatly affected by cerebral activity within their areas of coverage.

In a bipolar recording the difference between pairs of adjacent electrodes, which are otherwise known as as a *derivation*, is the quantity that is recorded. The longitudinal derivations most commonly viewed by electroencephalographers are shown in Figure 3-2. The electrode at the tip of an arrow is subtracted from that at the tail.
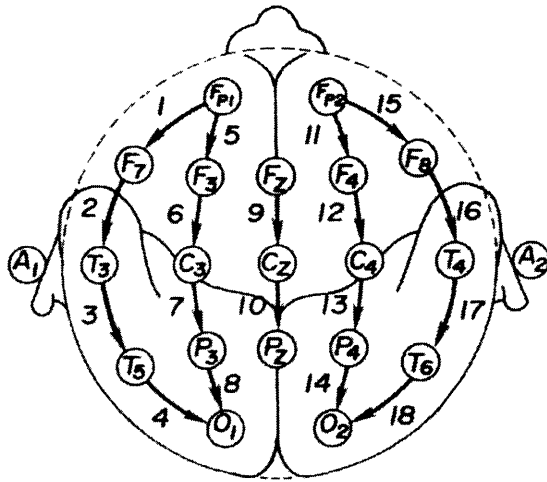
Figure 3-2: Derivations of a Bipolar Recording

An advantage of referential recordings is that a change or abnormality is always clearly observed since the absolute potentials of electrodes, rather than their differences, are the quantities recorded. The disadvantage of referential recordings is that they are very susceptible to common-mode noise as well as contamination of the reference electrode by artifact activity. Once the reference electrode is contaminated it becomes difficult to interpret the activity on electrodes measured against it.

Bipolar recordings overcome common-mode noise by subtracting potentials on contiguous electrodes. The consequence of this operation is a slight attenuation of changes or abnormalities observed in the EEG. An extreme case occurs when a derivations records a zero signal due to cerebral activity that equally affects its electrodes. Nevertheless, we chose to process bipolar EEG signals in our seizure onset detector since a higher resilience to artifacts outweighed the often slight attenuation of activity.

3.3 Characterizing EEG

EEG activity is characterized in terms of several quantitative and qualitative variables that must be considered in the context of a patient's age and state of consciousness.

33

These variables are *fundamental frequency, amplitude, morphology, localization,* and *reactivity.*

Fundamental Frequency

The fundamental frequency of an EEG waveform, measured in Hz, refers to the rate at which the waveform is repeated over a period of a second. The waveform can have an arbitrary shape and any number of subcomponents, all that matters is rate at which the unit as a whole repeats in the span of a second. For instance, the multi-component waveform in Figure 3-6 has a fundamental frequency of 3 Hz. An EEG waveform with a constant, stable fundamental frequency is called *rhythmic,* otherwise it is called *arrhythmic.* Figures 3-3 and 3-4 illustrate examples of these types of waveforms.

Figure 3-3: Rhythmic EEG Waveform
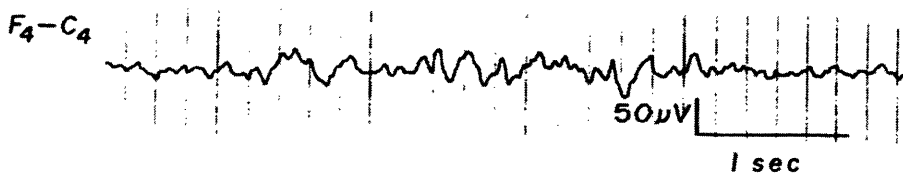
Figure 3-4: Arrhythmic EEG Waveform

Amplitude

The amplitude of a waveform in an EEG trace refers to its peak voltage, which is typically on the order of microvolts. For example, the the waveforms in the EEG trace

34

of Figure 3-3 have amplitudes smaller than $75\mu V$, and those in the trace of Figure 3-6 have an amplitude of approximately $100\mu V$. An EEG waveform demonstrating a sudden or gradual reduction in amplitude, such as that illustrated in Figure 3-8, is said to exhibit *suppression* or *depression*.

Morphology

The morphology of an EEG waveform describes its observed shape, which is a function of the amplitude and fundamental frequency of its constituent components. An EEG waveform that is composed of a single component is called *monomorphic*, and one that is composed of several different components is called *polymorphic*. Examples of these two different morphologies are shown in top and bottom panels of Figure 3-5.
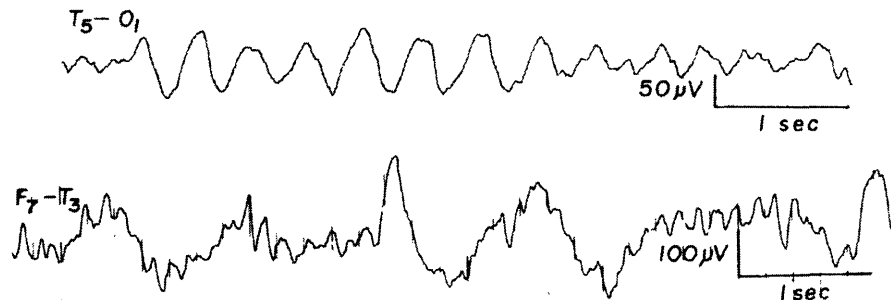
Figure 3-5: Monomorphic (Top) and Polymorphic EEG Waveforms (Bottom)

EEG waveforms consisting of two or more waveforms each with possibly different morphologies are called *complexes*. An example of a commonly observed abnormal complex is the *spike-and-slow-wave complex* shown in Figure 3-6. As its name implies, a spike-and-slow-wave complex is composed of a broad, slow wave and a *transient* spike.

35
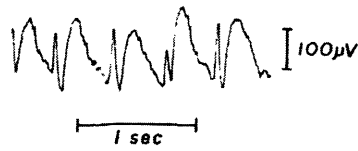
Figure 3-6: Spike-And-Slow-Wave Complex

Localization

The localization of EEG activity refers to the distribution of the activity over the head. EEG activity observed only in a limited region of the head is called *focal*, while activity observed in all regions is called *generalized*. Furthermore, EEG activity exhibiting equal fundamental frequency, amplitude, and morphology on the left and right sides of the head is *symmetric*, otherwise it is *asymmetric*. The clinical designations for different regions of the head are shown in Figure 3-7.
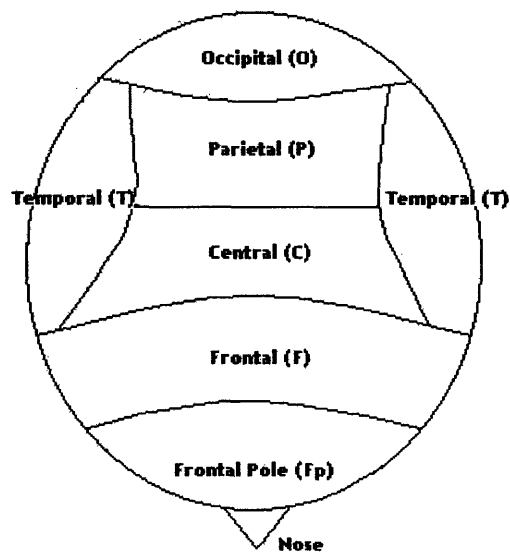
Figure 3-7: Regions of the Head

Reactivity

The reactivity of EEG waveforms refers to the degree of change in anyone of the preceding variables as a result of a stimulus. For instance, Figure 3-8 shows the suppression of 10 Hz occipital activity upon opening of the eyes.
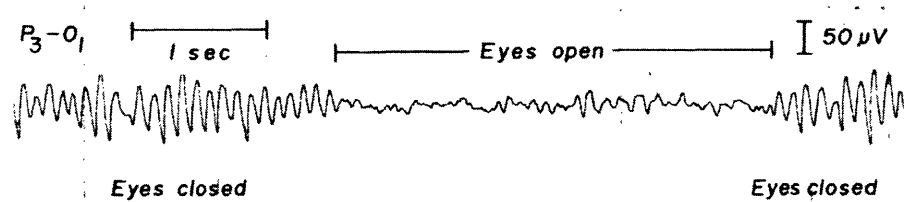
Figure 3-8: Reactivity of EEG Waveforms

3.4 Normal EEG

Normal EEG activity is any activity that qualitatively and quantitatively appears mostly in the EEG of subjects not affected by any disease. The following is a description of well-documented normal EEG activity in adults and children.

Alpha Rhythm

The alpha rhythm is EEG activity with frequency between 8-13 Hz that is prominent in the occipital regions of normal, relaxed adults whose eyes are closed. Alpha activity is attenuated by opening of the eyes, increased vigilance, or heightened awareness as shown in Figure 3-8. The mixture of the alpha rhythm with other rhythms results in *alpha variants*, which have different morphology but otherwise exhibit the same reactivity and localization.

The frequency of alpha rhythms in children gradually increases towards the rate observed in adults over the course of their development. The alpha rhythm may be as slow as 3 Hz at the age of two months and as fast as 7 Hz at the age of one year. Furthermore, the amplitude of alpha rhythms in children steadily increases until the age of one year and then declines towards the $10\mu V$ level observed in adults.

37

Beta Rhythm

The beta rhythm is EEG activity with frequency exceeding 13 Hz that is most promi-
nently observed in the frontal and central regions in adults, but may also be gener-
alized. Alertness and vigilance promotes the onset of beta activity, while voluntary
movement results in its suppression. Figure 3-3 illustrates rhythmic beta activity
recorded from the $F_3 - C_3$ central derivation. The beta rhythm also shows a gradual,
age-related increase in frequency for children.

Theta Rhythm

The theta rhythm is EEG activity with frequency between 4-7 Hz. This activity is
abnormal in awake adults, but commonly observed in sleep and children below the
age of 13 years. Theta activity is asymmetric since it is predominantly observed in
the central, temporal, and parietal regions of the left side of the head. Figure 3-9
shows the theta rhythm artificially placed in context of other normal EEG rhythms.

Delta Rhythm

The delta rhythm exhibits a frequency below 3 Hz and amplitudes that exceed those
of all other rhythms. It is most prominent frontally in adults and posteriorly in
children in the third and fourth stages of sleep. Figure 3-9 shows the delta rhythm
artificially placed in context of other normal EEG rhythms.
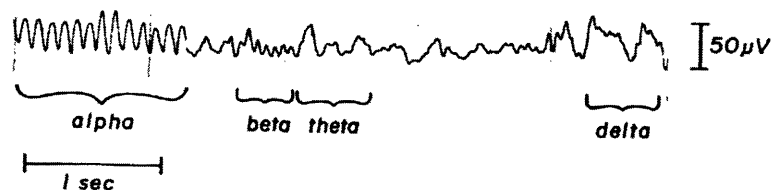
Figure 3-9: Normal EEG Rhythms

38

Mu Rhythm

The mu rhythm refers to EEG activity with frequency between 7-11 Hz that is most prominently observed in the central region. Mu activity is suppressed by movement (fist clenching), imagined movement, or tactile stimulation; in contrast, it is enhanced by immobility and heightened attention. While the frequency range of mu and alpha rhythms overlap, mu rhythms are differentiated by their localization, arch-like morphology, and reactivity. The suppression of mu activity following fist-clenching is shown in Figure 3-10.
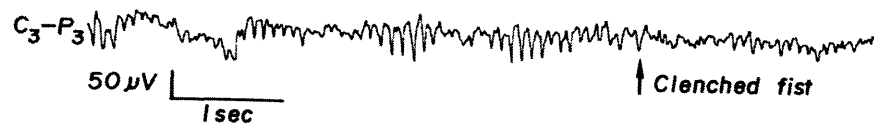
Figure 3-10: Mu Rhythm

Lambda Waves

Lambda waves are transient sharp waves lasting for a duration of approximately 0.25 seconds that occur in the occipital region whenever an adult scans a visual field with horizontal eye movement. Lambda waves are not seen when the eyes are closed, or opened in dark settings. Lambda waves exhibit the same localization and reactivity in children as in adults. Figure 3-11 illustrates several examples of occipital lambda waves.
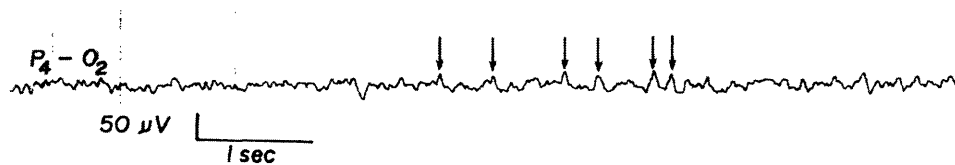
Figure 3-11: Lambda Waves

39

Sleep-Spindles, K-Complexes, and Vertex Waves

Sleep-spindles, K-complexes, and vertex waves are unique waveforms observed only during the four different stages of sleep. The salient characteristics of these waveforms and the four stages of sleep in both adult and children are discussed below.

In the first stage of adult sleep, alpha activity is typically attenuated; theta activity becomes more prominent in the temporal regions; and a series of positive occipital sharp transients may be observed. Deeper into the first stage of sleep *vertex waves*, which are the sharp waves shown in Figure 3-12, begin to appear centrally. For children between the ages of 6 months and 6 years, the first stage of sleep also exhibits high-amplitude bursts of 3-5 Hz waveforms over the central and frontal regions that can last anywhere between several seconds or several minutes. This activity, which is illustrated in Figure 3-13, can be easily mistaken for a seizure without knowledge of the child's state of consciousness.
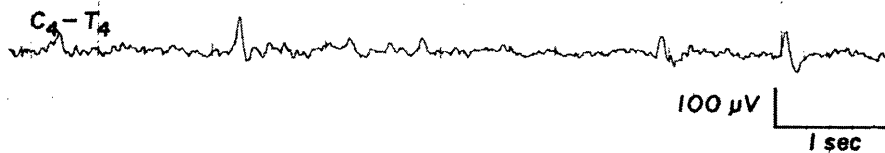
Figure 3-12: Vertex Waves

Figure 3-13: High-Amplitude 5 Hz Bursts in First Stage of Child Sleep

In the second stage, of adult sleep alpha activity is virtually absent; theta activity and vertex waves are more prominent; and rhythmic bursts called *sleep-spindles* with frequencies around 14 Hz appear centrally. Also common in the second stages of sleep are *k-complexes*, which are sharp, slow transients immediately followed by sleep-spindles. Examples of these waveforms are shown in Figure 3-14.

40
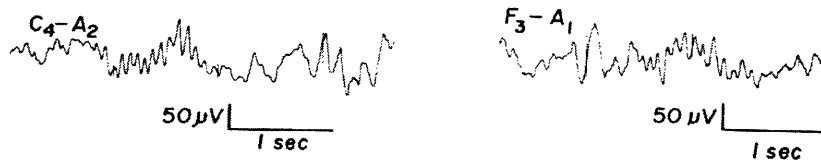
Figure 3-14: Sleep Spindles (Left) and K-Complexes (Right)

Sleep-spindles are absent from the EEG of children until sometime between 6 weeks and 2 months of age. When they first begin to appear in the second stage of sleep, the sleep-spindles of young children will exhibit sharper negative peaks than those of adults. K-complexes remain absent from the second stage of sleep in children until sometime between 3-4 months of age.

In the third stage of sleep, delta activity and slow frontal transients becomes increasingly prominent, while sleep-spindles and K-complexes are observed to a lesser degree. The fourth stage of sleep extends the activity of the third stage with sleep spindles slowing down to a frequency of 10 Hz.

3.5 Abnormal EEG

Abnormal EEG activity is any activity that is prevalent in the EEG of groups of people with neurological or other disease complaints, and absent from that of normal individuals. Abnormal EEG may be an unusual waveform as well as the absence or deviation of normal EEG from well-documented limits on frequency, amplitude, morphology, localization, and reactivity. For instance, an EEG recording exhibiting an absence or change in the nominal frequency and amplitude of sleep-spindles is considered abnormal. The following sections discuss several abnormal EEG waveforms that are commonly observed in the EEG of patient groups. For patients affected by epilepsy, these abnormalities are routinely observed during *interictal* periods, meaning between seizure episodes; however, they do not necessarily result in the clinical behavior observed during a seizure or match its electrographic signature. The characteristics of EEG in *ictal* periods, during seizure episodes, is reserved to section 3.6.

Spike and Sharp Waves

Spike waves are transients with pointed peaks exhibiting durations between 20-70 milliseconds. Sharp waves are similar to spike waves, but exhibit longer durations typically between 70-200 milliseconds as shown in Figure 3-15.
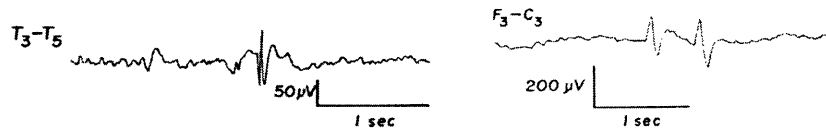
Figure 3-15: Spike Waves (Top) and Sharp Waves (Bottom)

A *spike-and-slow-wave complex* is a spike followed by a longer duration wave as shown in Figure 3-6. Multiple spikes may precede the slower wave and the entire complex may be repeated at a rates of 2.5-6 Hz with intervening periods of quiescence of various durations. A *sharp-and-slow-wave complex* is identical to the spike-and-slow-wave complex except that a sharp wave precedes the slower, broader wave and the complex is repeated at rates between 1-2 Hz.

Periodic Discharges

Periodic discharges refer to time-limited bursts that are repeated at a certain rate. These bursts may exhibit a variety of durations, frequencies, amplitudes, morphologies, and localizations. An example of a periodic discharge is *burst-suppression* activity, which is a discharge of theta or delta frequency waveforms with long intervening periods of very low-amplitude waves. Figure 3-16 shows and instance of burst-suppression activity.
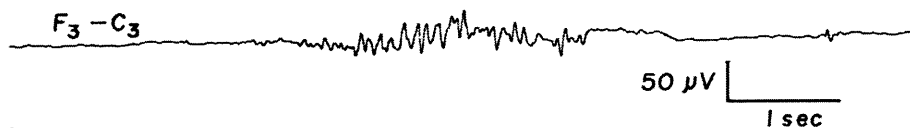
Figure 3-16: Burst-Suppression Activity

Rhythmic Hypersynchrony

Rhythmic hypersynchrony refers to rhythmic activity emerging from a quiescent background and exhibiting unusual frequency, amplitude, morphology and localization of any degree. The rhythmic activity may either be continuous or intermittent. Figure 3-17 shows an example of abnormal, high-amplitude, intermittent 2-3 Hz rhythmic activity on a frontal derivation.
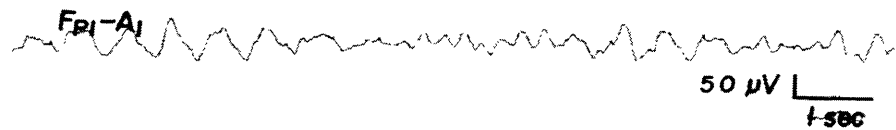
Figure 3-17: High-Amplitude Intermittent 2-3 Hz Activity

Electrocerebral Inactivity

Electrocerebral inactivity refers to a variable length period not caused by instrumental or physiological artifacts that exhibits extreme attenuation of the EEG relative to a patient-specific baseline as shown in Figure 3-18. To appreciate the reduced amplitude of this trace, note that a $10\mu V$ scale, rather than a $50\mu V$ scale, is being used for display. Furthermore, the transients in Figure 3-18 are not of cerebral origin, they are the result of electrocardiographic artifact.
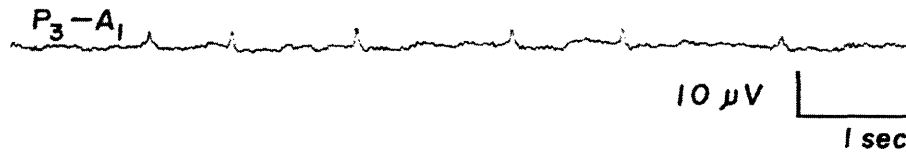
Figure 3-18: Electrocerebral Inactivity

43

3.6 Seizure EEG

Seizures are abnormal, continuous neuronal discharges with clinical correlates that include an involuntary alteration in behavior, movement, sensation, or consciousness. Seizures without clinincal correlates are called *subclinical seizures*. The electrographic signature of a seizure is composed of a continuous discharge of variable amplitude and frequency polymorphic waveforms; spike and sharp wave complexes; rhythmic hypersynchrony; or electrocerebral inactivity observed over a duration longer than the average duration of these abnormalities during interictal periods. Furthermore, the abnormalities observed during interictal periods need not be those that compose the seizure's electrographic signature.

The electrographic signature of a seizure for a given patient is stereotypical and distinguishable from their non-seizure activity. A patient can exhibit more than one type of seizure, however each type will have a stereotypical electrographic and clinical manifestation. The seizures of two different patients can exhibit very distinct morphology and localization; moreover, the characteristics of one patient's non-seizure activity can resemble the seizure activity of another. These clinical observations motivated us to design a patient-specific seizure onset detector, and suggested that the main risk associated with a nonpatient-specific, or generic detector is a high false positive-rate resulting from the similarity of seizure and non-seizure EEG across some patients.

Figure 3-19 illustrates the degree of similarity between two seizure onsets from the same subject. The first seizure onset, shown in the top panel after the dashed line, is characterized by a paroxysmal 10 Hz burst of sharp and monomorphic waves localizing primarily to the central derivations $\{F_Z - C_Z \ \ C_Z - P_Z\}$; the right fronto-central derivations $\{FP_2 - F_4 \ \ F_4 - C_4\}$, and the right frontal derivations $\{FP_2 - F_8 \ \ F_8 - T_8 \ \ T_8 - P_8\}$. The second seizure onset, shown in the bottom panel, matches the activity of the first except for less prominent discharges on the frontal derivations $\{FP_1 - F_7 \ \ FP_1 - F_3 \ \ FP_2 - F4 \ \ FP_2 - F_8\}$.
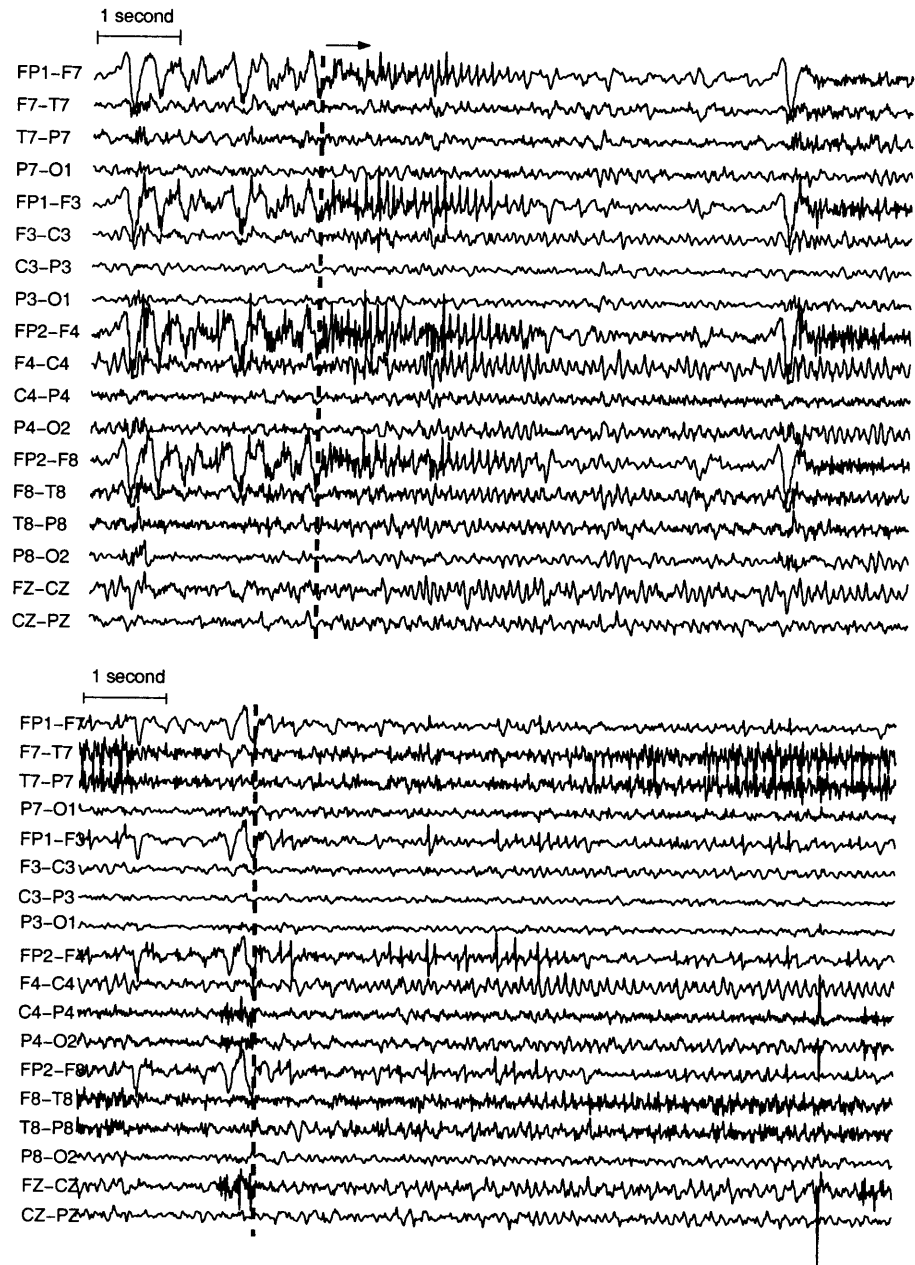
Figure 3-19: Comparing Seizure Onsets From the Same Patient. Dashed line marks electrographic onset of seizure
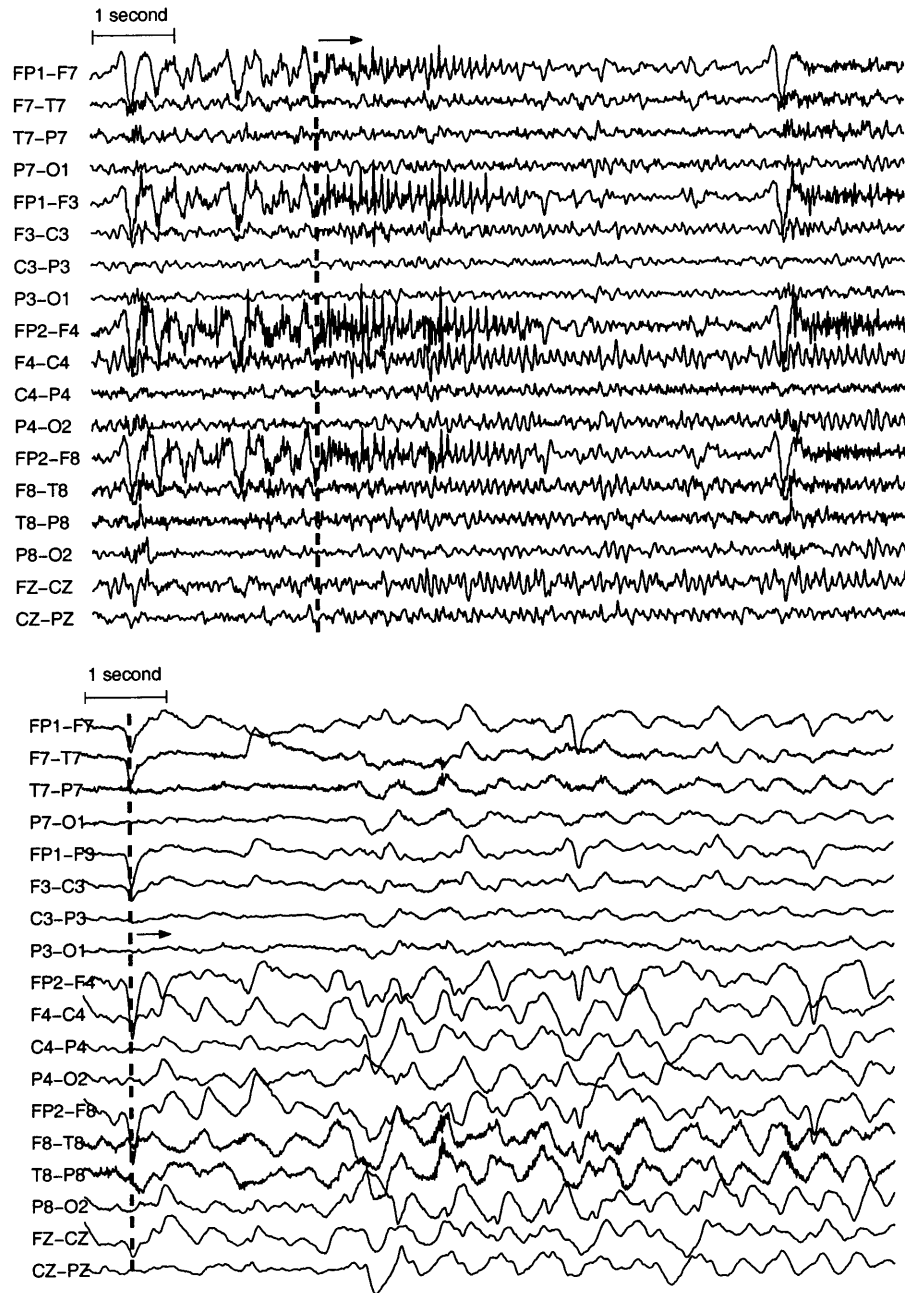
Figure 3-20: Comparing Seizure Onsets From Two Different Patients. Dashed line marks electrographic onset of seizure

Figure 3-20 illustrates the variability in morphology of seizure onsets for two different subjects. The seizure onset in the top panel is characterized by a paroxysmal 10 Hz burst of sharp and monomorphic waves, while the one in the bottom panel exhibits a higher-amplitude, paroxysmal 2 Hz burst of monomorphic waves. Coincidentally, the seizure onsets from both subjects localize to the same derivations.

3.7 Artifacts in EEG

Any electrical activity in EEG that is not of cerebral origin is labelled as an artifact. Artifacts of physiological origin may result from *muscle potentials*, *electrocardiographic potentials*, *eye movement potentials*, *glossokinetic potentials*, and *skin potentials*. Artifacts of nonphysiological origin result primarily from malfunctioning electrodes and electromagnetic interference. Learning the characteristics of these artifacts is crucial for both an electroencephalographer and an automated seizure detector, since artifacts are prevalent in EEG and can be easily confused with both abnormal and seizure activity.

3.7.1 Physiological Artifacts

Muscle Potentials

Artifacts caused by muscle potentials are very common in EEG recordings. They appear as high-frequency bursts in the frontal and temporal electrodes of a bipolar recording, and in all electrodes of a referential recording that uses the ear, chin, or mandible as a reference. Although muscle artifacts can never be completely eliminated, they can be attenuated with the use of a high frequency filter that limits the EEG bandwidth to 35 Hz activity. The risk associated with this strategy is that highly filtered muscle activity may be mistaken for normal beta activity. Figure 3-21 illustrates the high frequency activity associated with muscle artifacts.

Figure 3-21: Muscle Artifact

Electrocardiographic Potentials

Electrocardographic artifacts are those produced by the electrical activity of the heart. They resemble attenuated periodic sharp waves in both referential and bipolar recordings. Electrocardiographic artifacts cannot be easily removed through filtering, but can be distinguished from EEG activity by noting that their period perfectly matches the period of an accompanying EKG signal. Figure 3-22 shows the sharp waves associated with electrocardiographic potentials.
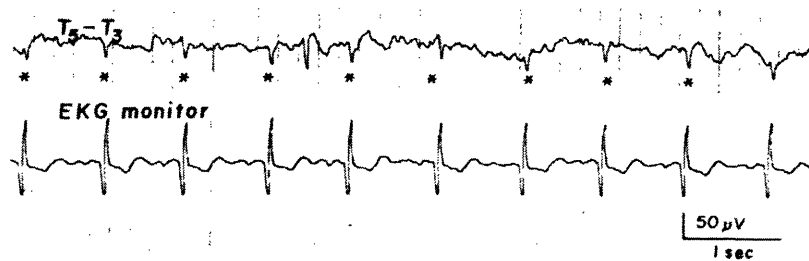
Figure 3-22: Electrocardiographic Artifact

Eye Movement Potentials

Eye movement, eye blinking, and eyelid fluttering gives rise to artifacts resembling transient or rhythmic EEG slow waves. These artifacts appear most prominently in the frontal channels of both bipolar and referential recordings, and can possibly be distinguished from EEG activity of frontal cerebral origin by the addition of electrodes around each eye. However, the extra electrodes are not often used in clinical practice and were not available to our detector. The mixture of eye movement and electrocardiographic artifacts results in rhythmic frontal activity with sharp and slow

48

components. Figure 3-23 illustrates the low frequency activity associated with eye blinking and the higher frequency activity associated with eye fluttering.
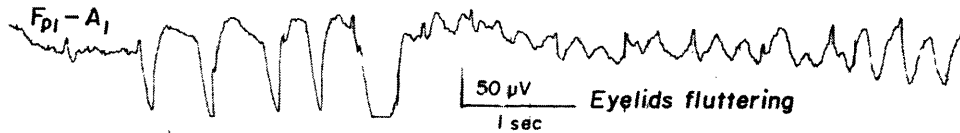
Figure 3-23: Eye Movement Artifact

Glossokinetic Potentials

Artifacts generated by glossokinetic potentials refer to artifacts generated by movement of the tongue. These artifacts appear as single rhythmic slow waves in the temporal regions and can be recognized by the addition of electrodes near the mouth. Chewing and sucking movements mix artifacts generated by muscle potentials and glossokinetic potentials, and can be identified by the addition of electrodes near the jaw. Finally, hiccups and sobbing can generate glossokinetic potentials that may appear in EEG as abnormal spike-and-wave discharges. Figure 3-24 shows the mixture of slow, fast, and spike activity resulting from glossokinetic and muscle potentials caused by chewing.
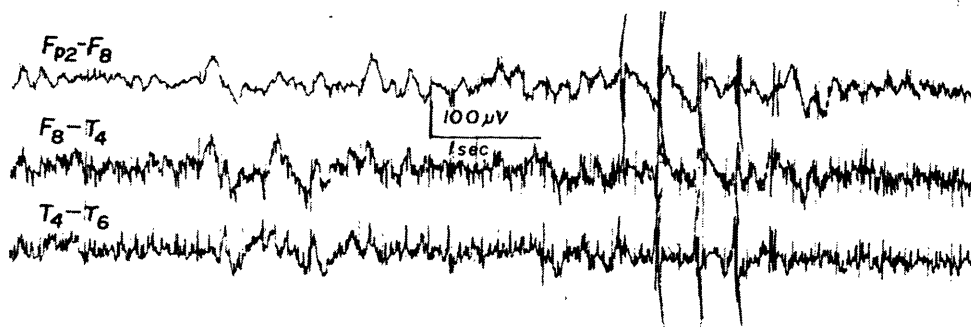
Figure 3-24: Chewing Artifact

49

Skin Potentials

Changes in skin potential produce low frequency baseline changes in the EEG. The potential of skin may change as a result of the electrical potential generated by active sweat glands, or because of sweat-related changes in electrolyte concentration between the skin and the EEG electrodes. Figure 3-25 shows a less than 1 Hz baseline variation in the referential recording of an F_7 electrode displayed on a 2 second, $50\mu V$ scale.
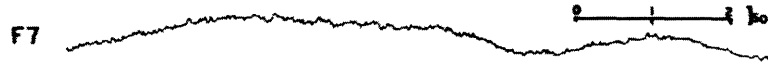
Figure 3-25: Low Frequency Baseline Change Caused by Sweat

3.7.2 Nonphysiological Artifacts

Electrodes that are poorly coupled mechanically or electrically to the skin can produce artifacts resembling EEG sharp waves, spike waves, or slow waves. Movement of the wires connecting electrodes to the EEG instrument simulates slow, rhythmic EEG activity with a frequency matching the movement of the wires.

Electromagnetic interference that is coupled electrostatically or inductively to recording electrodes can mask the underlying EEG activity. An example of this type of interference is 60 Hz and high frequency radiation from surrounding electronic and radio equipment. Furthermore, the movement of personnel around the wires of EEG electrodes generates electrostatically coupled artifacts that appear as high amplitude rhythmic waves as shown in 3-26.
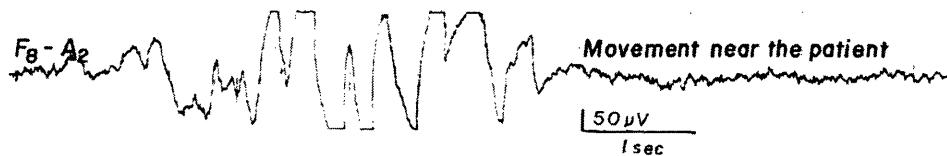
Figure 3-26: Artifact Caused by Movement Around Subject

3.8 Summary

The electroencephalogram (EEG) is a non-invasive, multi-electrode recording of time-varying potentials generated by the millions of cortical neurons. The electrodes are distributed symmetrically around the scalp to provide a temporal and spatial summary of brain surface activity. EEG activity is characterized by its fundamental frequency, amplitude, morphology, localization, and reactivity.

The alpha, beta, theta, delta, mu, and lambda rhythms are types of EEG activity observed in the normal EEG of adults and children; they are differentiated by their unique frequency, morphology, localization, or reactivity. Abnormal EEG is an unusual waveform as well as the absence or deviation of normal EEG activity from well-documented limits.

During interictal periods, or between seizure episodes, the EEG of patients affected by epilepsy will exhibit abnormalities like spike and sharp waves; periodic discharges, rhythmic hypersynchrony; or electrocerebral inactivity. In ictal periods, or during seizures, the EEG is composed of a continuous discharge of one of these abnormalities, but extended over a longer duration and typically accompanied by a clinical correlate.

The electrographic signature of a seizure for a given patient is stereotypical and distinguishable from their non-seizure activity. On the other hand, the seizures of two different patients can exhibit very distinct morphology and localization; moreover, the characteristics of the first patient's non-seizure activity can resemble those of the second patient's seizure activity. These clinical observations motivated us to design a patient-specific seizure onset detector; they also suggested that the main risk associated with a nonpatient-specific, or generic detector is a high false positive-rate resulting from the similarity of seizure and non-seizure EEG across some patients.

EEG is plagued by artifacts and signal attenuation due to the skull. Artifacts of physiological origin may result from muscle potentials, electrocardiographic potentials, eye movement potentials, glossokinetic potentials, and skin potentials. Artifacts of nonphysiological origin result primarily from malfunctioning electrodes and electromagnetic interference. Learning the characteristics of these artifacts is crucial for

both an electroencephalographer and an automated seizure detector, since can be easily confused with both abnormal and seizure activity.

Chapter 4

Related Work

The accurate diagnosis and characterization of epileptic syndromes involves clinical studies that require instant as well as extended interaction with a patient. For instance, electroencephalographic and video recordings lasting days or weeks are crucial for capturing physical and electrographic evidence of epilepsy, while ictal imaging studies require attending to the patient prior to, or immediately following the seizure's electrographic onset. These different clinical constraints have lead to the development of two types of EEG signal processing algorithms: those aimed at off-line detection and characterization of abnormal EEG activity in long-term recordings, and those aimed at on-line seizure detection and prediction.

4.1   Off-Line Detection of Seizures and Transients

Locating epileptiform activity in the form of seizures or transient abnormalities in EEG recordings lasting days or weeks is an arduous, time-consuming task because this activity constitutes a small percentage of the entire recording. This difficulty has motivated the development of automated systems that scan, identify, and then present to an electroencephalographer epochs containing epileptic events.

Gotman and Gloor [16] designed and implemented one of the earliest automated systems for the identification of epileptic activity in long-term EEG recordings. Their system detected epileptiform spikes and sharp waves on sixteen bipolar channels by

applying empirically determined thresholds to time-domain features. These features include the EEG wave duration, sharpness, and amplitude relative to a dynamically updated baseline. Gotman and Gloors system also rejected artifacts due to muscle activity and rapid eye blinks by noting the frequency, duration, and derivations characterizing these artifacts. The authors presented many anecdotal examples demonstrating detection of epileptic transients and rejection of artifacts, but did not perform a systematic evaluation of the method. Furthermore, the authors did not statistically justify the chosen detection thresholds. Nevertheless, the features and artifact rejection schemes proposed by Gotman and Gloor have proven to be well suited for the detection of epileptiform transients, and have been included in many later systems.

To bypass manually determining detection thresholds, Tarassenko et al. [20] trained a neural network classifier to detect epileptic spike and sharp waves using both time-domain and frequency-domain features. The time-domain features were similar to those used in [16], while the frequency domain features captured the dominant frequencies of the EEG signal. Tarassenko reported a sensitivity between 83-97% and a specificity between 85-95% when their classifier was trained and tested on various patient-specific data sets. The classifier exhibited similar sensitivity when trained and tested on data sets that combine many patients, but a much greater number of false-detections due to muscle bursts, sleep spindles, and vertex waves. The authors argued that increasingly sophisticated processing would not lower false positives significantly, rather the integration of contextual information (patient-awake, asleep, or moving), spatial information (comparison of different channels), and training sets that combine both wakeful and asleep EEG would be a more effective strategy.

Glover et al. [13] demonstrated that false-positives could be reduced by integrating contextual information into an automated system for the detection of epileptiform transients. The authors implemented a system that examines cardiac activity using the electrocardiogram (EKG); muscular activity using the electromyogram (EMG); and ocular activity using the electrooculogram (EOG) signals before confirming the presence of an epileptiform transient in any of twelve available EEG channels. The

rules used to interpret relations between these signals were drawn from an electroencephalographer and encapsulated into a rule-based decision system. The automated system was tested on two patients and demonstrated successful rejection of a large number of epileptiform-like transients caused by sleep spindles, muscle movement, eye-blinks, and electrocardiographic artifacts.

Gotman [14] also developed one of the earliest systems for the automatic detection and recording of seizure events in long-term EEG recordings. The system was designed to identify seizures with paroxysmal rhythmic activity at any point in their evolution. No attempt was made to detect seizures consisting of several mixed frequencies, low amplitude fast activity, or spike-and-slow-wave bursts. Gotman's system declared the presences of a seizure when time-domain features from a minimum of two EEG channels exceeded empirically determined thresholds. The time-domain features were EEG wave rhythmicity as well as amplitude relative to a dynamically updated baseline. The system rejected artifacts due to muscle activity and movement using techniques similar to those in [16]. Gotman's system was tested on 16 recordings with an average length of 12.4 hours, and exhibited considerable variability in performance across patients. Gotman reported that 22% of all detections were caused by epileptiform discharges, 58% were due to non-epileptiform discharges, and 20% were due to artifacts. He also noted that tailoring the detection thresholds to suite the EEG of each patient resulted in improved performance.

4.2   On-Line Detection and Prediction of Seizures

Interacting with a patient just prior to, or immediately following the onset of a seizure requires continuous monitoring of the patient as well as the EEG in order to capture the earliest clinical and electrographic signs of a seizure's actual or imminent onset. This difficult task has encouraged the development of automated systems that can quickly and reliably alert the patient, caregiver, or clinician of ongoing or future electrographic events.

4.2.1 Noninvasive Methods

Qu and Gotman [11] designed one of the earliest patient-specific seizure onset detection algorithms. The algorithm used a nearest-neighbor classifier to differentiate between a patient's normal and abnormal EEG on a manually-selected list of EEG derivations. The EEG signal was described using a total of six weighted time and frequency domain features extracted from 2.56 sec epochs. These features included average EEG wave amplitude and duration, as well as the frequency and width of the dominant spectral peak. 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 alarms/hr on a dataset that included 12 patients and 47 seizures. For each patient the detector was trained on the first twenty seconds of a seizure and 40 minutes of normal EEG sampled over a 24-hour period; then the trained algorithm was tested on 2-6 seizures from the same patient and 160 minutes of normal EEG sampled over another 24 hour period. The authors demonstrated worse performance when their algorithm was made semi-patient-specific (ie. trained on seizure EEG from the test patient and normal EEG from several other patients) as well as generic (i.e. trained on seizure and normal EEG from many different patients). Finally, Gotman showed in an earlier publication [10] with a different version of the algorithm the possibility of reducing false-positive rates at the expense of a higher number of missed seizures by including patient-specific, seizure-like interictal activity in the classifiers training set.

Boashash el al. [3] developed an algorithm for the detection of seizures in newborn EEG that is based on a computational model of cortical brain activity derived from histological and biophysical principles. In particular, the algorithm dynamically estimates the parameters of the model from 10 second epochs of EEG, and then uses the updated model to estimate the energy contribution of seizure and normal activity in the observed EEG epoch. If the ratio of the energy in the seizure component to the energy in the normal component exceeds a certain threshold a seizure is detected. When this approach was evaluated on two channels of EEG data from two newborn babies, a 76% detection rate and 15.6% false-positive rate was observed,

but there was no mention of the average detection latency. Since this method does not detect seizures by thresholding the values of extracted features, no training on patient-specific or generic data sets is required. The computations used by this approach are very intensive, which limits the algorithm to processing one EEG channel and forgoing the use of spatial information from other channels.

Hively and Protopopescu [17] designed an algorithm that predicts seizure onsets by inferring from nonlinear dynamical indicators of the spatio-temporal evolution of EEG relative to a patient-specific baseline, 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-predictions/hour on a dataset of 260 hours of EEG from 41 patients. The authors recognize that only limited conclusions can be drawn from their results since their data was not partitioned into independent sets for the purposes of testing and selection of generic prediction parameters.

4.2.2   Invasive Methods

Meng et al. [19] designed and implemented a seizure onset detection algorithm that processes one manually-selected channel of an ECoG (electrocorticogram, or invasive EEG) recording. The algorithm used a maximum-likelihood classifier with gaussian mixture model conditional densities to differentiate between a patient's normal and abnormal ECoG. The ECoG was described using 24 features extracted from the subband signals of a three-level wavelet decomposition. The features include various measures of the energy in the subband signals; for instance, the percentage and gradient of the energy. For a 0% false-positive rate, the detector exhibited slightly less than a 60% detection rate and an onset detection latency of 5 seconds. Also, when the algorithm was made patient-specific a 0% false-positive rate resulted in a 80% detection rate and an average prediction of seizure onset by 5 seconds. The algorithm was trained on 19 patients and tested on another 18 patients.

D'Alessandro et al. [22] designed a patient-specific algorithm that automatically selects both the optimal features and channels for the prediction of seizure episodes 10 minutes prior to the electrographic onset. 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 optimization over the selection of features and channels was accomplished using a genetic algorithm and an objective function over the training error. When tested on intracranial recordings from four patients, the algorithm correctly predicted the onset of 62.5% of seizures at a false-prediction rate 0.27 false-predictions/hour.

Iasemidis et al. [18] 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 towards 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-predictions/hour.

4.3 Comparison

Our non-invasive, patient-specific seizure onset detector differs from that presented by Qu and Gotman [11] in several important respects. Our detector does not require that an experienced electroencephalographer identify EEG derivations relevant to detecting a patient's seizure onset; instead, derivations are selected automatically. Furthermore, our detector uses wavelet-based features to capture the detailed morphology of seizure and non-seizure EEG rather than general features like EEG waveform amplitude, duration, and dominant frequency. Finally we preserve the spatial properties of the EEG signal either through classifying activity on derivations independently and then imposing spatial constraints, or by simultaneously classify-

ing activity from all derivation in a single context that is sensitive to their spatial relations. This contrasts with Qu and Gotman's detector which classifies activity on derivations independently and in a context blind to their spatial relations.

4.4 Summary

An impressive array of signal processing and pattern recognition techniques have been used in algorithms meant to process EEG and ECoG signals. Furthermore, extensive research has gone into the design of two types of EEG signal processing algorithms: those aimed at off-line detection and characterization of abnormal EEG activity in long-term recordings, and those aimed at on-line seizure detection and prediction.

Chapter 5

Seizure Onset Detection

5.1 Overview

We chose to design a patient-specific seizure onset detector because of the clinically observed consistency of seizure and non-seizure EEG characteristics within patients, and their great heterogeneity across patients. Furthermore, we decided to treat patient-specific seizure onset detection as a binary classification problem. In such a problem one 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 a multi-channel EEG signal; its features include amplitude, fundamental frequency, morphology, and spatial localization on the scalp; and it is classified as an instance of non-seizure or seizure EEG based on the learned features of training examples from a single patient.

The block diagrams in Figure 5-1 present two different processing architectures for the patient-specific seizure onset detector. Under both architectures, a two-second epoch from each of twenty-one bipolar derivations is individually passed through a feature extractor in order to compute four features characterizing the amplitude, fundamental frequency, and morphology of its waveforms. In the Spatially Independent Processing (SIP) architecture, the four features extracted from each derivation are assembled into a distinct *feature vector* and assigned to the seizure or non-seizure class independently of other derivations. This is accomplished by way of a classifier

trained only on the EEG of the feature vector's source derivation. 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 as discussed in section 5.4. In the Spatially Dependant Processing (SDP) architecture, the features extracted from all derivations are grouped into a large feature vector that captures whatever interdependencies exist between derivations. This feature vector is then assigned to either the seizure or non-seizure class by way of a classifier trained on EEG from all the derivations. Finally, seizure onset is declared once the classification result satisfies the temporal constraint discussed in section 5.4.

Spatially Independent Processing (SIP) Architecture   Spatially Dependant Processing (SDP) Architecture
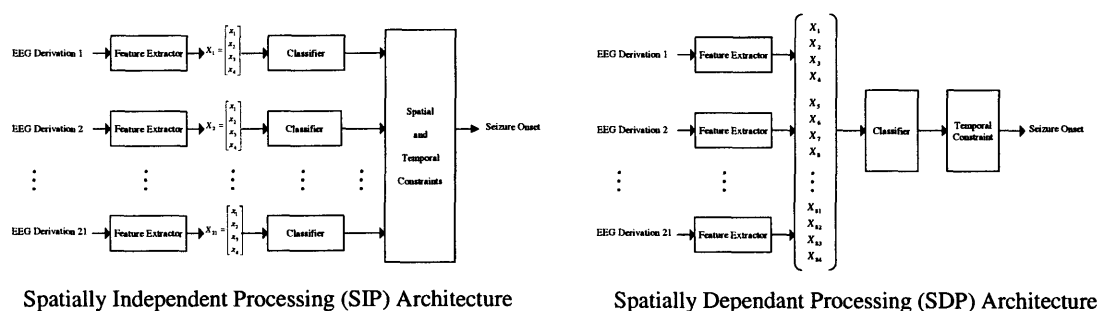
Figure 5-1: Seizure Onset Detector Architectures

The SIP and SDP architectures differ primarily in the stage in which patient-specific spatial localization constraints are captured or enforced. In the case of the SIP architecture, localization constraints are imposed using explicit rules in the final element of the detector. This allows for the independent classification of activity on each derivation 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 derivations. While this obviates the need to explicitly enforce localization constraints, it hides from the user which derivations are being used for detection; and causes classification to take place in a higher dimensional feature space.

The following sections explore the computational elements of the seizure onset detector. Specifically section 5.2 discusses how EEG waveforms are analyzed in or-

der to extract features characterizing their amplitude, fundamental frequency, and morphology, while section 5.3 discusses how the class membership of feature vectors under both architectures is determined using patient-specific seizure and non-seizure training examples. Section 5.4 outlines the temporal and patient-specific localization constraints used in the SIP architecture to determine whether or not classified feature vectors are indicative of seizure onset. The patient-specific training examples are discussed in section 5.5, and detection case studies are presented in section 5.6. A discussion and comparison of the performance of both architectures is in Chapter 6.

5.2 Feature Extraction

An electroencephalographer relies on alterations in amplitude, fundamental frequency, and morphology to discriminate between normal and seizure EEG activity on a single derivation. This naturally suggests that features capable of expressing these quantities with a high-degree of fidelity, and which are efficient to compute are crucial for accurate, real-time seizure onset detection. These requirements are to a great extent satisfied by features produced using a multi-level wavelet decomposition [9] of the EEG signal. Section 5.2.1 discusses the wavelet decomposition scheme while section 5.2.2 outlines how the results of the decomposition are used to construct feature vectors.

5.2.1 Multi-level Wavelet Decomposition of EEG Signals

A multi-level wavelet decomposition of an EEG waveform extracts *subband signals* containing components contributing to the waveform morphology at specific timescales. For instance, a spike-and-slow-wave pattern can be decomposed into a subband signal containing the short time-scale (high-frequency) "spike" component; and another subband signal containing the long time-scale (low-frequency) "wave" component as illustrated in Figure 5-2. Fourier analysis of the same pattern would not be as sensitive to the short time-scale "spike" component because it provides a description of a signal's global regularities, rather than its local, singular irregularities

63

or *non-stationarities*[1]. More generally, the wavelet transform is better suited for analyzing non-stationary signals like the EEG in comparison to the Fourier transform, which assumes signal stationarity. Our detector demonstrated higher sensitivity and specificity using wavelet-based features as opposed to fourier-based features.
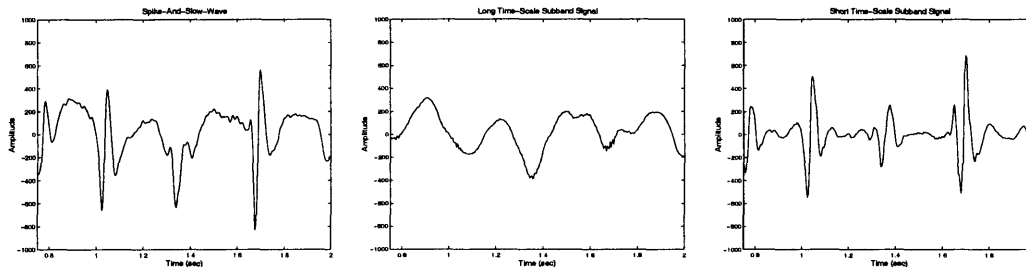
Figure 5-2: Wavelet Decomposition of Spike-And-Slow-Wave Signal

The subband signals of a multi-level wavelet decomposition are computed by passing the EEG signal through an iterated filterbank structure linked by downsampling operations ($\downarrow$ 2) as shown in Figure 5-3. The time-scale or frequency of activity captured by a particular subband signal is predetermined by the iteration level producing it and the choice of analysis filters $H_1(z)$ and $H_0(z)$. Generally, the time-scale resolved by a subband signal increases the higher its iteration level, which is equivalent to the frequency of the resolved activity decreasing.

In the case of the detection algorithm, $H_1(z)$ and $H_0(z)$ were chosen to be the filters associated with the fourth member of the Daubechies wavelet family since those filters are only four taps long and exhibit a maximally flat response in their passband as well as little spectral leakage in their stopbands. Furthermore, only the subband signals $\{d_4 \, d_5 \, d_6 \, d_7\}$ are computed because collectively these signal faithfully represents activity at time-scales corresponding to frequencies between 0.5-25 Hz; which is a frequency band Gotman et al. [15] has shown captures seizure onsets of various electrographic manifestations. The remaining subband signals primarily resolve activity

---

[1]This argument can be further developed by noting that the basis functions used to compute the coefficients of a Fourier transform are sinusoids with infinite temporal extent, as opposed to the basis functions of the wavelet transform which are localized and limited in time. Basis functions with localized and limited temporal extent are naturally more suited for representing short time-scale non-stationarities.

64

Level N

Level N-1

$H_0(z) \downarrow 2$ a_N

Level 2

$H_0(z) \downarrow 2$

$H_1(z) \downarrow 2$ d_N

Level 1

$H_0(z) \downarrow 2$

$H_0(z) \downarrow 2$ • • •

$H_1(z) \downarrow 2$ d_{N-1}

EEG Signal

$H_0(z) \downarrow 2$

$H_1(z) \downarrow 2$ d_2

$H_1(z) \downarrow 2$ d_1
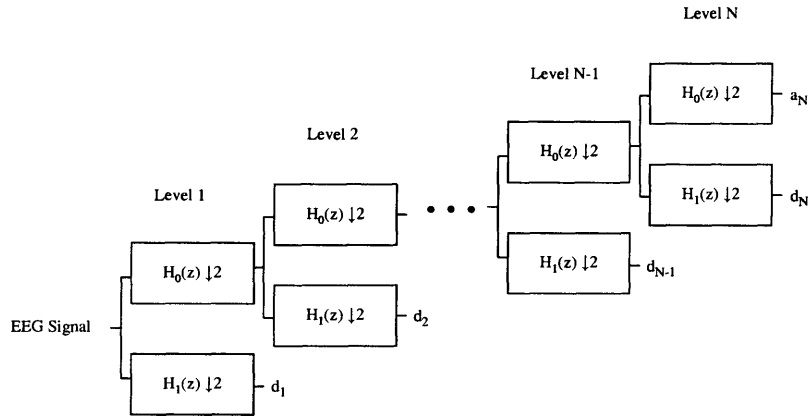
Figure 5-3: Multi-Level Wavelet Decomposition Filter Bank

of no clinical relevance. In particular, the subband signal $\{a_7\}$ captures slow baseline variations like those caused by sweating, while the subband signals $\{d_1\, d_2\, d_3\}$ capture high frequency artifacts similar to those resulting from muscular contractions.

To better appreciate the time-scales or frequencies captured within the subband signals $\{d_4\, d_5\, d_6\, d_7\}$, one can examine the overall *impulse* or *frequency response* of the cascade of filters between the input and each of the output subband signals. The frequency response illustrates the frequencies that will pass through the cascade of filters and appear in a given subband signal; while the the impulse response highlights the time-scale, or duration of activity to which the cascade of filters is most sensitive, and consequently appears in the output subband signal.

Figure 5-4 shows the overall impulse and frequency responses producing each of the subband signals $\{d_4\, d_5\, d_6\, d_7\}$. The impulse responses are progressively stretched for higher level subband signals so that activity of longer time-scales can be represented. This is equivalent to the observed decrease in center frequency and bandwidth of frequency responses associated with filter cascades producing higher level subband signals. Finally, the overall impulse responses are of interest because they simplify the computation of the subband signals from a real-time stream by collapsing each cascade of filters into a single filter that can be used with the overlap-add method of convolution [2].
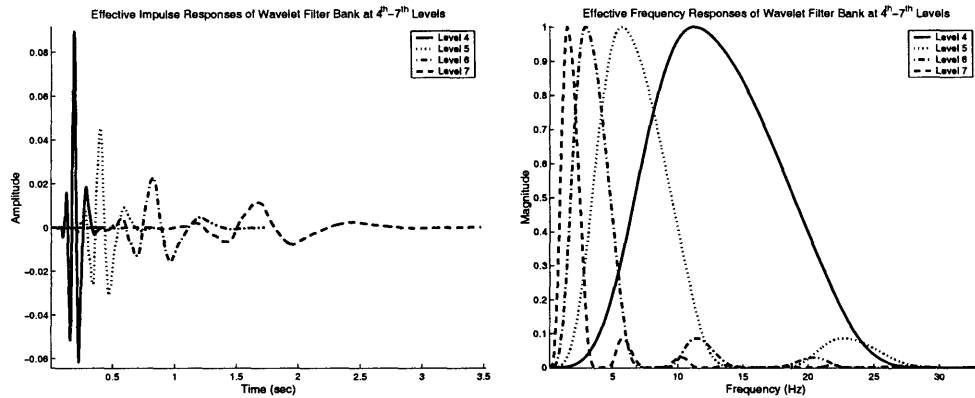
Figure 5-4: Effective Impulse and Frequency Responses of Wavelet Filter Bank

5.2.2 Feature Vector Construction

The subband signals $\{d_4 \, d_5 \, d_6 \, d_7\}$ should not be used directly as the entries of a feature vector since such an exact representation of the input EEG waveform would be too sensitive to both noise, and the slight variations in electrographic morphology commonly observed in the instances of a patient's seizures. Consequently it is useful to introduce four statistics that more generally summarize the information about waveform components within the four subband signals, and which can then be used as the entries of a four-dimensional feature vector X.

The statistics used by the detector correspond to the absolute, rather than normalized, log-energies in each of the subband signals $\{d_4 \, d_5 \, d_6 \, d_7\}$. These quantities are attractive since they are sensitive to the amplitude of waveform components within each subband signal, an important discriminating factor; and very effecient to compute. Moreover, the nonlinear log operator used in computing these quantities amplifies small differences separating feature vectors of the seizure and non-seizure classes. An explicit representation of the feature vector produced by the feature extraction stage is shown in equation 5.1.

66

$$X = \begin{bmatrix} \log(\sum_n |d_4(n)|) \\ \log(\sum_n |d_5(n)|) \\ \log(\sum_n |d_6(n)|) \\ \log(\sum_n |d_7(n)|) \end{bmatrix} \qquad (5.1)$$

In summary, the feature extraction stage begins with a wavelet decomposition of an EEG waveform to produce subband signals that capture components contributing to the waveform morphology at specific time-scales or frequencies. Next the energy in each of these subband signals is computed to form a statistic that summarizes their activity while still being robust to noise and commonplace variations in the electrographic morphology of a patient's seizure onset.

5.3 Classification

In the classification stage of the detection algorithm, feature vectors are assigned to either the seizure or non-seizure class by way of a classifier. The classifier must reliably make this binary assignment even though the feature vectors represent more than two classes of activity. Specifically, the non-seizure class represents normal as well as artifact-contaminated EEG observed in different states of consciousness; while the seizure class represents EEG activity observed during seizure onset. A probabilistic *Maximum-Likelihood* classifier (section 5.3.1) and a non-probabilistic *Support-Vector Machine* (section 5.3.2) classifier were considered for the task of determining the class membership of observed feature vectors under both the SIP and SDP architectures.

5.3.1 Maximum-Likelihood Classifiers

The maximum-likelihood classifier determines the class membership of a feature vector X by first computing the likelihood that the observation belongs to the seizure or non-seizure class, and then assigning the observation to the class with the greater likelihood [27]. This classification criterion can be modified so that the observation is assigned to the class with a likelihood exceeding that of the other class by a specific

factor as shown in equation 5.2. The conditional probability density $p(X|\text{seizure})$ is the likelihood that the observed feature vector X belongs to the seizure class, while the conditional probability density $p(X|\text{non-seizure})$ is the likelihood that it belongs to the non-seizure class.

$$\text{If } \frac{p(X|\text{seizure})}{p(X|\text{non-seizure})} \geq \gamma \text{ then } X \in \text{seizure} \tag{5.2}$$

The multi-dimensional likelihood functions $p(X|\text{seizure})$ and $p(X|\text{non-seizure})$ are a priori unknown, so their values for any observed feature vector X is estimated by the classifier using the associated class' training examples and the nonparametric method of product-kernel density estimation [29]. In essence, this density estimation technique equates the likelihood of a feature vector X to a sum of kernel functions $K(z)$ that are stretched and shifted according to the spatial distribution of training samples $\tilde{X}$ as shown in equation 5.3, and graphically illustrated for the one-dimensional case in Figure 5-5. The figure shows instances of a gaussian kernel centered over samples drawn from a one-dimensional random variable with unknown distribution, as well as the resulting bimodal density estimate that results from summing over the kernels. The bimodal density estimate explains well the clustering of the samples. The advantage of a nonparametric density estimate is that it makes no assumptions about the form of the likelihood functions in terms of the number or volume of modes, instead it extracts them from the training samples. Nevertheless, a nonparametric density estimate can be computationally taxing when a large number of training examples are used to generate the estimate.

$$p(X) = \frac{1}{n * h_1 * \cdots * h_d} \sum_{i=1}^{n} \prod_{j=1}^{d} K\left(\frac{X_j - \tilde{X}_{ij}}{h_j}\right) \quad K(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \tag{5.3}$$

In the SIP architecture a value for the threshold γ is automatically chosen by each classifier to limit its individual probability of false-positive classification to a specified tolerance level α. Specifically, each classifier searches for a γ that satisfies equation 5.4 using nonparametric estimates of the likelihood functions.
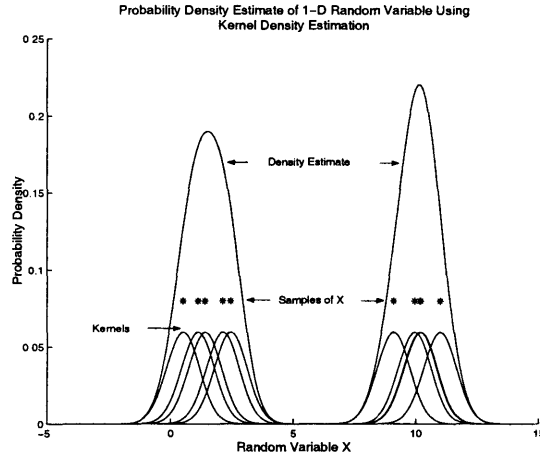
Figure 5-5: Probability Density Estimation using Kernels

$$Z = \left\{ X \mid \frac{p(X|\text{seizure})}{p(X|\text{non-seizure})} \geq \gamma \right\} \quad P_{FP} = \int_Z p(X|\text{non-seizure}) \, dX \leq \alpha \quad (5.4)$$

Equation 5.4 states that any value of γ defines a decision region Z where the classifier will assign all observed feature vectors X to the seizure class; the decision region Z may be one region or the union of several disjoint regions. Furthermore, the probability of a false-positive classification given a value γ is given by an integral over the region Z of the likelihood of X belonging to the non-seizure class. The value of γ must be chosen by the classifier so that this integral results in a probability of false-positive classification that is less than α. Once the appropriate γ is determined by each classifier, their individual probabilities of true-positive classification is given by equation 5.5. These probabilities are used in section 5.4 for the purpose of spatially localizing a patient's seizure onset.

$$P_{TP} = \int_Z p(X|\text{seizure}) \, dX \quad (5.5)$$

In the SDP architecture the high dimensional feature vectors prohibit the approximation of the integrals in equation 5.4. Consequently the value of value of γ is

not set according to a specified tolerance on false-positive classification, instead it is determined empirically and fixed across patients as explained in Chapter 6.

Maximum-Likelihood Classification Example

In this section the decision region computed by a maximum-likelihood classifier using a sample training set is visualized in a two-dimensional space. The two-dimensional feature vectors X' within this space are synthesized by projecting the four-dimensional feature vectors X used by the SIP architecture onto the directions of greatest variance ϕ_1 and ϕ_2 computed using principle components analysis [26].

$$X' = \begin{bmatrix} X'_1 \\ X'_2 \end{bmatrix} = \begin{bmatrix} \phi_1 \cdot X \\ \phi_2 \cdot X \end{bmatrix} \tag{5.6}$$

The patient-specific training feature vectors used by the maximum-likelihood classifier to determine a decision region are illustrated in Figure 5-6. These feature vectors were computed by passing seizure and non-seizure epochs from one derivation through the feature extraction stage, and then transforming the resulting four-dimensional feature vectors X into lower-dimensional feature vectors X'. Note the greater number of non-seizure training examples; this is typical of any training set since there is always more non-seizure EEG to sample from a patient than seizure EEG.
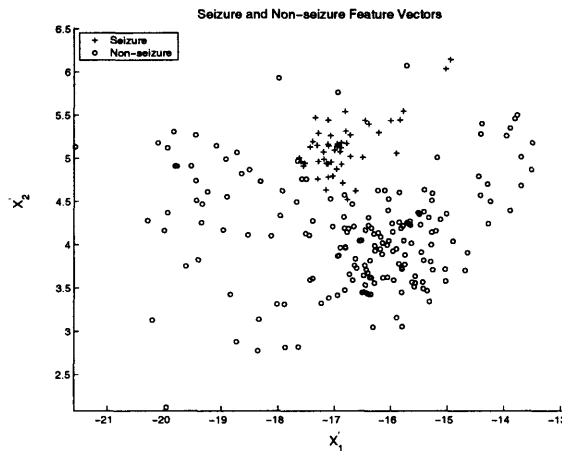
Figure 5-6: Training Seizure and Non-Seizure Feature Vectors

The first step in determining the decision region Z involves using the training feature vectors and kernel density estimation to construct estimates of the seizure and non-seizure likelihoods as shown in the left panel of Figure 5-7. These estimates are then used in equation 5.4 to compute the decision region Z that limits the probability of a false-positive classification to a maximum value of α; the region Z is outlined in the right panel of Figure 5-7. Increasing the value of α will result in a decision region with a greater radius, and consequently the correct classification of more seizure examples at the expense of the incorrect classification more non-seizure examples.
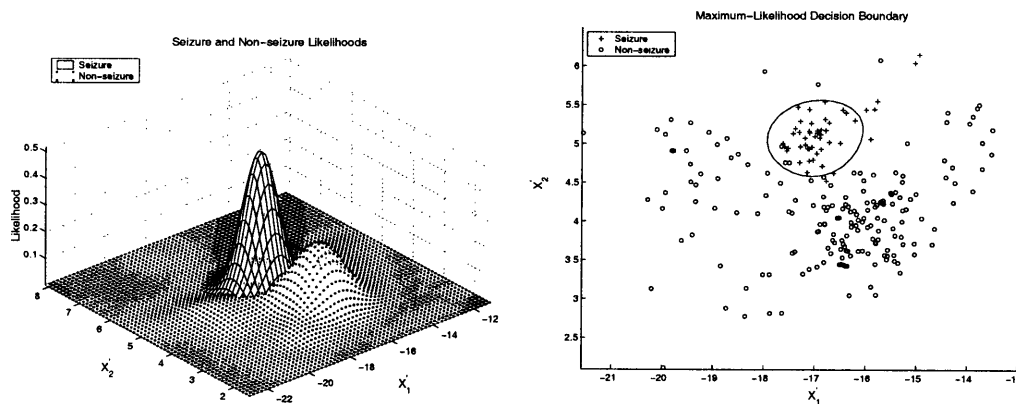
Figure 5-7: Likelihood Estimates (Left) and Decision Region (Right)

5.3.2 Support-Vector Machines

A support vector machine [23] determines the class membership of a feature vector X based on which side of an optimal hyperplane the observation lies. In the case of linearly separable classes this optimal hyperplane is the one that is maximally distant from *support-vectors*. These are the training examples from both classes corresponding to boundary cases, and consequently the ones carrying all relevant information about the classification problem. If the classes are not linearly separable the optimal hyperplane can be determined in a higher-dimensional feature space where they are linearly separable; this translates to computing a nonlinear decision boundary in the original space.

71

A *kernel* is a function that allows support-vector machines to define the optimal hyperplane in a kernel-specific, higher-dimensional space without the explicit construction of high-dimensional feature vectors. In the seizure detection algorithm, the Radial-Basis Kernel expressed in equation 5.7 was chosen since determination of the optimal hyperplane in its associated high-dimensional feature space yields nonlinear decision boundaries that may be discontinuous when necessary. This means that the decision region of a Radial-Basis Kernel need not be one region, instead it can be the union of several disjoint regions.

$$\text{Radial-Basis Kernel:} \quad K(X_i, X_j) = exp\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) \quad \sigma = (0 \, \infty) \qquad (5.7)$$

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 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 encourages the determination of a decision boundary that correctly classifies those examples; the penalties can be specified independently for each class through the two entries of a vector parameter C[2]. Extreme choices for both of these variables increases the risk of overfitting; that is creating a classifier that correctly identifies a high percentage of the training set, but performs 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

Finally in the SIP architecture, as described in section 5.4, the probabilities of true and false-positive classification of each classifier are used to localize a patient's seizure onset. In the case of support vector machines, these probabilities are approx-

---

[2]The absolute value of the penalties, as opposed to their ratio, is important. This means that the choice C=[20 10] is not equivalent to C=[40 20].

imated using the equations 5.8. The quantity $N_{correct\ seizure}$ is the number of correctly classified seizure examples in a training set that includes a total number of seizure examples $N_{total\ seizure}$; similarly $N_{incorrect\ normal}$ is the number of incorrectly classified normal examples in a training set that includes a total number of non-seizure examples $N_{total\ normal}$. The equations state that the probability of correctly classifying a newly observed feature vector as a seizure is approximated by the percentage of seizure training examples that are correctly classified, and the probability of incorrectly classifying the observed feature vector as seizure is approximated by the percentage of non-seizure training examples incorrectly classified.

$$P_{TP} \approx \frac{N_{correct\ seizure}}{N_{total\ seizure}} \qquad P_{FP} \approx \frac{N_{incorrect\ normal}}{N_{total\ normal}} \tag{5.8}$$

Support Vector Machine Classification Example

In this section the decision region computed by a support vector machine classifier using a sample training set is visualized in a two-dimensional space. As with the previous classification example, we synthesize the two-dimensional feature vectors X' within this space by projecting the four-dimensional feature vectors X used by the SIP architecture onto the directions of greatest variance ϕ_1 and ϕ_2 computed using principle components analysis [26].

$$X' = \begin{bmatrix} X'_1 \\ X'_2 \end{bmatrix} = \begin{bmatrix} \phi_1 \cdot X \\ \phi_2 \cdot X \end{bmatrix} \tag{5.9}$$

The patient-specific training feature vectors used by the support vector machine to determine a decision region are illustrated in Figure 5-8; and are equivalent to those used in the classification example of the maximum-likelihood classifier. The feature vectors were computed by passing seizure and non-seizure epochs from one derivation through the feature extraction stage, and then transforming the resulting four-dimensional feature vectors X into lower-dimensional feature vectors X'.

73
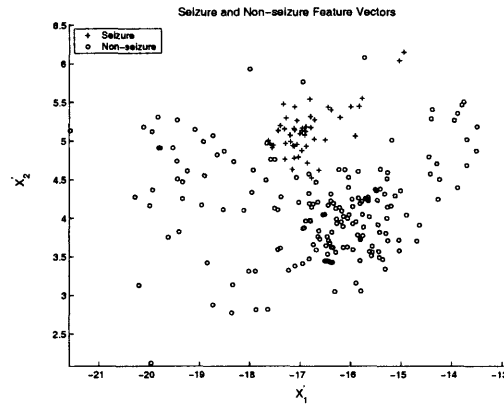
Figure 5-8: Training Seizure and Non-Seizure Feature Vectors

The support-vector machine classifier uses the training feature vectors to compute the coefficients parameterizing the optimal hyperplane in either the original or kernel-induced feature space. Computing the hyperplane in the original feature space leads to the linear decision boundary shown in the left panel of Figure 5-9, while computing the hyperplane in the feature space induced by a radial basis kernel with parameter $\sigma = 1$ is shown in the right panel of the figure. The nonlinear decision boundary computed by the support vector machine is very different from that determined by the maximum-likelihood classifier, which is not unexpected given the vastly different theoretical foundation of each classification scheme.
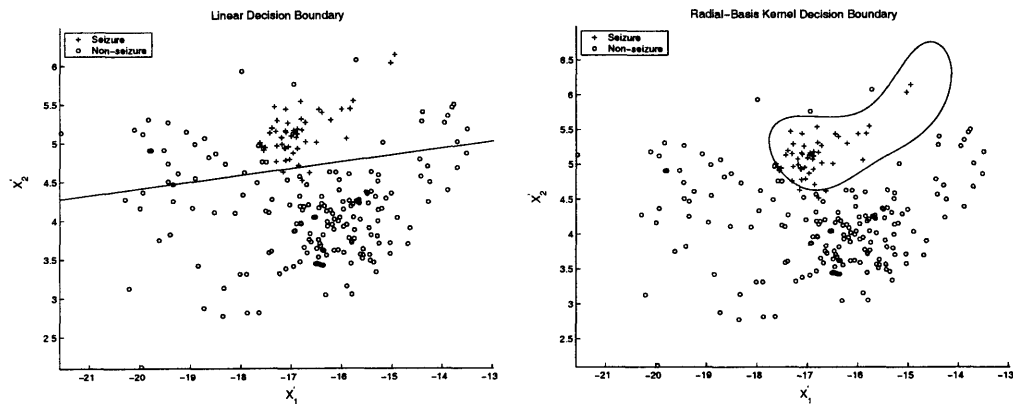
Figure 5-9: Support-Vector Machine Linear and Non-linear Decision Boundaries

5.4 Spatial and Temporal Constraints

In the SIP architecture the assigned class memberships of the twenty-one feature vectors are examined in the context of temporal and patient-specific localization constraints in order to make a final decision regarding seizure onset. Specifically, a detector with the SIP architecture declares seizure onset only after K derivations are assigned to the seizure class for a duration of T seconds. The K derivations must all belong to one of the groups illustrated in Figure 5-10; which one depends on the nature of each patient's seizures and is determined automatically by the detector. The groups in Figure 5-10 were chosen because they provide coverage of possible centers of focal seizure activity; moreover, in the case of generalized seizures any one of these groups can be used for the purpose of detection since all derivations will be active at the seizure's onset.
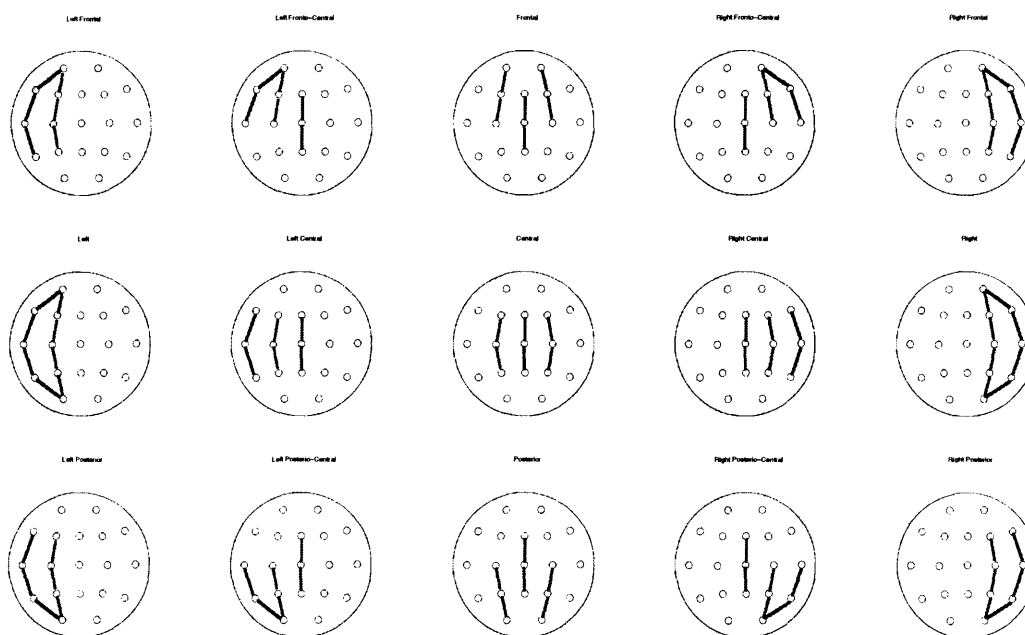
Figure 5-10: Groups of Contiguous EEG Derivations

For a given patient the detector chooses the group exhibiting the highest discriminability between seizure and non-seizure activity on its constituent derivations. This is accomplished by first assigning each derivation a weight based on the ability of its classifier to differentiate between seizure and non-seizure activity, and then selecting the group with the maximal total weight. The weight a_i assigned to derivation i is computed using its classifier's probability of true and false-positive classification as expressed in equation 5.10, while the optimal group G_j is the one with the greatest total weight w_j shown in equation 5.11. This selection procedure is similar in spirit to that used by an electroencephalographer, who identifies derivations active at seizure onset by noting how visually distinct their waveforms are during seizure and non-seizure periods.

$$a_i = P_{TP,i} - P_{FP,i} \quad i = 1, \ldots, 21 \tag{5.10}$$

$$w_j = \sum_{k \in G_j} a_k \quad j = 1, \ldots, 15 \tag{5.11}$$

5.5 Training

During training the detector's classifiers use a diverse set of examples from the seizure and non-seizure classes to determine decision boundaries. The training examples are patient-specific, non-overlapping sets S_i $i = 1, \ldots, 21$ each containing two-second epochs of labelled activity from a single EEG derivation. The epochs that correspond to seizure activity are labelled as examples of the seizure class, while those corresponding to both normal and artifact-contaminated activity from different states of consciousness are labelled as examples of the non-seizure class.

The training procedure begins by converting the labelled sets S_i into a collection of feature vectors $\{X\}$ by passing their epochs through the feature extraction stage. The feature vectors are used by the classifiers for the purpose of estimating quantities necessary for defining a decision boundary. In the case of maximum-likelihood classifiers, these quantities correspond to the conditional densities $p_i(X|\text{seizure})$ and

76

$p_i(X|\text{non-seizure})$; while for support-vector machines the quantities are the coefficients of the hyperplane in the kernel-induced feature space.

5.6 Case Studies

5.6.1 Case 1

This case study explores in detail the detector's training process as well as consequences of our method for automatically determining spatial localization constraints.

Consider detecting the electrographic onset of the seizure illustrated in Figure 5-11 using a detector with the SIP architecture. This seizure's onset is characterized by a paroxysmal, 10 Hz burst of sharp and monomorphic waves localizing to the central derivations $\{F_Z - C_Z \ C_Z - P_Z\}$, the right fronto-central derivations $\{FP_2 - F_4 \ F_4 - C_4\}$, and the right frontal derivations $\{FP_2 - F_8 \ F_8 - T_8 \ T_8 - P_8\}$. With the exception of $\{FP_1 - F_7 \ FP_1 - F_3\}$, the derivations on the left side of the head, which are odd-numbered, show no appreciable change in behavior after the onset. All of this implies that the seizure originates from a region towards the front and right-side of the head.

The first step in the detection process is to train the detector not only on 2-4 previous occurrences of seizure onsets similar to that illustrated in Figure 5-11, but also on the non-seizure EEG separating these occurrences. Figure 5-12 shows one of the training seizures presented to the detector; the training seizure is very similar to the one we hope to detect except for less prominent activity on the frontal derivations $\{FP_1 - F_7 \ FP_1 - F_3 \ FP_2 - F4 \ FP_2 - F_8\}$. This difference illustrates the variability between the instances of a seizure, and explains why the detector requires more than one training seizure in order to discover the derivations that are consistently active following the electrographic onset. The training seizure is not used as it is shown in the figure, instead it is segmented into two-second epochs that are grouped into the training sets $S_i \ i = 1, \ldots, 21$ according to their source derivation.

Onset of Test Seizure
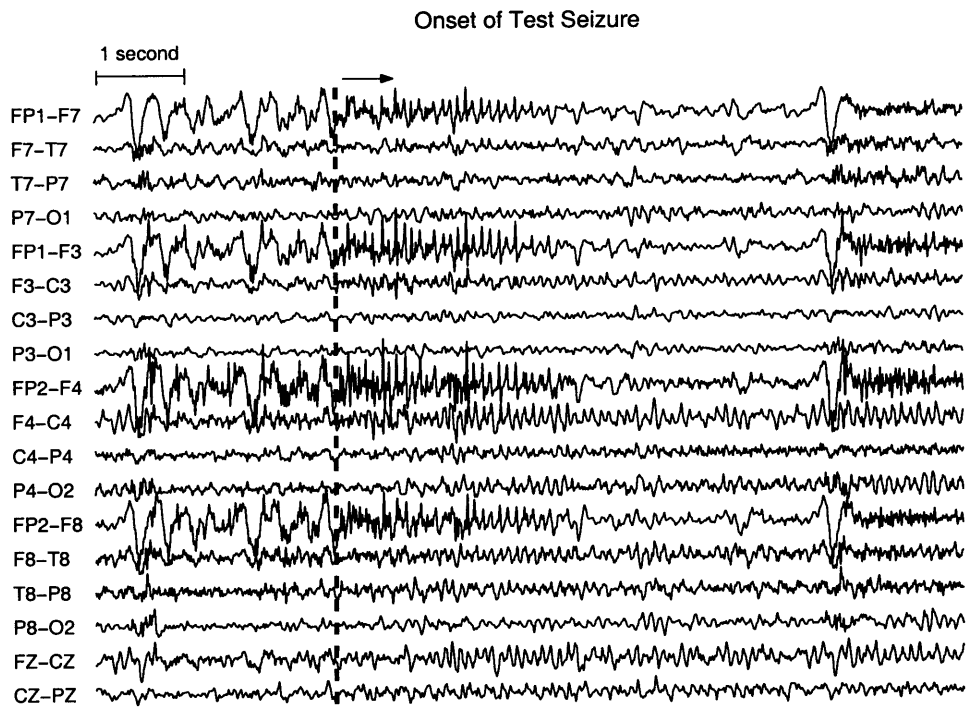
Figure 5-11: Case 1: Electrographic Onset of Test Seizure Following Dotted Line
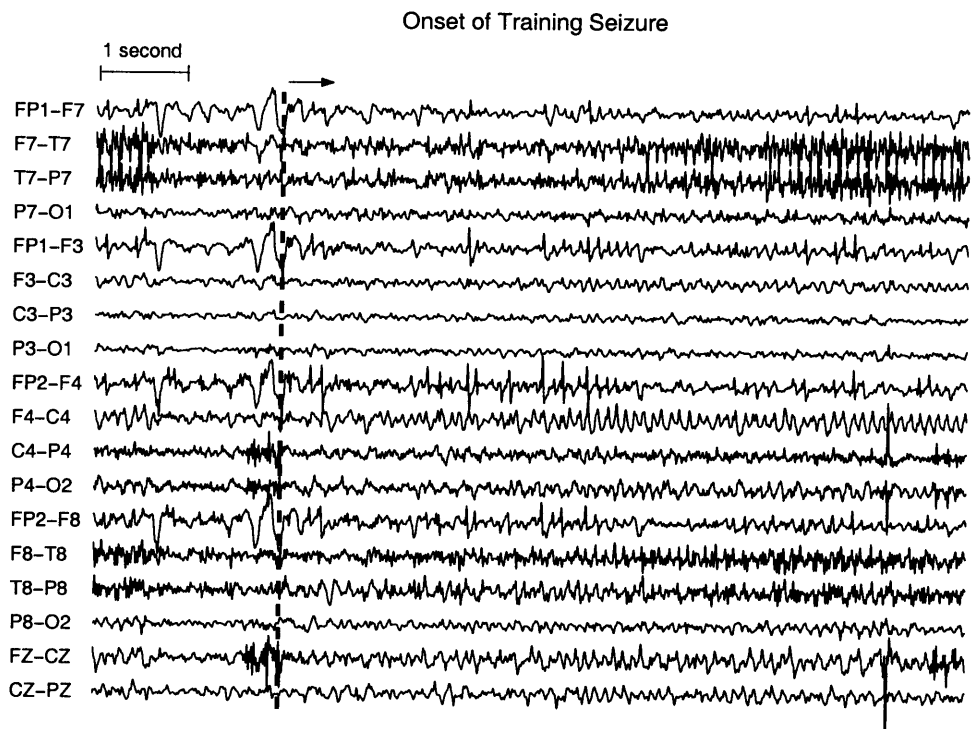
Figure 5-12: Case 1: Electrographic Onset of Training Seizure Following Dotted Line

The non-seizure EEG included as part of the detector's training consists of the baseline EEG; rhythms from different states of consciousness such as the normal alpha rhythm; physiological artifacts like those caused by eye flutter or chewing; and nonphysiological artifacts like those introduced by movement of EEG electrodes. Since nonphysiological artifacts are not necessarily limited to the derivations on which they are observed, they are artificially introduced into the training set S_i of each classifier. In all other cases, the training sets S_i only contain epochs of EEG from a single derivation. Figure 5-13 shows various non-seizure EEG training epochs.
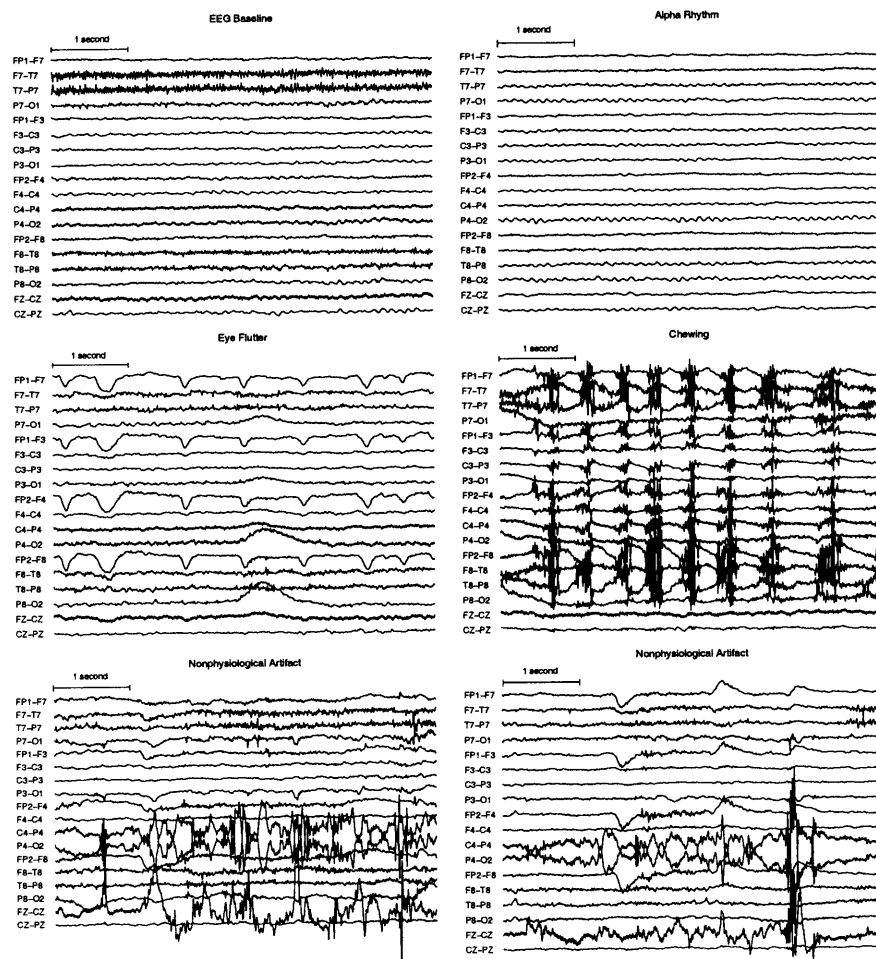
Figure 5-13: Case 1: Non-seizure Training EEG

After the epochs within the training sets S_i are converted to sets of feature vectors, the detector determines the decision boundary associated with each classifier. For instance, the maximum-likelihood and support vector machine decision boundaries for the derivation $\{F_4 - C_4\}$ are shown in Figures 5-7 and 5-9. The detector uses the decision boundaries to compute the probabilities of true and false-positive classifications $P_{TP,i}$ and $P_{FP,i}$ for the purpose of localizing the seizure's onset to one of the groups in Figure 5-10. In this example, the detector selected the right-central derivations shown in Figure 5-14. All the selected derivations exhibit a change in their waveforms following seizure onset with the possible exception of $\{C_4 - P_4\}$; this result illustrates that a consequence of selecting derivations as a group is the possible inclusion of irrelevant derivations, and also explains why the detector performs poorly when declaration of a seizure event is conditioned on observing seizure activity on K=6 rather than K < 6 derivations. Note that specifying a minimum number of derivations for declaring a seizure event is not required by the SDP architecture since spatial localization constraints are encapsulated within its feature vectors, rather than explicitly imposed as in the SIP architecture.

Right Central
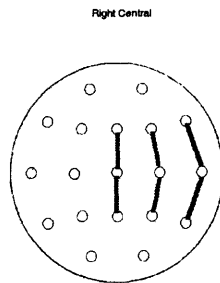
Figure 5-14: Case 1: Selected Group of Derivations

When the trained detector was used to detect the test seizure using K=4 derivations and T=6 seconds, a seizure event was declared seven seconds following the electrographic onset as shown in Figure 5-15. The derivations responsible for triggering the detection included $\{F_4 - C_4 \quad F_8 - T_8 \quad T_8 - P_8 \quad F_Z - C_Z \quad C_Z - P_Z\}$. On the other hand, the abnormal activity on the frontal derivations $\{FP_1 - F_3 \quad FP_1 - F_7 \quad FP_2 -$

81

F_4 $FP_2 - F_8$} was not used for the purpose of detection because these derivations are not members of the selected group. Even if the frontal derivations were members of the selected group they would not have triggered a detection since their seizure activity does not persist for the required T=6 seconds.
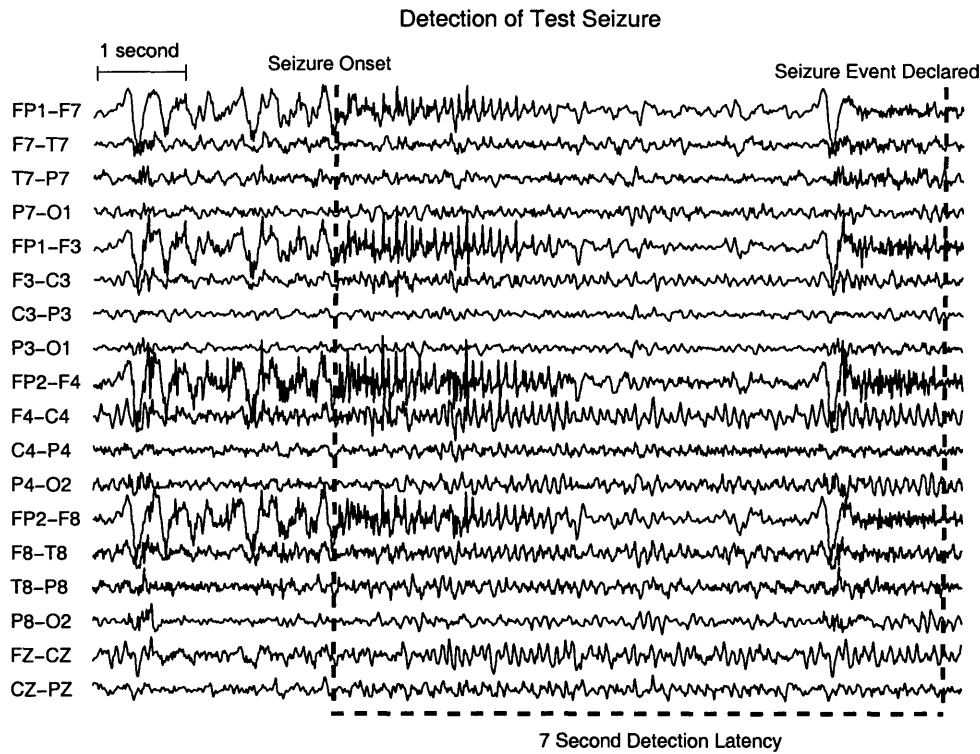
Figure 5-15: Case 1: Detection of Seizure Onset

5.6.2 Case 2

This case study highlights the importance of both localization and morphology to seizure detection, and the possibility of sharing certain types of non-seizure activity across the training sets of patients.

Consider detecting the electrographic onset of the seizure illustrated in Figure 5-16 again using a detector with SIP architecture. This seizure's onset is characterized by a paroxysmal 2 Hz burst of monomorphic waves localizing to the central derivations

82

$\{F_Z - C_Z \ \ C_Z - P_Z\}$, and all derivations on the right-side of the head $\{FP_2 - F_4 \ \ F_4 -$ $C_4 \ \ C_4 - P_4 \ \ P_4 - O_2 \ \ FP_2 - F_8 \ \ F_8 - T_8 \ \ T_8 - P_8 \ \ P_8 - O_2\}$. The baseline EEG can be observed on derivations from the left-side of the head, which are odd-numbered, since they exhibit no change after the onset. This electrographic evidence indicates that the seizure originates from the right-side of the head.
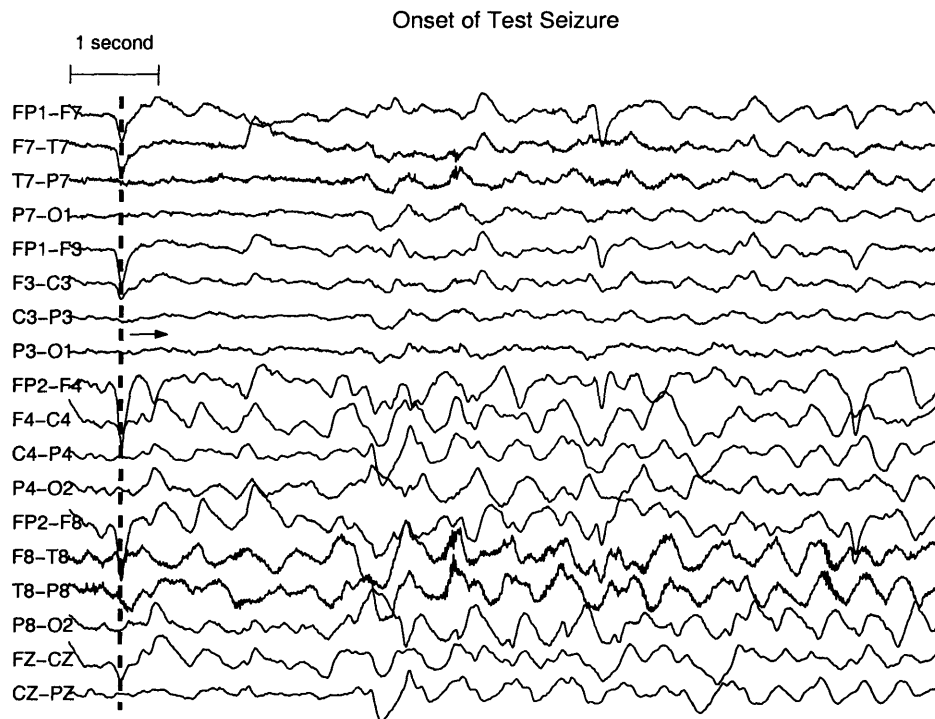
Figure 5-16: Case 2: Electrographic Onset of Test Seizure Following Dotted Line

To detect the seizure in 5-16 the detector must be trained on previous instances of the seizure as well as on the non-seizure EEG separating these instances as was done in Case 1. It is interesting to note that the baseline EEG included as part of the non-seizure training must be specific to the case; in contrast, physiological and nonphysiological artifacts as well as hallmark activity from different states of consciousness can be shared across cases within similar age groups. This is supported by the fact that an electroencephalographer can identify these activities solely based on morphology, localization, and reactivity; reference to the baseline EEG associ-

ated with the case is not necessary. In contrast, an electroencephalographer cannot be certain whether an epoch of activity includes seizure onset without reference to the baseline EEG, which argues for the necessity of baseline and seizure EEG to be case-specific [5]. Practically this means that a diverse library of case-independent physiological and nonphysiological activity can be compiled and saved, and then used to supplement the baseline and seizure EEG that are specific to the case under consideration. This is an approach we adopted while testing our algorithm as described in Chapter 6. Figure 5-17 shows one of the training seizures presented to the detector.
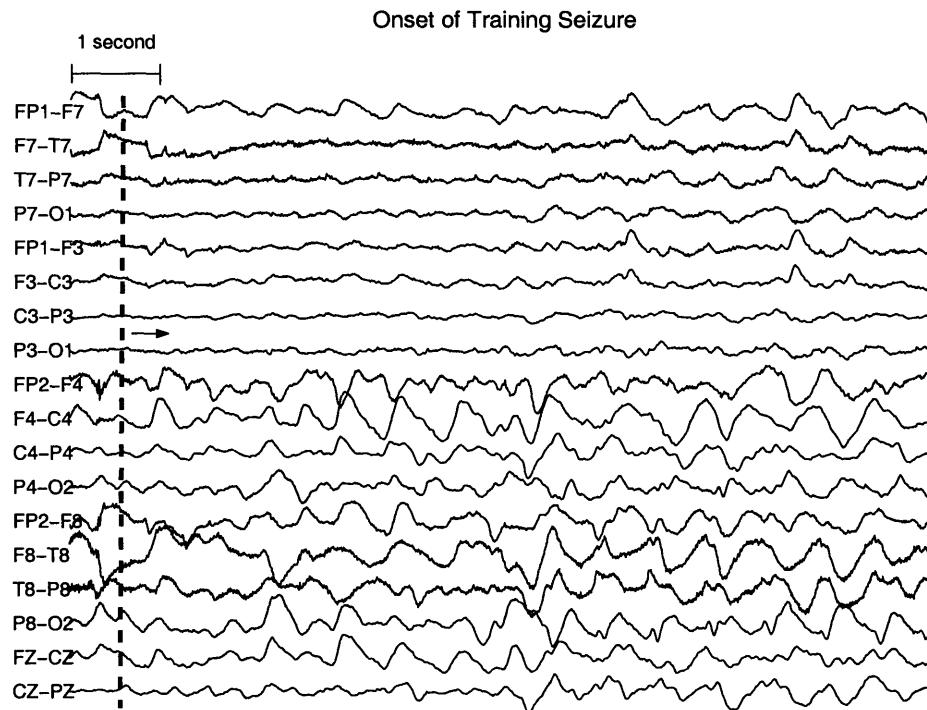
Figure 5-17: Case 2: Electrographic Onset of Training Seizure Following Dotted Line

Following training and completion of the localization procedure of section 5.4, the detector selected the right-central derivations shown in Figure 5-14. While the selected group of derivations matches that of Case 1, the detector from Case 2 fails to detect the test and training seizures from Case 1 because of the very different waveform morphologies. This demonstrates the importance of both morphology and

localization to seizure onset detection.

When the trained detector of this case was used to detect the test seizure in 5-16 using K=4 derivations and T=6 seconds, a seizure event was declared seven seconds following the electrographic onset as shown in Figure 5-18. The six derivations responsible for the detection included $\{F_4-C_4 \quad C_4-P_4 \quad F_8-T_8 \quad T_8-P_8 \quad F_Z-C_Z \quad C_Z-P_Z\}$.
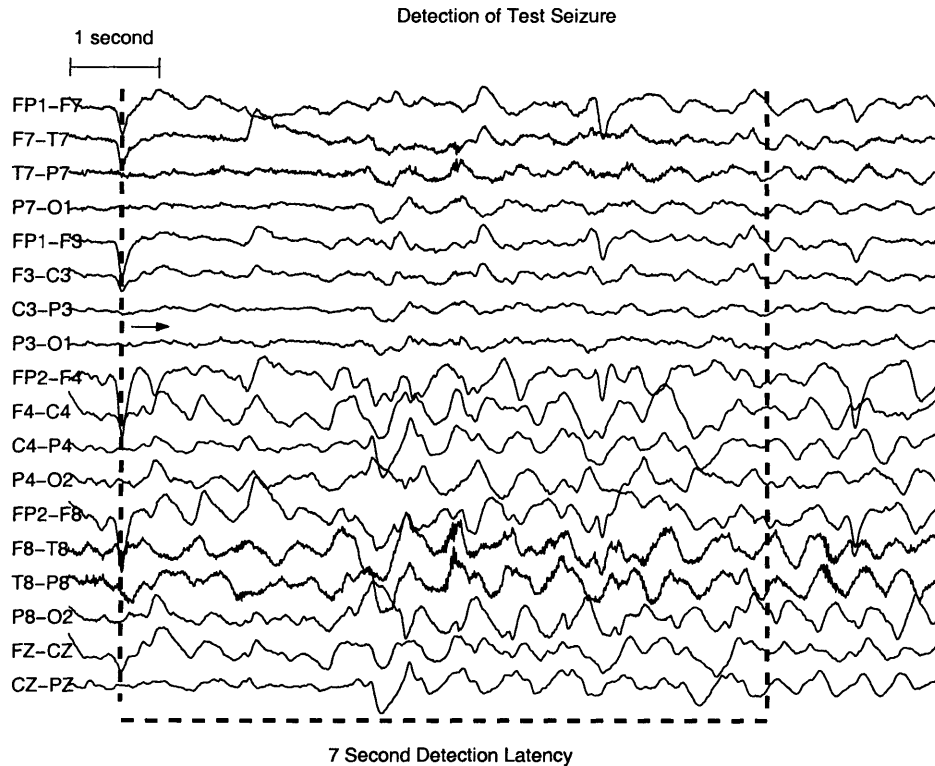
Figure 5-18: Case 2: Detection of Seizure Onset

5.6.3 Case 3

This case illustrates a type of patient-specific, non-seizure activity that often leads to the false declaration of seizure events by the detector. The activity corresponds to the the abnormal discharges discussed in section 3.5, these are observed between seizure events and may have similar morphology and localization to actual seizures.
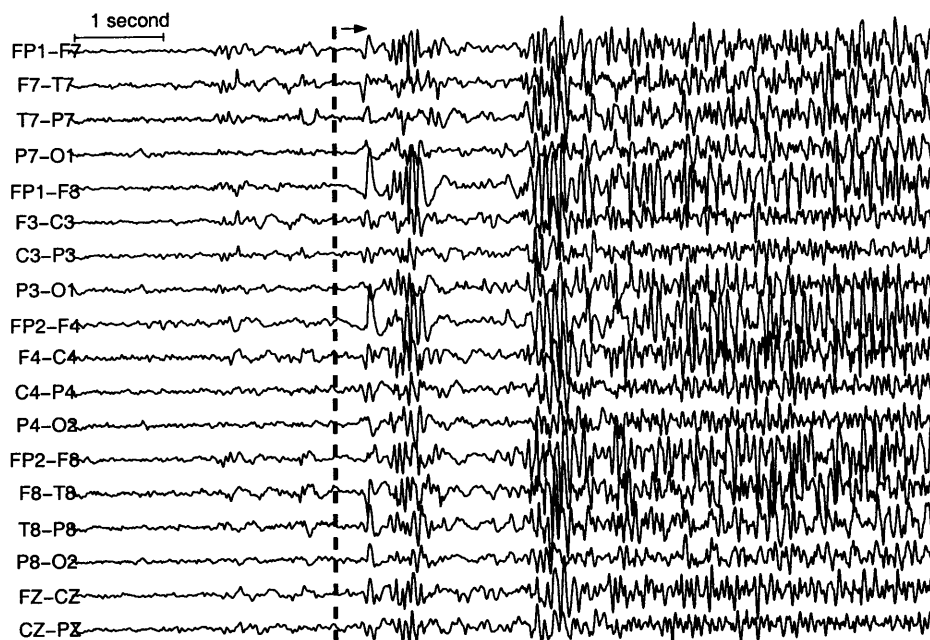
Figure 5-19: Case 3: Electrographic Onset of Training Seizure Following Dotted Line

Consider a detector with the SIP architecture that is trained on several electrographic seizure onsets similar to that shown in Figure 5-19. Since the onsets are generalized, the detector can select any of the group of derivations illustrated in Figure 5-10 for the purpose of subsequent detections.

When the trained detector was presented with non-seizure EEG between seizure occurrences, a false seizure event was declared upon analyzing the generalized, periodic discharge of sharp-wave groups boxed in Figure 5-20 following the dotted line. Visually one can distinguish the sharp wave groups in Figure 5-20 from those in Figure 5-19 by their temporal spacing. To the detector both activities are similar since the spacing between any two groups of sharp waves does not exceed two-seconds, the duration with which EEG is analyzed.

86
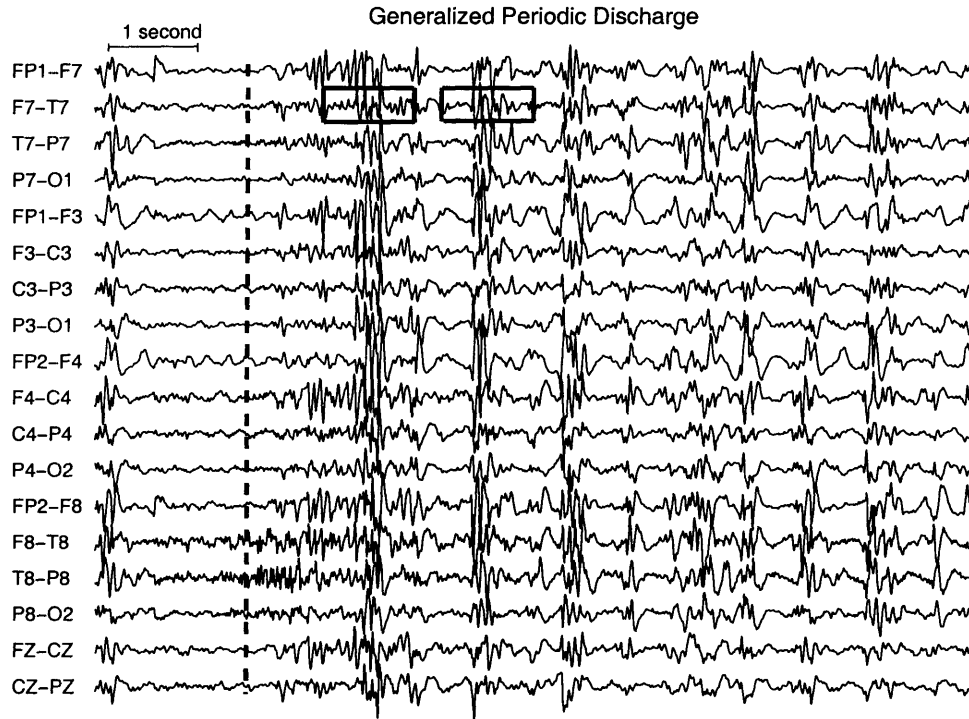
Figure 5-20: Case 3: Periodic Discharge Leading to False Seizure Event

5.7 Summary

In this chapter we described two possible processing architectures for the patient-specific seizure onset detector. In the Spatially Independent Processing (SIP) architecture, feature vectors extracted from each derivation are assigned to the seizure or non-seizure class independently by way of classifiers trained on the EEG of an associated derivation. A final decision regarding seizure onset is declared after all classifications are examined in the context of automatically extracted spatial and temporal constraints. In the Spatially Dependant Processing (SDP) architecture, the features extracted from all derivations are grouped into a large feature vector that captures whatever interdependencies exist between them. This feature vector is then assigned to the seizure or non-seizure class by way of a classifier trained on EEG from all the derivations.

We described how a multi-level wavelet decomposition is used to extract feature vectors that capture the amplitude, fundamental frequency, and morphology of EEG waveforms. Then we discussed how a maximum-likelihood or support vector machine classifier can be used to determine the class membership of these vectors based on training examples of seizure and non-seizure EEG. As mentioned in the second case study, the training seizure and baseline EEG must be specific to a patient, but physiological and nonphysiological artifacts as well as hallmark activity from different states of consciousness may be shared across them.

Finally, we concluded the chapter with case studies highlighting properties of the detector. The first case study explored in detail the detector's training process as well as consequences of our method for automatically determining spatial localization constraints. The second case study highlighted the importance of both localization and morphology to seizure detection, and the possibility of sharing certain types of non-seizure activity across the training sets of patients. The final case study illustrated a type of patient-specific, non-seizure activity that often leads to the false declaration of seizure events by the detector.

Chapter 6

Performance

In this chapter the performance of our seizure onset detector on thirty-six, de-identified test subjects is presented. Performance is gauged in terms of the following three metrics computed for *each* subject:

- **Detection Latency**: The average time elapsed between the electrographic onset of a seizure and the declaration of a seizure event.

- **True-Detections**: The number of test seizures declared as seizure events.

- **False-Detections**: The number of false-positives declared during analysis of non-seizure EEG.

In general, a detector cannot be biased to improve performance in terms of all three metrics simultaneously; instead, improving performance as measured by one or two of these metrics implies forgoing performance as measured by the third. For example, decreasing the detection parameter T will result in shorter detection latencies; a possible increase in the number of true-detections; and an almost certain increase in the number of false-detections. The extra false-detections will result from short-duration, seizure-like discharges commonly observed in the EEG separating seizure events. The number of true-detections will increase or remain unchanged depending on whether or not the original value of T resulted in misses of very short-duration seizure events.

The testing methodology used to compute the performance metrics is described in section 6.1. This is followed by a summary and comparison of the detector's performance on thirty-six test subjects using the SIP and SDP architectures in section 6.2

.

6.1 Testing Methodology

For each test subject four or five bipolar EEG recordings sampled at 256 Hz, and each containing a seizure event with an onset labelled by an experienced electroencephalographer were available. The recordings lasted approximately 20 minutes for twenty-four subjects; 40 minutes for six subjects; 150 minutes for four subjects; and 12 hours for two other subjects.

For each subject a *leave-one-out* cross-validation testing scheme was followed: The detector was trained on a training set that includes the seizure and non-seizure epochs from all but one of the subject's recordings, and was then used on the excluded recording. This was repeated until each recording had been excluded once. We also added to the training set a library of epochs that included generic artifacts and hallmark activity from various states of consciousness; for example, sleep spindles from the second stage of sleep. This compensates for the possible under representation of any type activity in the training recordings; more practically, it implies that training records can be assembled quickly and without a great deal of concern over whether or not they are truly representative.

In short, a subject with recordings {A B C D} would require the following four testing trials

- **Trial 1** Train on {A B C EEG Library} and test on recording D

- **Trial 2** Train on {A B D EEG Library} and test on recording C

- **Trial 3** Train on {A C D EEG Library} and test on recording B

- **Trial 4** Train on {B C D EEG Library} and test on recording A

The performance metrics we would ideally measure to characterize the detector include the expected detection latency; the percentage of seizures likely to be detected; the hourly rate at which false-detections occur; and the ratio of true to false-detections in a given period of time. The final metric exposes the frequency with which true and false-detections occur in a routine clinical monitoring session. Unfortunately, we cannot accurately report the hourly rate of false-detections since we do not have long-term recordings for each patient. We also cannot report the ratio of true to false-detections since the number of seizure events occurring in a given period of time is lost through recordings with unknown temporal sequencing. This means we can neither determine if the seizure event in one recording occurred an hour or a day after the seizure event in another, nor whether other seizure events occurred in between.

The performance metrics we actually report are the average detection latency; the number of test seizures detected; and the total number, as opposed to the hourly rate, of false-detections. For a given subject the reported detection latency is the *average* of latencies measured in each testing trial, while the reported number of true and false-detections is the *sum* of seizures and false-positives declared in all the testing trials. The average detection latency corresponds closely to the desired "expected latency" metric. Also, once the number of test seizures detected is normalized by the total number of available test seizures, it will closely approximate the desired metric "percentage of seizures likely to be detected".

Reporting the total number of false-detections equally weighs false-detections declared in the short length recordings of one patient with those in the long length recordings of another. In other words, a false-detection caused by a movement artifact in a twenty-minute recording is not treated differently from the same false-detection in a thirty-minute recording. If we had decided to compute an hourly rate, we would have estimated a false-detection rate of 3/hour for one recording and a rate of 2/hour for the other recording; however, nothing about the recordings suggests that extending them to an hour would have resulted in one or two more false-detections.

6.2 Results

In this section the performance of the seizure onset detector under both the SIP and SDP architectures is reviewed and compared for thirty-six test subjects. Specifically, section 6.2.1 examines the performance and optimal parameter settings of the SIP architecture when either support vector machines or maximum-likelihood classifiers are used in the classification stage; the same discussion for the SDP architecture is in section 6.2.2. Finally, architectures are compared in section 6.2.3.

All detection architectures satisfy our performance requirements. In particular, the detector that combines the SDP architecture with the support-vector machine classifier exhibited an average detection latency of 8.0 ± 3.2 seconds while correctly declaring 131 of 139 seizure events; and declared only 11 false-detections during 49 hours of randomly selected non-seizure EEG.

6.2.1 Spatially Independent Processing Architecture

In the SIP architecture, the detector's performance is influenced by the choice of several parameters that directly control when seizure onset is declared. These parameters are: the required duration time T of an abnormality; the minimum number of derivations K exhibiting the abnormality; the allowable probability of false-positive classification α for maximum-likelihood classifiers; and the radial-basis kernel parameter σ and vector parameter C for support vector machines. The parameters α, σ, and C may be freely set for each classifier in the SIP architecture, but to reduce the detector's degrees of freedom one value for each parameter is used across all of them.

The change in performance of a detector that combines the SIP architecture with maximum-likelihood classifiers due to different choices of the parameters T, K, and α is illustrated in Figure 6-1. Not surprisingly, the figure shows that for a given choice of T and K increasing the probability of false-positives α results in a decrease in the average detection latency, and an increase in both the true and false-detections measured for twenty-eight of the thirty-six subjects.

Detection Latency, True-Detections, and False-Detections vs Parameters
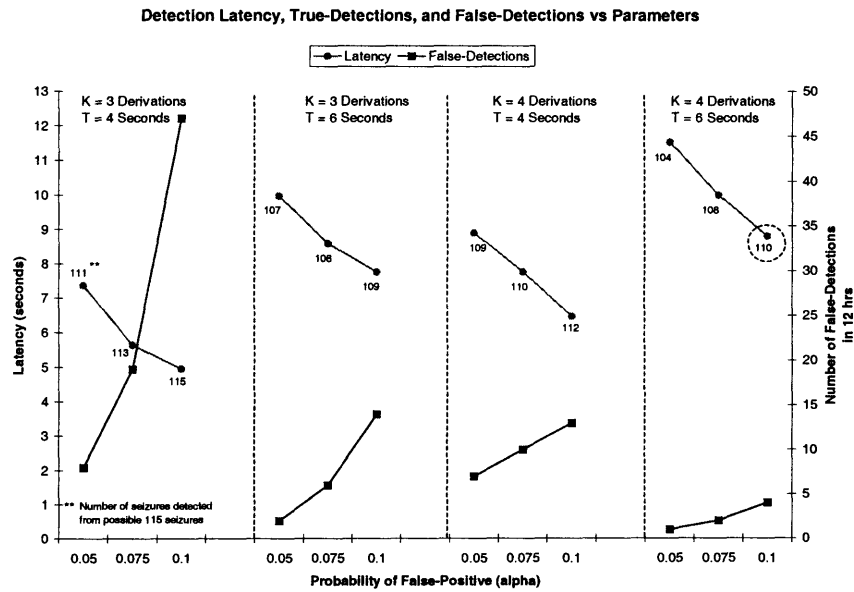
Figure 6-1: SIP Architecture Sensitivity with Maximum-Likelihood Classifier

The optimal choice of parameter settings depends entirely on the detector's application. For instance, if the detector is used to activate harmless stimulation of brain regions upon detecting a seizure, then false-detections are not costly but minimizing latency is crucial. In such a case, the parameter settings $T = 4$ seconds, $K = 3$ derivations, and $\alpha = 0.10$ may be appropriate. In our application minimizing both latency and false-detections is crucial, which is achieved by the parameter settings $T = 6$ seconds, $K = 4$ derivations, and $\alpha = 0.10$ as shown by the circled data point in Figure 6-1. A high true-detection rate is also desirable; however, since the cost of a miss in the ictal SPECT application is conducting the procedure using existing hospital protocols, we decided to choose parameters that primarily maximize performance in terms of latency and false-detections.

The sensitivity of a detector that combines the SIP architecture with support vector machines to changes in T, K, σ, and C is illustrated in Figures 6-2 through 6-4. For a given choice of the vector C, whose first and second entries corresponds to the cost of misclassifying seizure and non-seizure training examples respectively,

93

Detection Latency, True-Detections, and False-Detections vs Parameters for C = [10 10]
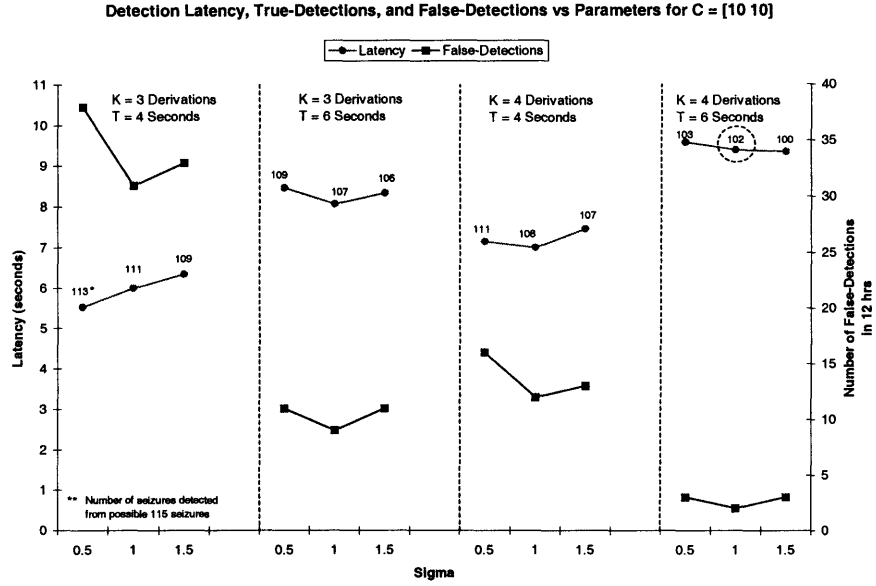
Figure 6-2: SIP Architecture Sensitivity with Support Vector Machine C=[10 10]

Detection Latency, True-Detections, and False-Detections vs Parameters for C = [30 10]
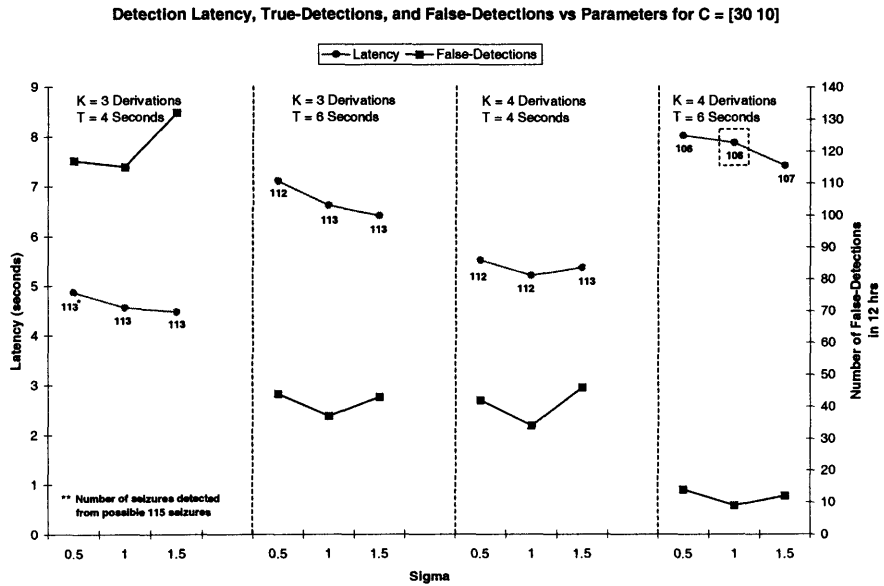
Figure 6-3: SIP Architecture Sensitivity with Support Vector Machine C=[30 10]

94
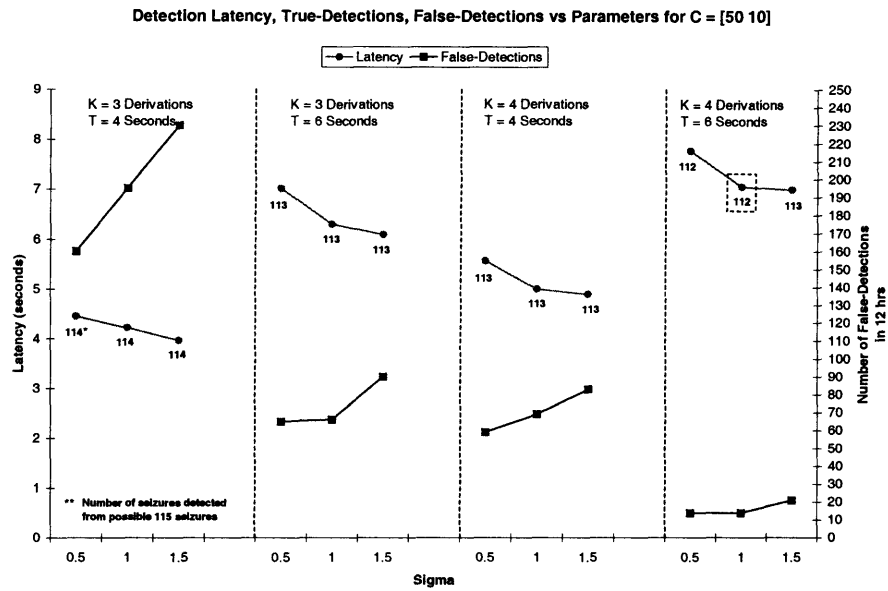
Figure 6-4: SIP Architecture Sensitivity with Support Vector Machine C=[50 10]

the values of T and K are responsible for changes in the average detection latency and total number of true and false-detections. In contrast, the performance metrics remain almost constant for changes in σ. The values of σ were chosen so that decision boundaries required between 10%-40% of the training data to be support vectors, a percentage that limits the prospect of overfitting. The parameter settings $C = [10\ 10]$, $T = 6$ seconds, $K = 4$ derivations, and $\sigma = 1$ minimize both latency and false-detections as measured for twenty-eight of the thirty-six subjects; this data point is circled in Figure 6-2.

Although the detector can exhibit a lower detection latency and a higher true-detection rate with $C = [30\ 10]$ and $C = [50\ 10]$ as shown by the boxes in Figures 6-3 and 6-4, the circled parameter settings that include $C = [10\ 10]$ exhibit a lower number of false-detections.

For the parameter settings $T = 6$ seconds, $K = 4$ derivations, $\alpha = 0.10$, $\sigma = 1$, and $C = [10\ 10]$, Figure 6-5 compares the average detection latency of a detector that combines the SIP architecture with maximum-likelihood classifiers and support vector

95

SIP Architecture Detection Latency with
Maximum-Likelihood and Support Vector Machine Classifiers
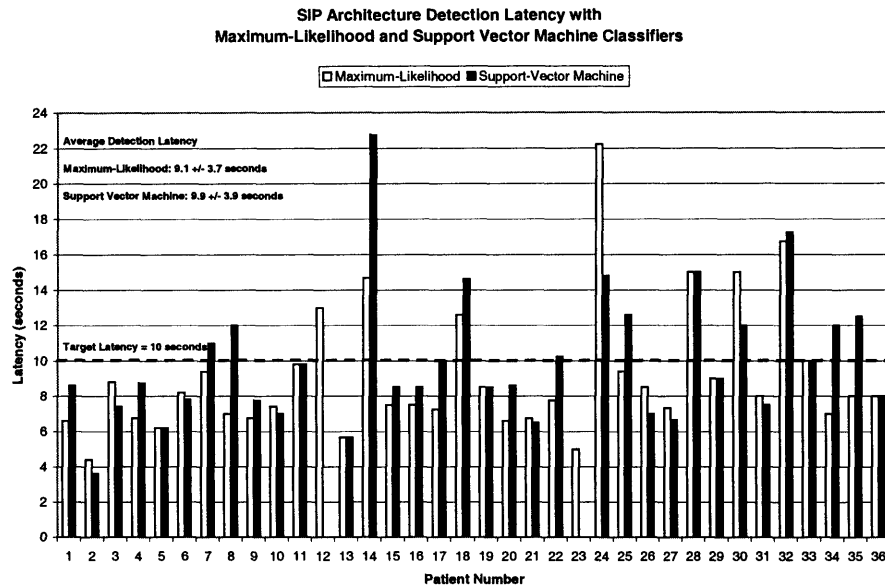
Figure 6-5: SIP Architecture Detection Latency

machine classifiers. The detection latencies for both configurations are similar, which argues that the detector is not grossly sensitive to the classifier type. Furthermore, the detection latencies of most subjects are less than the proposed target latency of ten seconds by more than one second. For subjects 12 and 23 a zero detection latency is shown since the support vector machine based detector failed to identify any seizure events. However, when the parameter C was changed from $C = [10\ 10]$ to $C = [30\ 10]$, the support vector machine tries harder to correctly classify seizure waveforms and does so with a latency matching that of the maximum-likelihood classifier, but at the expense of two extra false-detections on subject 12. The same change in C also reduces the latency of the support vector machine based detector on subject 14 to the level shown for the maximum-likelihood based detector. Finally, the large latencies shown for subjects 14 and 24 are a result of gradual seizure onsets localizing to a number of derivations less than the required detection minimum of K=4 before spreading to include a greater number of derivations.

Figure 6-6 shows the false-detections declared on each test subject for both detector configurations. With the exception of subject 30 whose false-detections were a

**SIP Architecture False-Detections with
Maximum-Likelihood and Support Vector Machine Classifiers**

□ Maximum-Likelihood ■ Support Vector Machine

**SIP Architecture True-Detections with
Maximum-Likelihood and Support Vector Machine Classifiers**

□ Maximum-Likelihood ■ Support Vector Machine
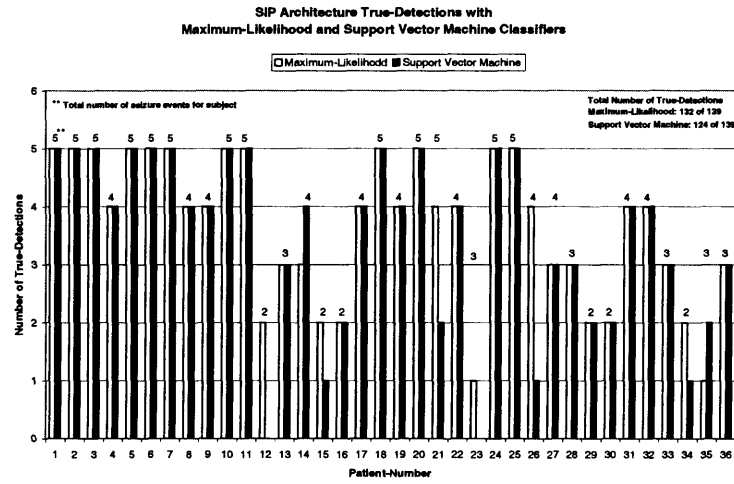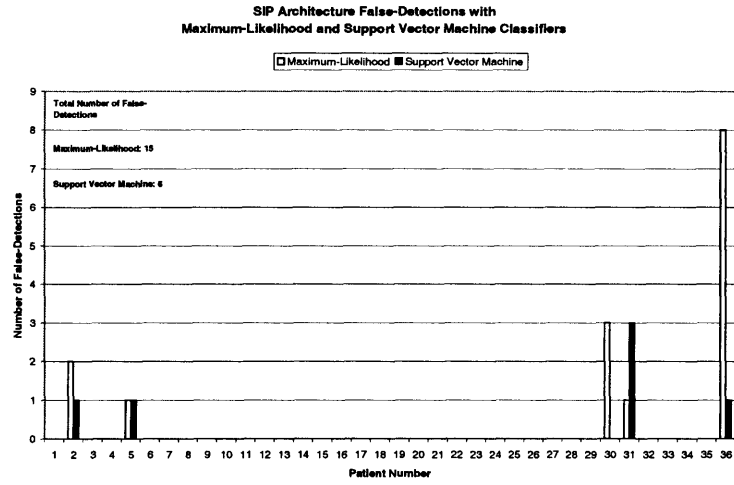
Figure 6-6: SIP Architecture False-Detections (Top) and True-Detections (Bottom)

97

result of non-physiological artifacts, all the false-detections declared by both detector types were caused by periodic discharges resembling the seizure onset activity of the particular subject. The maximum-likelihood classifier based detector was especially sensitive to the periodic discharges of subject 36, this lead to eight false-detections in twelve hours of processing.

Figure 6-6 also shows the true-detections declared on each test subject for both detector configurations; the number over each bar denotes the number of test seizures for a given subject. The discrepancy in true-detections between detector types is caused by the conservative choice of $C = [10 \; 10]$, which leads the support vector machine based detector to miss more seizures from subjects 12, 21, and 23. When $C = [30 \; 10]$ is used, the support vector machine based detector identifies the same number of seizures for these subjects as the maximum-likelihood detector, but at the expense of more false-detections on other subjects.

6.2.2 Spatially Dependant Processing Architecture

In the SDP architecture, localization constraints are encapsulated within feature vectors, so the detector's performance is influenced only by the required duration time T of an abnormality; the likelihood ratio threshold γ in the case of maximum-likelihood classifiers; and both the radial-basis kernel parameter σ and vector parameter C in the case of support vector machines.

The sensitivity in performance of a detector with the SDP architecture and maximum-likelihood classifiers due to different choices of the parameters T and γ is illustrated in Figure 6-7. The figure shows that for a given choice of T increasing the threshold γ results in an increase in the average detection latency, and a decrease in both the true-detections and false-detections measured for twenty-eight of the thirty-six subjects. Since we have chosen to optimize performance primarily in terms of latency and false-detections, we chose the parameter settings $T = 6$ seconds and $\gamma = 10^2$ because they provide an appropriate tradeoff between these metrics as shown by the circled data point in Figure 6-7.
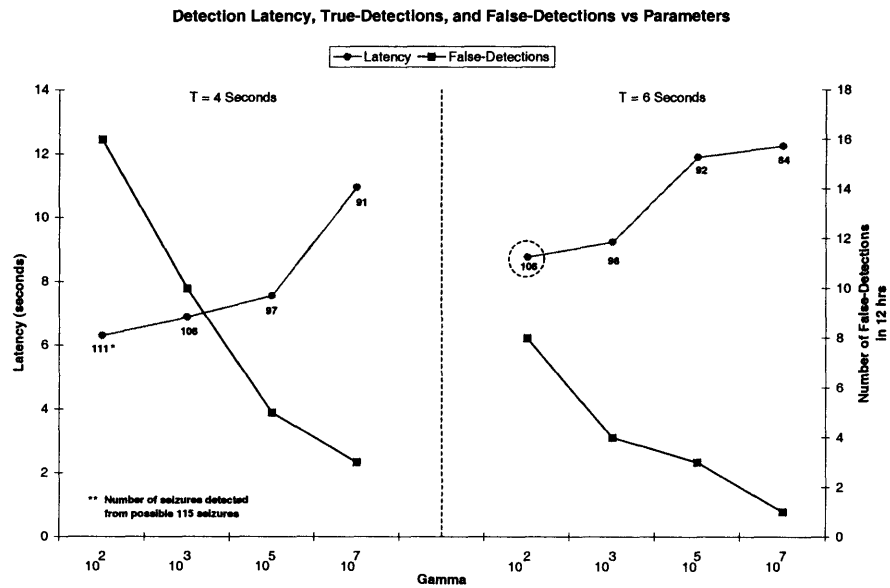
Figure 6-7: SDP Architecture Sensitivity with Maximum-Likelihood Classifier

Figure 6-8 illustrates the sensitivity of a detector that combines the SDP architecture with a support vector machine to different values of the parameter T; the settings $\sigma = 1$ and $C = [10\ 10]$ were fixed having observed their effects on performance through the SIP architecture. Figure 6-8 shows that increasing the parameter T increases the average detection latency and decreases both the true and false-detections measured for twenty-eight of the thirty-six subjects. For this detector configuration the parameter settings $T = 6$ seconds, $C = [10\ 10]$, and $\sigma = 1$ result in a tradeoff between detection latency and false-detections suitable for our application as shown by the circled data point in Figure 6-8.

Detection Latency, True-Detections, and False Detections vs Parameters with C = [10 10] and Sigma = 1
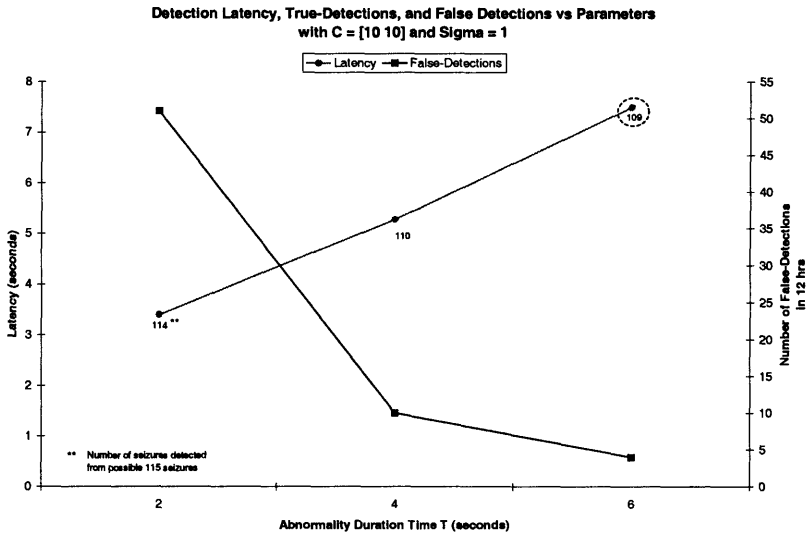
Figure 6-8: SDP Architecture Sensitivity with Support Vector Machine C=[10 10]

For the parameter settings $T = 6$ seconds, $\gamma = 10^2$, $C = [10\ 10]$, and $\sigma = 1$, Figure 6-9 shows the average detection latency of a detector that combines the SDP architecture with a maximum-likelihood classifier or a support vector machine. The latencies of both detector configurations are similar, which argues that the detector is not too sensitive to the classifier type. Furthermore, the detection latencies of most subjects are less than the proposed target latency of ten seconds by more than two seconds. The conservative choice of $C = [10\ 10]$ as well as gradual seizure onsets

**SDP Architecture Detection Latency with
Maximum-Likelihood and Support Vector Machine Classifiers**
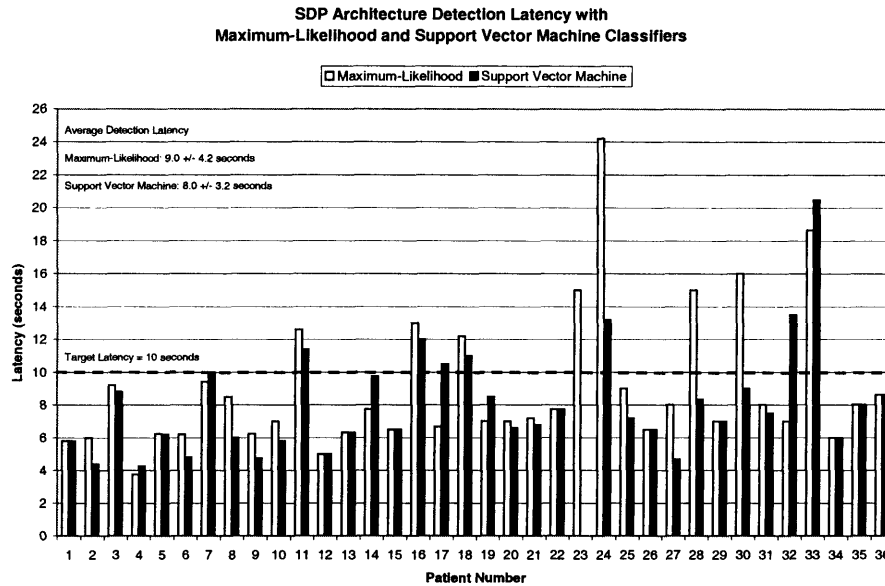
Figure 6-9: SDP Architecture Detection Latency

resulted in poor performance on subjects 23 and 24, while an artifact masking seizure onset activity on a number of derivations resulted in poor performance on subject 33. Coincidentally, the artifact did not affect the performance of the SIP architecture since it was not on the selected derivations. On the other hand, the SDP architecture does not exhibit a latency for subject 14 that is as large as that of the SIP architecture since there is no explicit setting for the minimum number of derivations required for a detection.

Figure 6-10 shows the false-detections declared on each test subject for both detector configurations. With the exception of subjects 9, 29, and 30 whose false-detections are a result of non-physiological artifacts, all other false-detections are a result of periodic discharges that resemble the seizure onset of a particular subject. The support vector machine based detector was more sensitive to discharges of subject 36.

Figure 6-10 also shows the true-detections declared on each test subject for both detector configurations. The difference in true detections is primarily caused by the three seizure events from subject 32 that were missed by the maximum-likelihood based detector. Unfortunately, the encapsulation of localization constraints within

101

the feature vectors of the SDP architecture makes it very difficult to explain why the seizure events were missed. Lowering the value of γ would most likely allow for the detection of these seizures at the cost of more false-detections.

Figure 6-10: SDP Architecture False-Detections (Top) and True-Detections (Bottom)

6.2.3 Comparison

The fundamental difference between the SIP and SDP architectures is the manner of representing and enforcing spatial localization constraints. In the case of the SIP

102

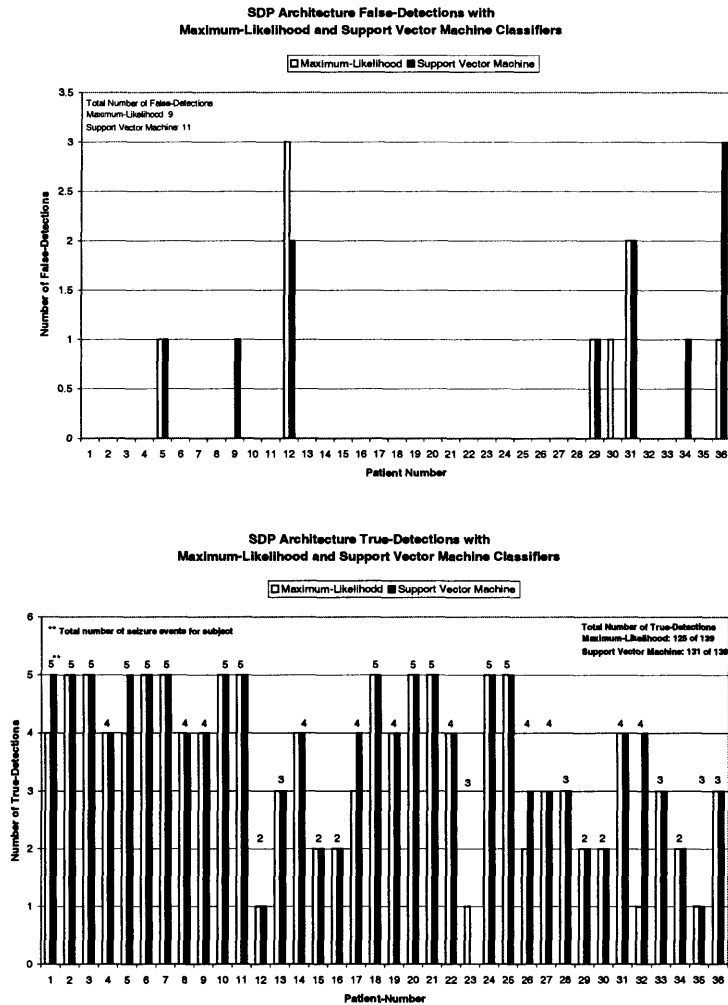architecture these constraints are imposed using explicit rules in the final element of the detector. This permits independent classification of activity on each derivation in a low dimensional feature space, and the skipping of derivations that are irrelevant to the detection of a seizure's onset. In contrast, the SDP architecture expresses spatial constraints through the interrelations of elements within a large feature vector summarizing activity from all derivations. While this obviates the need to explicitly enforce localization constraints, it hides from the user which derivations are being used for detection; and causes classification to take place in a higher dimensional space that includes features irrelevant to the detection of a given seizure's onset. Comparing the performance of these architectures will shed light on the question of which scheme of representing spatial constraints is more effective, and will also illustrate the robustness of the maximum-likelihood and support vector machine classifiers to feature vectors with irrelevant data.

SIP and SDP with Maximum-Likelihood Classifier

Figure 6-11 compares the performance of the SIP and SDP architectures when combined with the maximum-likelihood classifier. The two architectures exhibit similar detection latencies across all subjects, but the SIP architecture exhibits a slightly higher number of true-detections and six extra false-detections. All of the additional false-detections result form the periodic discharges of subject 36. The close performance of both detectors in terms of latency suggests that the maximum-likelihood classifier in the SDP architecture to a great extent ignored features from irrelevant derivations, and effectively exploited those crucial for detection of seizure onset. The results also argue that the SDP architecture does not exploit inter-derivation relations masked or lost by the independent processing of the SIP architecture.

The ability of a maximum-likelihood classifier to ignore features irrelevant to determining the class membership of an observed feature vector can be shown by reexpressing the likelihood ratio that the classifier compares to a threshold. To observe this, consider classifying a two-dimensional feature vector $X = [x_1 \ x_2]$ as an instance of the classes C_1 or C_2 when the feature x_1 is identically distributed conditioned on

103

Detection Latency of SIP and SDP Architectures with Maximum-Likelihood Classifier

□ SIP Architecture ■ SDP Architecture
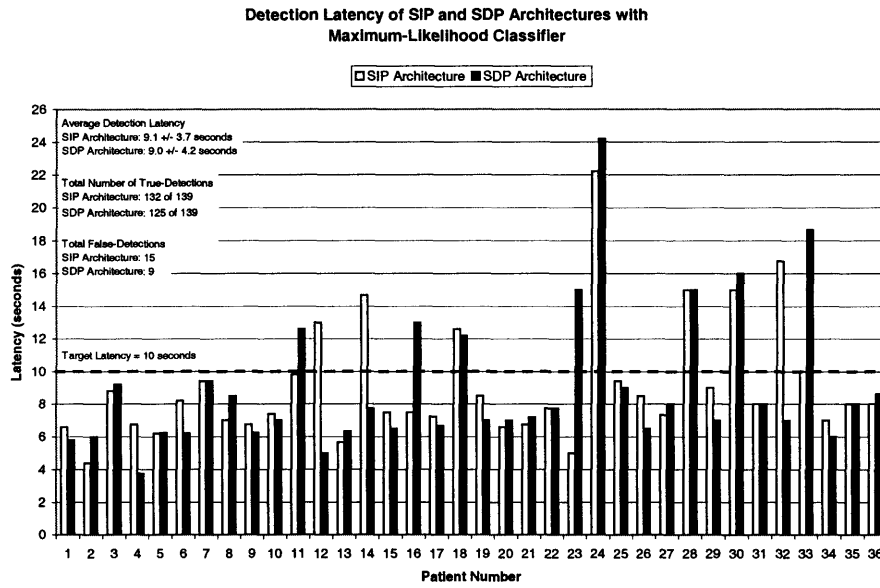
Figure 6-11: SIP and SDP Architecture Latency with Maximum-Likelihood Classifier

both classes, and is also independent of x_2. The decision rule for this problem is

$$\text{If } \frac{p(X|C_1)}{p(X|C_2)} \geq \gamma \text{ then } X \in C_1$$

Since the likelihood of X in this case can be reexpressed as $p(X) = p(x_1|x_2)p(x_2) = p(x_1)p(x_2)$, the decision rule can be rewritten as

$$\text{If } \frac{p(x_1|x_2, C_1)p(x_2|C_1)}{p(x_1|x_2, C_2)p(x_2|C_2)} = \frac{p(x_1|C_1)p(x_2|C_1)}{p(x_1|C_2)p(x_2|C_2)} \geq \gamma \text{ then } X \in C_1$$

Because x_1 is identically distributed conditioned on both classes, the likelihood $p(x_1|C_1) = p(x_1|C_2)$ and the decision rule simplifies to one that relies only on the feature x_2 for classification.

$$\text{If } \frac{p(x_2|C_1)}{p(x_2|C_2)} \geq \gamma \text{ then } X \in C_1$$

More generally x_1 and x_2 need not be independent. In such a case x_1 needs to be identically distributed conditioned on both classes and the feature x_2 for the above

104

result to hold since $p(X) = p(x_1|x_2)p(x_2) \neq p(x_1)p(x_2)$. In other words, for the decision rule to reduce to one that only relies on x_2, one must satisfy the stronger condition $p(x_1|x_2, C_1) = p(x_1|x_2, C_2)$.

SIP and SDP with Support-Vector Machine

Figure 6-12 compares the performance of the SIP and SDP architectures when each is combined with support vector machine classifiers. The SDP architecture exhibits a smaller detection latency and a higher number of true-detections relative to the SIP architecture, but a greater number of false-detections. The smaller average detection latency of the SDP architecture suggest 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.
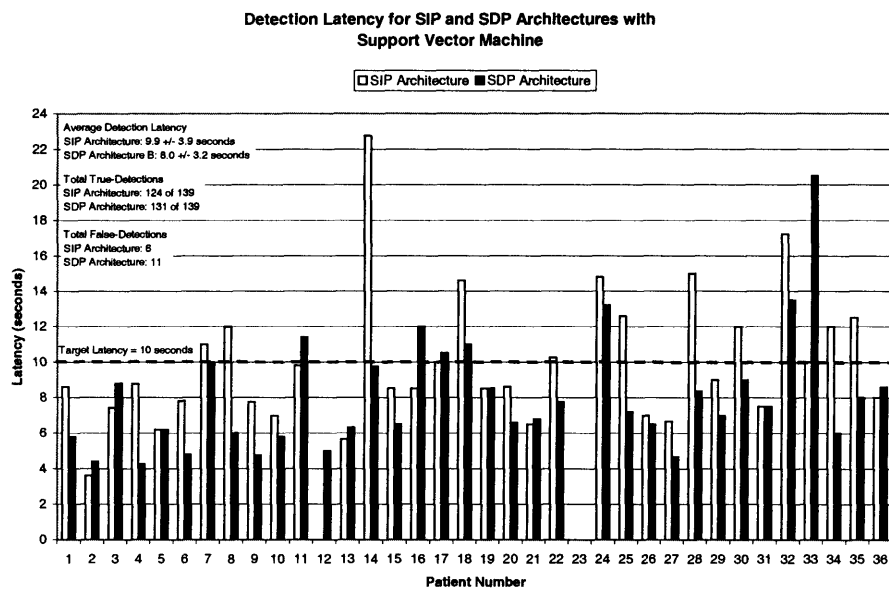

Figure 6-12: Latency of SIP and SDP Architectures with Support Vector Machine

6.3 Summary

In this chapter we defined the metrics used to characterize our detector's performance and outlined how they are computed. In terms of these metrics, we showed that both the SIP and SDP architectures satisfy our application's performance requirements. In particular, the detector that combines the SDP architecture with the support-vector machine classifier exhibited an average detection latency of 8.0± 3.2 seconds while correctly declaring 131 of 139 seizure events; and declared only 11 false-detections during 49 hours of randomly selected non-seizure EEG.

We also demonstrated how different choices for the detection parameters of the SIP and SDP architectures affect performance. We highlighted settings that optimize performance primarily in terms of latency and false-detections due to the lower cost of missing seizure events in our target application. In particular, we chose the parameter setting $T = 6$ seconds for both architectures to avoid a great number of false-detections caused by short-length, seizure-like periodic discharges occurring between true seizure events.

Finally, we observed that the SIP and SDP architectures perform similarly regardless of the classifier type. This allowed us to conclude that for the purpose of seizure onset detection, there is no gain or loss in performance resulting from the simultaneous or independent processing of derivations. This conclusion does not extend to seizure onset prediction since researchers have clearly demonstrated the need to consider the relation between derivations [18].

Nevertheless, the SDP architecture is both simpler to implement and more flexible due to its ability to automatically embed any spatial localization constraint within feature vectors. The SDP architecture is not limited to specifying localization constraints in terms of the group of derivations illustrated in Figure 5-10, and does not require the user to set a minimum number of derivations K that trigger a detection.

Chapter 7

Patient-Specific and Generic

Seizure Detection

This chapter contrasts the properties and performance of our detector when used in a patient-specific and nonpatient-specific, or generic mode. In the patient-specific mode, the detector is trained solely on previous examples of seizure and non-seizure EEG from the test subject. In the generic mode, it is trained on seizure and non-seizure EEG from a collection of subjects that excludes the test subject.

Section 7.1 highlights a property of the patient-specific detector that allows it to be of immediate, practical utility in a clinical setting; specifically, a high learning rate that results in excellent performance following observation of a very small number of seizures from the test subject. This issue is not considered for the generic detector since training data is always available and plentiful. Next, section 6.2.3 compares the performance of the two detector types, and highlights and important drawback to the generic approach.

7.1   Learning Rate

Figure 7-1 illustrates the improvement in a patient-specific detector's average detection latency and true-detection rate as function of the number of 20 minute EEG training recordings observed; a training recording includes a single seizure event as

well as non-seizure activity from a given subject. The figure highlights 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 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 performance beyond that obtained using two training recordings. In particular, a detector trained on three recordings detects on average 97% of a subjects future seizures with a mean latency of 7.1 ± 1.9 seconds. It is important to note both the decreasing mean latency, and the decreasing deviation about the mean as the number of training records is increased. These numbers were compiled using twenty-one[1] of the thirty-six test-subjects, which explains the deviation of the true detection rates and average detection latencies from those presented in Chapter 6. False-detections are not greatly affected by the number of training records observed, but primarily by the prevalence of a patient's seizure-like, interictal abnormalities and diversity of artifacts collected for inclusion in the training set.

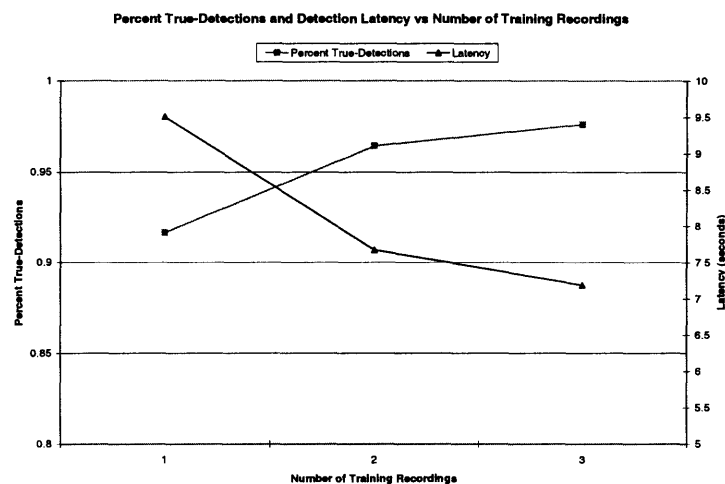Figure 7-1: Effect of Training on Patient-Specific Detector's Performance

---

[1]These twenty-one patients each had four or more recordings, which allowed us to evaluate and compare the detector's performance after training on one, two, or three recordings. If a patient were to have three or fewer recordings, then we would have been able to compare performance after training on one and two recordings, but not three.

The results presented in Figure 7-1 support the notion of seizures from a given subject being electrographically stereotypical, and argue that our patient-specific detector can reliably and quickly detect seizure onsets using as few as two training seizures. This is crucial in a clinical setting due to the paucity of data collection time, and the possible rarity of seizure events in some patients. To put these results in perspective, note that the generic detector discussed in the following section exhibits a 76% detection rate with an average detection latency of 12.3 ± 7.4 seconds even when trained on more than forty training recordings.

7.2   Comparison

A generic seizure onset detector is expected to perform poorly in terms of total number of true and false detections when compared to a patient-specific detector. This is primarily due to the heterogeneity of seizure onset patterns across patients, and the possible similarity between the non-seizure EEG of one subject and the seizure EEG of another. On the other hand, prior to this comparative study, the expected disparity between the average detection latency of a generic and patient-specific detector was not clear.

Figure 7-2 compares the performance of a patient-specific and generic detector. For each test subject the patient-specific detector was trained as described in section 6.1; while the generic detector was trained on the seizure and non-seizure EEG from all subjects excluding the one being tested. All performance metrics were computed using the manner described in section 6.1.

Performance tests were conducted for patients numbered 1-34, but results from patients $\{4, 29-34\}$ are omitted from Figure 7-2 due to exceedingly poor performance by the generic detector. In particular, the generic detector declared an excess of 50 false-detections on the EEG recordings of these subjects due to the similarity of seizure EEG from the training subjects and non-seizure EEG from the test subjects. In short, a generic detector always runs the risk of declaring many false-detections due to the possible similarity of training seizure and testing non-seizure EEG.

Patient-Specific vs. Generic Seizure Onset Detection
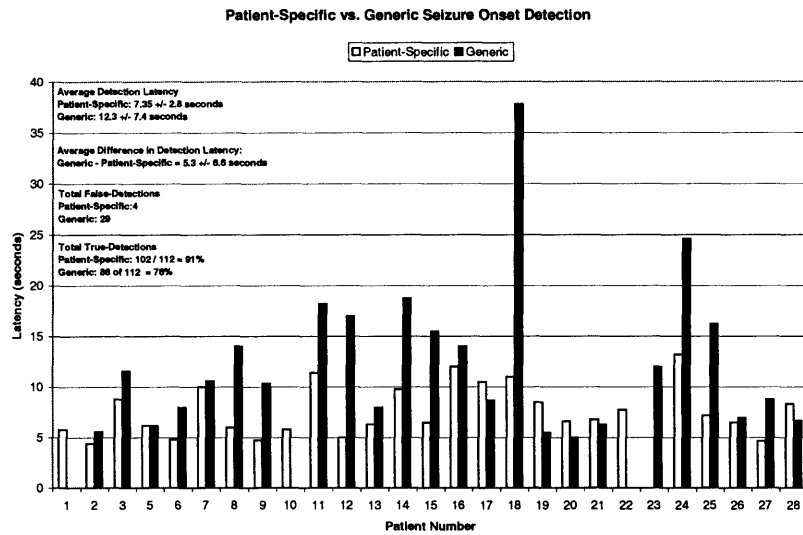
□ Patient-Specific ■ Generic

Figure 7-2: Comparison of Patient-Specific and Generic Seizure Detection

To convince ourselves that poor performance resulted from the seizure EEG of training subjects matching the non-seizure EEG of the test patients {4, 29 − 34}, we conducted the following experiment: We recorded the number of false-detections for each of {4, 29 − 34} when the detector was trained on a single training subject at a time, and then noted the training subjects resulting in fifteen or more false detections. Next, we constructed a training set for the test patients that exclude 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, 29 − 34} with only 2-10 false-detections.

On the remaining test subjects of Figure 7-2, 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 to the 91% detection rate and 4 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; the ±6.6 deviation from the smaller 5.3 average implies that there were subjects like 19 for

110

whom the generic detector outperformed the patient-specific detector. Finally, subject 27 is noteworthy since 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 phenomena may explain why the generic detector is able to identify one of three seizures from subject 23, while the patient-specific detector found none.

7.3 Summary

In this chapter we provided experimental evidence of our patient-specific detector's high learning rate; a property that allows it to exhibit excellent performance even when trained on as few as two seizure events from the test subject. Moreover, we demonstrated that a patient-specific detector exhibits a lower average detection latency; a lower total number of false-detections; and a higher total number of true-detections than a generic seizure onset detector. Our comparative study also underscored the likely event of a generic detector performing very poorly when the seizure EEG of a subject in the training set matches the non-seizure EEG of the test subject.

Chapter 8

Conclusion

In this chapter, we conclude the thesis with a summary of its goals and contributions followed by proposed improvements and directions for future work.

8.1 Goals and Contributions

The goal of this work was to design a real-time detector that identifies electrographic seizure onsets in scalp EEG for the purpose of initiating time-sensitive clinical procedures like ictal SPECT.

The detector's performance requirements were set according to the nature of clinical settings and the ictal SPECT procedure. In particular, the hectic clinical environment necessitated that the detector require minimal assistance from experienced hospital staff. Moreover, we required the detection and initiation of protocols to be within ten seconds of electrographic onset to improve upon the current ability to localize epileptogenic foci using ictal SPECTs. Finally, we required a low false-positive rate, or the ability to tradeoff more false-negatives for fewer false-positives due to the low cost of missed seizures in our target application.

We proposed a patient-specific seizure onset detector that extracts wavelet-based features from the EEG, and then employs a classifier to determine whether those features are indicative of seizure onset based on the learned features of a patient's seizure and non-seizure EEG. Our research makes the following contributions:

- **Provides an Effective Clinical Tool**

  Our detector exhibits the properties of an effective clinical tool. In particular the detector is simple to operate; broadly applicable; and has a high sensitivity and specificity.

  - **Simple To Operate:** To use the detector an electroencephalographer only needs to mark seizure onsets in EEG training records that are approximately thirty-minutes in duration; the number of records maybe as small as one and need not exceed three as demonstrated in Chapter 7. The detector automatically combines sections of the training records not marked as seizure with both artifact and generic EEG from various states of consciousness to form a representation of non-seizure activity, and to extract derivations to which the seizure onset localizes as discussed in Chapter 5.

  - **Broadly Applicable:** The detector's use of multi-scale, wavelet-based features allows it to detect seizure onsets with diverse electrographic manifestations. These manifestations include bursts of sharp waves, spike-and-slow-wave complexes, polymorphic waves, and rhythmic hypersynchrony of variable amplitude and frequency.

  - **Highly Sensitive and Specific:** When tested on the seizures of thirty-six de-indentified test subjects, the detector exhibited an average delay of $8.0\pm$ 3.2 seconds while correctly declaring 131 of 139 seizure events. Furthermore, the detector only declared 11 false-detections during 49 hours of randomly selected non-seizure EEG from these subjects. Most false-detections were caused by seizure-like, interictal abnormalities as illustrated by the third case study of Chapter 5.

- **Demonstrates Utility of Patient-Specificity in Seizure Detection.**

  We demonstrated through the comparative study of Chapter 7 the improved performance of a patient-specific seizure detector over a generic detector. The

114

patient-specific detector exhibited a lower average detection latency; a lower total number of false-detections; and a higher total number of true-detections. Our study also underscored the likely event of a generic detector performing very poorly when the seizure EEG of a subject in its training set matches the non-seizure EEG of the test subject.

- **Provides a Novel Perspective on Artifact Rejection.**

  We included various classes of EEG artifacts as part of the detector's non-seizure training set so that they can be identified and avoided through a learning methodology. This is in contrast to the more common approach of removing artifacts using traditional linear, nonlinear, or adaptive signal processing techniques.

- **Compares Alternative Approaches to Seizure Onset Detection**

  In Chapters 5-6, we designed and compared the performance of two different detector architectures that differ in how they capture and enforce the spatial localization constraints of seizure and non-seizure EEG. We also compared the efficacy of different classification schemes within each architecture; specifically, we experimented with both support vector machine and a maximum-likelihood classifiers. Our results demonstrated similar performance for the different architectures and classifiers.

8.2   Future Work

In this section, we present possibilities for future work that are likely to improve the current performance of our patient-specific seizure detection algorithm.

Making Detection Parameters Patient-Specific

In the SIP architecture seizure events are declared when K derivations exhibit abnormal activity for a duration T. The derivations vary for each subject, but their number K is fixed. This leads to a large detection latency for any subject with seizure onsets

that localize to a number of derivations less than K before spreading to include a number that is greater. By varying K automatically for each subject, large latencies resulting from an inappropriate fixed setting can be avoided.

One approach for automatically selecting K on a per-subject basis would use cross-validation on the training set. Specifically, for each candidate choice of K the detector would evaluate its performance on the available training set using leave-one-out cross-validation; then it would select the value of K with the best overall performance for use in the detection of future seizures from the same subject.

The selection procedure just outlined could also be used to select T for each subject. However, the prevalence of short-duration, patient-specific pre-seizure abnormalities that otherwise resemble seizure onset activity is likely to result in final values of T that are very similar across patients, and close to our fixed choice.

Using Pre-Seizure Abnormalities to Enhance Seizure Onset Detection

Prior to the electrographic onset of a seizure, the EEG in some subjects exhibits patient-specific abnormalities that may or may not resemble the electrographic signature of the onset. For instance, the electrographic onset of a subject may be preceded by a combination of time-limited discharges and epochs of electrocerebral inactivity that remain constant or change in duration and frequency of repetition as the onset nears.

A detector that not only recognizes this patient-specific pre-seizure activity, but also estimates from training data the duration of time separating the beginning of this activity and the actual seizure onset would be an effective seizure forewarning device. However, even for a specific subject, the time between the beginning of pre-seizure activity and seizure onset may be variable; and for any given observation of the pre-seizure activity the seizure may or may not actually occur. These difficulties will directly influence how long prior to a seizure event the detector can reliably make a prediction.

If a prediction is not desirable, the detector could instead enter a mode of higher-sensitivity for an allotted duration of time upon detecting pre-seizure activity. The

116

higher-sensitivity mode may result in smaller detection latencies, and can be attained by decreasing the minimum detection duration T, the minimum number of derivations K, or by increasing the tolerance α for the probability of a false-detection. If no seizure is detected within the allotted window of time, the detector can return to its original sensitivity.

Using Generic Data to Overcome Deficient Patient-Specific Training Sets

Our study of generic seizure detectors underscored the possibility of poor performance whenever the seizure EEG of a subject in the training set matches the non-seizure EEG of the test subject. The study also highlighted that a generic detector could exhibit performance close to that of a patient-specific detector when seizure and non-seizure EEG in the training set closely matched that of the test subject. These results argue for augmenting a patient-specific detector with data from a group of patients with similar EEG characteristics whenever the training set from the test subject is insufficient.

The process of automatically choosing other subjects to include in the training set of a patient-specific detector is known as *active learning*. Its main challenge is to devise metrics that accurately reflect similarities between the seizure and non-seizure EEG of two patients in terms of amplitude, fundamental frequency, morphology, and localization of seizure and non-seizure activity. Once these metrics are developed, a large collection of patients can be easily screened for subjects to augment the training set of a patient-specific detector.

Deployment

Figures 8-1 and 8-2 illustrate two strategies for using our detector to initiate Ictal SPECTs following electrographic seizure onsets. In the expert-assisted strategy of Figure 8-1, the detector is first trained on the seizure and non-seizure EEG of a patient and then set to monitor their streaming EEG. When seizure activity is detected, an electroencephalographer is immediately notified and automatically shown both the suspected electrographic pattern and a live video of the patient. The elec-

117

troencephalographer uses this information to confirm with the detector whether or not to infuse the Ictal SPECT radioisotope. Although the expert-assisted strategy introduces a delay by requiring confirmation of a seizure event, it has the advantage of producing no false injections of the SPECT radioisotope by the detector.
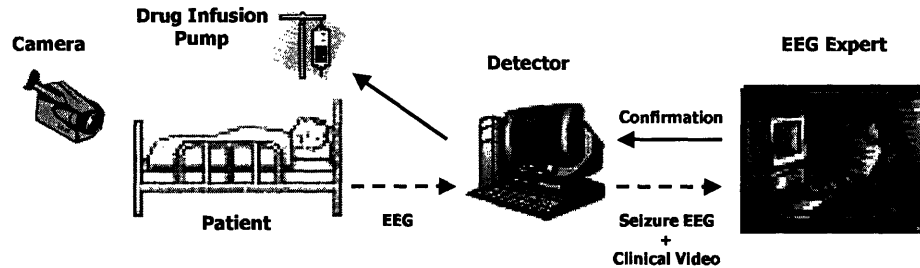
Figure 8-1: Expert-Assisted Strategy For Initiating Ictal SPECTs Using Detector

In the automated strategy of Figure 8-2, the detector is again trained on the seizure and non-seizure EEG of a patient and then set to monitor their streaming EEG. When seizure activity is noted the detector automatically initiates the infusion of the ictal SPECT radioisotope without incurring further delay by awaiting a seizure event confirmation; the decrease in latency comes at the expense of an incorrect injection occurring at the rate of the detector's false-detection rate.

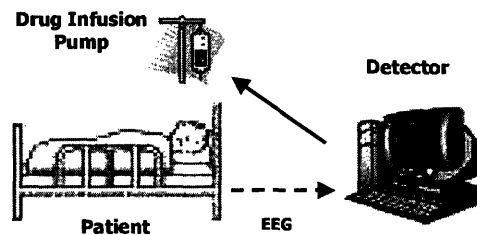Figure 8-2: Automated Strategy For Initiating Ictal SPECTs Using Detector

To determine the efficacy of our detector as well as which deployment strategy is more effective, one needs to conduct a clinical trial comparing the accuracy and consistency of seizure focus localization obtained across several patients using the two deployment methods. If the trial demonstrates that the extra delay incurred

by confirming seizure events in the expert-assisted strategy does not greatly affect localization, then that strategy should be used since it greatly limits the possibility of a false-injection. On the other hand, if the trial demonstrates that the minimal delay of the automated strategy leads to much more accurate localizations, then one may need to accept the possibility of an incorrect injection occurring at the rate of the detector's false-detection rate.

8.3 Summary

This thesis discusses the design and validation of a real-time, patient-specific method that can be used to detect the onset of epileptic seizures in non-invasive EEG, and then initiate time-sensitive clinical procedures like ictal SPECT. We adopt a patient-specific approach because of the clinically observed consistency of seizure and non-seizure EEG characteristics within patients, and their great heterogeneity across patients. Furthermore, we treat patient-specific seizure onset detection as a binary classification problem. Our observation is a multi-channel EEG signal; its features include amplitude, fundamental frequency, morphology, and spatial localization on the scalp; and it is classified as an instance of non-seizure or seizure EEG based on the learned features of training examples from a single patient as well as a library that includes generic EEG artifacts and hallmark activity from different states of consciousness.

We use a multi-level wavelet decomposition to extract features that capture the amplitude, fundamental frequency, and morphology of EEG waveforms. These features are then classified using a support vector machine or maximum-likelihood classifier trained on a patients seizure and non-seizure EEG; non-seizure EEG includes normal and artifact contaminated EEG from various states of consciousness. The outcome of the classification is examined in the context of automatically extracted spatial and temporal constraints before the onset of seizure activity is declared.

During validation tests our method exhibited an average latency of 8.0± 3.2 seconds while correctly identifying 131 of 139 seizure events from thirty-six, de-identified

test subjects; and only 11 false-detections over 49 hours of randomly selected non-seizure EEG from these subjects. The validation tests also highlight the high learning rate of the detector; a property that allows it to exhibit excellent performance even when trained on as few as two seizure events from the test subject.

We also demonstrate through a comparative study that our patient-specific detector outperforms a nonpatient-specific, or generic detector in terms of a lower average detection latency; a lower total number of false-detections; and a higher total number of true-detections. Our study also underscores the likely event of a generic detector performing very poorly when the seizure EEG of a subject in its training set matches the non-seizure EEG of the test subject.

This research has lead to an effective clinical tool that has the potential to initiate ictal SPECTs in a more timely and consistent fashion, and consequently improve the localization of epileptogenic foci.

Bibliography

[1] A.B.Geva and D.H.Kerem. Forecasting generalized epileptic seizures from the eeg signal by wavelet analysis and dynamic unsupervised fuzzy clustering. *IEEE Transactions On Biomedical Engineering*, 45(10):1205–1216, October 1998.

[2] A.V.Oppenheim, R.W.Schafer, and J.R.Buck. *Discrete-Time Signal Processing*, chapter 8, pages 582–588. Prentice-Hall Inc., second edition, 1989.

[3] B.Boashash, A.M.Zoubir, and M.Roessgen. On-line detection of seizure in newborn eeg using signal processing tools. *Proceedings of IEEE International Conference on Digital Signal Processing*, 1:79–82, 1997.

[4] B.Litt. Engineering devices to treat epilepsy: A clinical perspective. *2001 Proceedings of the 23rd Annual EMBS International Conference, October 25-28, Istanbul, Turkey*.

[5] Private communication with the staff of the Boston Children's Hospital.

[6] D.W.Klass and D.D.Daly. *Current Practice of Clinical Electroencephalography*. Raven Press, 1979.

[7] E.Waterhouse. New horizons in ambulatory electroencephalography. *IEEE Engineering in Medicine and Biology Magazine*, pages 74–80, May/June 2003.

[8] F.S.Tyner, J.R.Knott, and W.B.Mayer Jr. *Fundamentals of EEG Technology: Basic Concepts and Methods*, volume Volume 1. Lippincott Wiliams & Wilkins, 1983.

[9] G.Strang and T.Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1997.

[10] H.Qu and J.Gotman. Improvement in seizure detection performance by automatic adaptation ot the eeg of each patient. *Electroencephalography and Clinical Neurophysiology*, 86:79–87, 1993.

[11] H.Qu and J.Gotman. A patient-specific algorithm for the detection of seizure-onset in long-term eeg monitering: Possible use as a warning device. *IEEE Transactions On Biomedical Engineering*, 44(2):115–122, February 1997.

[12] I.E.Leppik. *Contemporary Diagnosis and Management of the Patient with Epeilepsy*. Handbooks in Health Care, Newton,Pennsylvania,USA, fifth edition, 2000.

[13] J.Glover, N.Raghavan, P.Ktonas, and J.Frost. Context-based automated detection of epileptogenic sharp transients in the eeg: Elimination of false postives. *IEEE Transactions On Biomedical Engineering*, 36(5):519–527, May 1989.

[14] J.Gotman. Automatic recognition of epileptic seizures in the eeg. *Electroencephalography and Clinical Neurophysiology*, 54:530–540, 1982.

[15] J.Gotman, J.R.Ives, and P.Gloor. Frequency content of eeg and emg at siezure onset: Possibility of removal of emg artifact by digital filtering. *Electroencephalography and Clinical Neurophysiology*, 52:626–639, 1981.

[16] J.Gotman and P.Gloor. Automatic recognition and quantification of interictal epileptic activity in the human scalp eeg. *Electroencephalography and Clinical Neurophysiology*, 41:513–529, 1976.

[17] L.Hively and A.Protopopescu. Channel-consistent forewarning of epileptic events from scalpe eeg. *IEEE Transactions On Biomedical Engineering*, 50(5):584–592, May 2003.

[18] L.Iasemidis, D.Shiau, W.Chaovalitwongse, J.Sackellares, P.Pardalos, J.Principe, P.Carney, A.Prasad, B.Veeramani, and K.Tsakalis. Adaptive epileptic seizure prediction system. *IEEE Transactions On Biomedical Engineering*, 50(5):616–625, May 2003.

[19] L.Meng, M.Frei, I.Osorio, G.Strang, and T.Nguyen. Detection and short-term prediction of epileptic seizurs from ecog signal by wavelet analysis and gaussian mixture model. February 2000.

[20] L.Tarassenko, Y.U.Khan, and M.R.G Holt. Identification of inter-ictal spikes in the eeg using neural network analysis. *IEEE Proc.-Sci. Meas. Technol.*, 145(6):270–278, November 1998.

[21] M.A.Hearst. Support vector machines. *IEEE Intelligent Systems*, pages 18–28, 1998.

[22] M.D'Alessandor, R.Esteller, G.Vachtsevanos, A.Hinson, J.Echauz, and B.Litt. Epileptic seizure prediction using hybrid feature selection over multiple intracranial eeg electrode contacts: A report of four patients. *IEEE Transactions On Biomedical Engineering*, 50(5):603–615, May 2003.

[23] N.Cristianini and J.Shawe-Taylor. *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[24] N.Hazarika, J.Z.Chen, A.C.Tsoi, and A.Sergejew. Classification of eeg signals using the wavelet transform. *Signal Processing*, 59:61–72, 1997.

[25] P.E.McSharry, L.A.Smith, and L.Tarassenko. Comparison of predictability of epileptic seizures by a linear and nonlinear method. *IEEE Transactions On Biomedical Engineering*, 50(5):628–633, May 2003.

[26] R.O.Duda, P.E.Hart, and D.G.Strok. *Pattern Classification*. John Wiley & Sons Inc., second edition, 2000.

[27] S.M.Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*, volume 2, chapter 3, pages 77–80. Prentice Hall Inc., 1993.

[28] T.Betts. Epilepsy at the millenium. Technical report, Birmingham University Seizure Clinic and Epilepsy Liaison Service, 1998.

[29] W.L.Martinez and A.R.Martinez. *Computational Statistics Handbook with Matlab*, chapter 8, pages 280–287. Champman & Hall/CRC, 2000.