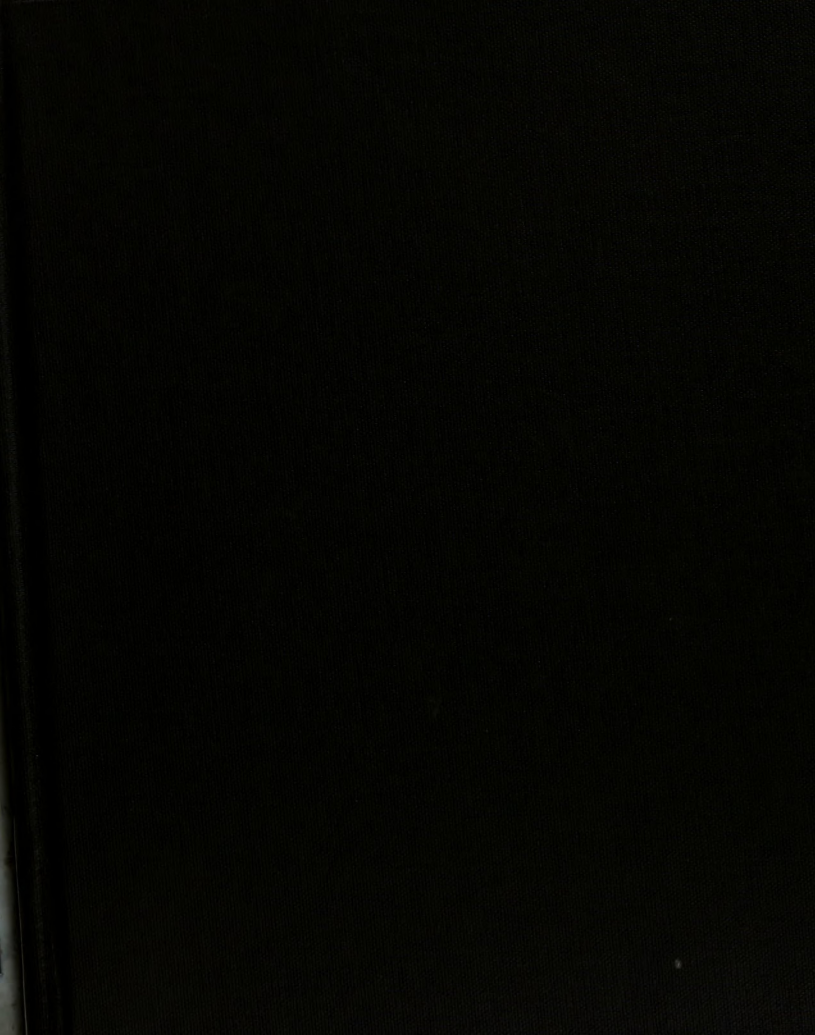






3 1293 01710 2801

This is to certify that the

thesis entitled

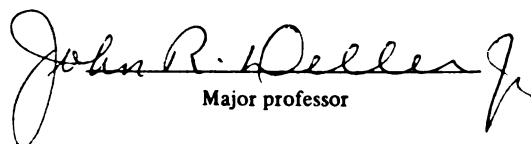
**DEVELOPMENT OF THE DETECTOR STAGE OF AN
ADAPTIVE ALERTING DEVICE FOR PEOPLE
WITH HEARING DISABILITIES**

presented by

VIKTOR ADUT

has been accepted towards fulfillment
of the requirements for

 M.S. degree in ELECTRICAL ENG.


Major professor

Date 2 April 98

LIBRARY

Michigan State University

PLACE IN RETURN BOX
 to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
SEP 12 2002 08 23 02		
NOV 03 2003 12 09 04 APR 10 2004		
MAR 10 2006		

**DEVELOPMENT OF THE DETECTOR STAGE
OF AN ADAPTIVE ALERTING DEVICE FOR
PEOPLE WITH HEARING DISABILITIES**

By

Viktor Adut

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Electrical Engineering

1998

ABSTRACT

DEVELOPMENT OF THE DETECTOR STAGE OF AN ADAPTIVE ALERTING DEVICE FOR PEOPLE WITH HEARING DISABILITIES

By

Viktor Adut

Since their inception, devices which signal deaf people to the occurrence of alerting signals (fire alarms, doorbells, etc.) have changed little. Their use is limited to highly constrained environments and relatively few sounds. This thesis reports the results of the first phase of a commercial research project between Michigan State University's Speech Processing Laboratory and Silent Call Corporation whose goal is to develop a digital-signal-processing-based portable device which "listens" to the environment and detects the presence of alerting signals. The device will also have the capability to "learn" new sound sets. In this work, algorithms for detection of alerting signals under realistic noise environments are developed, and initial simulation results quantifying the performance of these algorithms are presented.

ACKNOWLEDGMENTS

I am indebted to Professor John R. Deller, Jr. for the guidance he provided. I would like to thank to my committee members, Professor Marvin Siegel and Professor Robert Nowak, for their insightful comments. I am also thankful to Mr. Dale Joachim (M.S., Ph.D. candidate) for numerous invaluable discussions, especially during the initial stages of the project.

This study is based in part on work supported by State Research Fund grant monies awarded to Silent Call Corporation by the Michigan Jobs Commission. Any opinions, findings, conclusions, or recommendations expressed in this thesis are solely mine and do not necessarily reflect the views of the Michigan Jobs Commission.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
1 Introduction	1
2 Detector Architecture	3
2.1 Overview	3
2.2 Operating Environment	4
2.2.1 Classification of Alerting Signals	4
2.2.2 Noise Environment	13
2.3 Sampling Stage	19
2.4 Choosing a Detection Strategy	20
2.4.1 Spectral Thresholding	22
2.4.2 Dynamic Time Warping and Hidden Markov Models	22
2.4.3 Artificial Neural Networks	23
2.5 A Functional Block Description of Detector Architecture	23
3 Detector Design	26
3.1 Design Methodology	26
3.2 Detection Filters	27
3.2.1 Matched Filter	29
3.2.2 Eigenfilter	30
3.2.3 Prediction Error Filter	30
3.2.4 Autocorrelation Detector	31
3.3 Decision Block	33
3.3.1 Synchronization	33
3.3.2 Decision Rule	36
3.4 Postprocessor	38

4	Performance Evaluation	40
4.1	Introduction	40
4.2	Benchmarking with the SNR Improvement Factor	41
4.2.1	Formal Developments	41
4.2.2	Simulation Results	43
4.3	Benchmarking Using Detection Rates	44
4.3.1	Introduction	44
4.3.2	Performance of the Decision Rule	47
4.3.3	Performance of the Overall Detector	49
5	Conclusions and Considerations for Future Research	63
5.1	Conclusions	63
5.2	Future Research	64
	BIBLIOGRAPHY	65

LIST OF TABLES

4.1	Design details of the detection filters used in the simulations.	45
4.2	SNR improvement factors in decibels for matched filter.	46
4.3	SNR improvement factors in decibels for prediction error filter.	46
4.4	SNR improvement factors in decibels for eigenfilter.	46
4.5	SNR improvement factors in decibels for autocorrelation detector. . .	46

LIST OF FIGURES

2.1	A fire alarm signal consisting of a stationary region and a region of silence.	5
2.2	A doorbell signal consisting of two stationary regions.	7
2.3	A police siren signal consisting of two stationary signal segments connected by a transition region.	9
2.4	A crying baby signal consisting of two stationary signal segments connected by transition regions.	11
2.5	Several example stationary signal segments and their frequency spectra.	14
2.6	Several example noise waveforms.	17
2.7	The sampling stage.	21
2.8	The basic detector architecture.	25
3.1	Framewise processing of a stationary signal segment.	35
4.1	Power curves of the matched filter detector for $L=750$ samples and several values of α	50
4.2	Power curves of the matched filter detector for $L=1000$ samples and several values of α	51
4.3	Power curves of the matched filter detector for $L=1500$ samples and several values of α	52
4.4	Power curves of the eigenfilter detector for $L=750$ samples and several values of α	53
4.5	Power curves of the eigenfilter detector for $L=1000$ samples and several values of α	54
4.6	Power curves of the eigenfilter detector for $L=1500$ samples and several values of α	55
4.7	Power curves of the autocorrelation detector for $L=750$ samples and several values of α	56
4.8	Power curves of the autocorrelation detector for $L=1000$ samples and several values of α	57

4.9	Power curves of the autocorrelation detector for $L=1500$ samples and several values of α	58
4.10	Overall error rate of the matched filter detector for $L=1500$ samples and several values of α	60
4.11	Overall error rate of the eigenfilter detector for $L=1500$ samples and several values of α	61
4.12	Overall error rate of the autocorrelation detector for $L=1500$ samples and several values of α	62

CHAPTER 1

Introduction

During the past decades advances in biomedical engineering combined with breakthroughs in microelectronics resulted in significant improvements in the devices for people with hearing disabilities. For people with mild to profound hearing impairments, the most widely used solution is a classical hearing aid fitted in the ear canal. If the auditory nerves are damaged, a cochlear implant consisting of clusters of electrodes which directly excite the nerves in cochlea can be considered [8]. However, these techniques are of limited success when hearing loss is severe. They result in poor sound quality and the problems can be so serious that the owner may opt not to use them unless absolutely necessary, or reject them totally [18].

If neither of these solutions helps, the inability to respond to emergency signals becomes a potentially life threatening problem. In this case the deaf person must use alerting devices that signal the occurrence of acoustic events like doorbells, telephones, smoke detectors, *etc.* Such alerting devices have changed little over the past fifty years. They usually consist of a transmitter hard-wired to the signaling device which activates a blinking light or a vibratory device carried by the system user. The operation of these devices is extremely constrained and their use is limited to the everyday environments of the deaf person.

Recently, a digital-signal-processing (DSP) based adaptive alerting device which solves these limitations was proposed, and its feasibility was studied by [5]. Using DSP technology, one can develop a portable device which “listens” to the environment and detects the presence of alerting signals. One can also equip this device with the capability to “learn” new sound sets. The current advances in microprocessor technology provide the required computational power for real-time implementation of such algorithms at low cost. Given the fact that there are 28 million individuals with with hearing impairments in United States alone [18], the world-wide utility of this device is enormous.

This thesis was written as part of a joint research project between Michigan State University’s Speech Processing Laboratory and Silent Call Corporation whose ultimate goal is to develop a commercially-available adaptive alerting device for mass markets. The results of the first phase of this project are presented here.

Before studying the development of detection algorithms, we must concisely state the specifications of the alerting device. Our goal is to develop a device which can detect the presence of alerting signals (sirens, doorbells, crying babies, *etc.*) under realistic noise environments encountered in everyday life (people talking, highway noise, air conditioner noise, *etc.*). The algorithm must be computationally efficient, work with limited memory, and lend itself to trainability.

In this study we pursue a detection theoretic approach for the development of the alerting device, and focus on the development of the detector stage of the adaptive alerting device. In Chapter 2, the basic structure of the detector stage is developed in terms of functional blocks. In Chapter 3, several detection strategies are examined from a theoretical standpoint. Results of simulation studies are presented in Chapter 4. The final chapter discusses open points, and outlines the course of future research.

CHAPTER 2

Detector Architecture

2.1 Overview

The goal of this chapter is to describe the detector stage of the alerting device in terms of functional blocks. The primary factors affecting any detection algorithm are the properties of the signals to be detected and those of the noise environment; therefore we start by classifying alerting signals according to the complexity of their time-domain waveforms and performing an analysis of the diverse acoustic noise environments under which the alerting device is expected to operate. In Section 2.3, we turn our attention to the sampling stage and explain how the automatic gain controller (AGC) used in this stage complicates the detection procedure. Next, in Section 2.4, we justify the decision to apply a detection-theoretic technique to the development of the adaptive alerting device, and discuss the shortcomings of other possible approaches. Finally, in Section 2.5, our observations lead to the formulation of an alerting signal detector consisting of three stages.

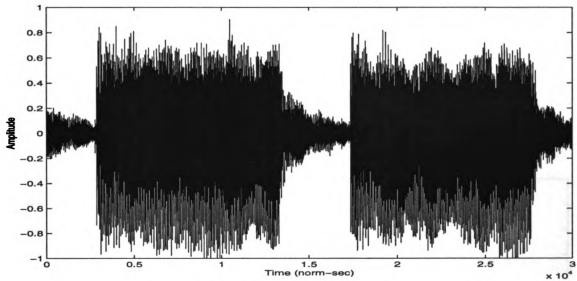
2.2 Operating Environment

2.2.1 Classification of Alerting Signals

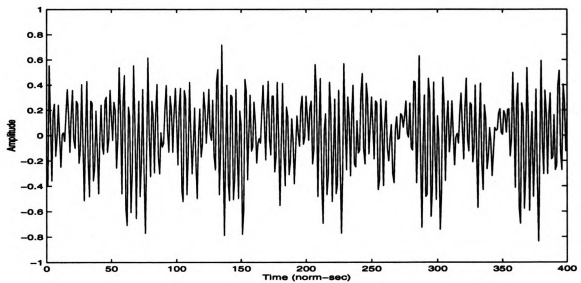
Let us consider Figures 2.1-2.4 which show frames of several alerting signals. Each frame is indexed beginning from time 0 and is long enough to represent the salient properties of the individual alerting signals. As every example signal in this thesis, they were sampled at 8 kHz and quantized to 8 bits. First, let us focus on the fire alarm signal in Figure 2.1. Every realization of this signal consists of the repetition of a stationary signal segment followed by a region of silence. The duration of these regions is fixed for every realization. Next, let us examine the doorbell signal shown in Figure 2.2. It is composed of two stationary signal segments connected. Depending on how quickly one pushes and releases the doorbell button, the length of each signal section in the generated signal will be different; however their order will be same. Finally, let us analyze the crying baby signal shown in Figure 2.4. This signal consists of several quasi-stationary signal segments connected by transition regions; both their duration and temporal order will be different every time the baby cries.

Thus we conclude that alerting signals are *locally stationary*; based on the above observations we can classify them according to increasing complexity as follows:

- Type I alerting signal: The temporal order and duration of each stationary signal segment is fixed for every realization (*e.g.*, fire alarm).
- Type II alerting signal: The temporal order of stationary signal segments is same across realizations; however their durations are different (*e.g.*, doorbell).
- Type III alerting signal: Both the temporal order and durations of stationary signal segments are variable across realizations (*e.g.*, crying baby).



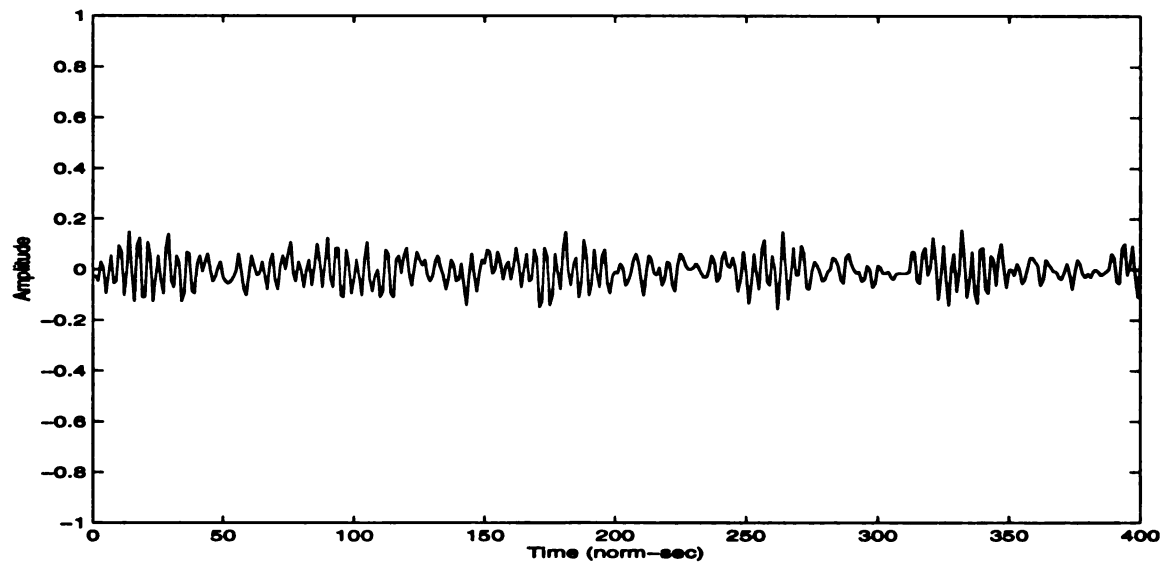
(a) A fire alarm signal.



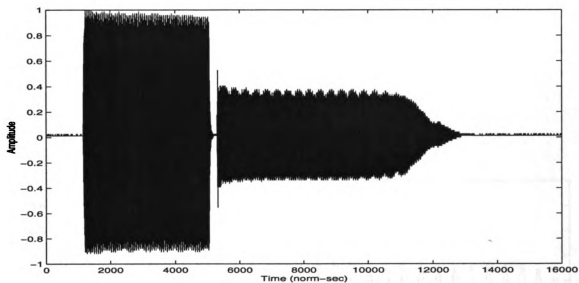
(b) A stationary signal segment.

Figure 2.1. A fire alarm signal consisting of a stationary region and a region of silence.

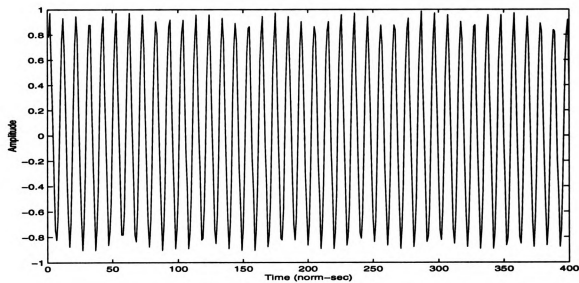
Figure 2.1. (cont'd).



(c) A region of silence.



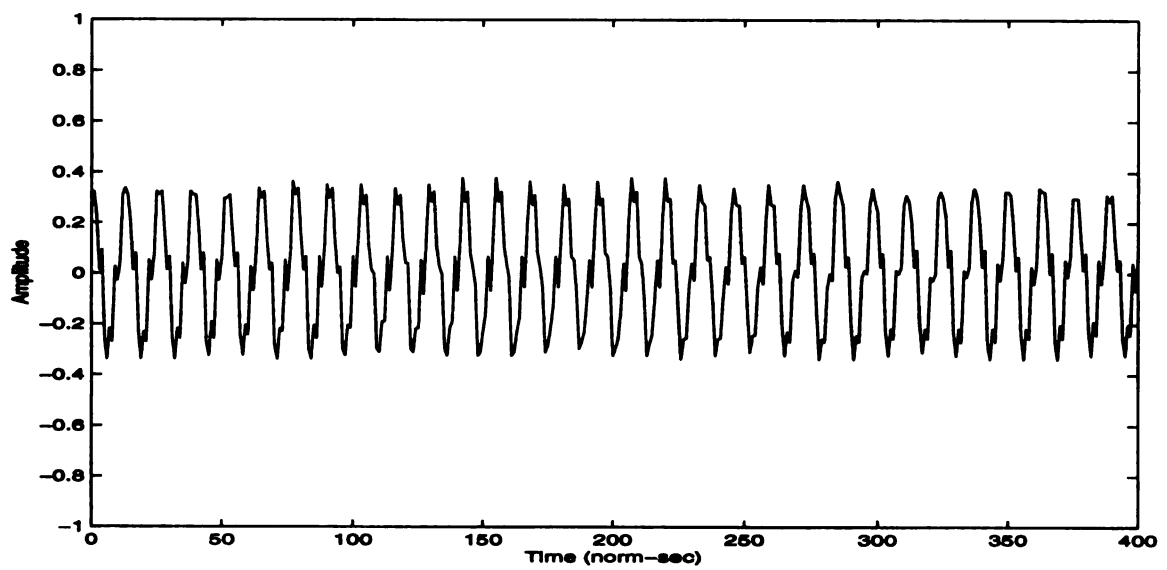
(a) A doorbell signal.



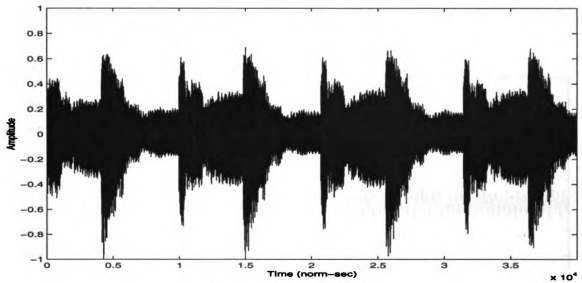
(b) First stationary signal segment.

Figure 2.2. A doorbell signal consisting of two stationary regions.

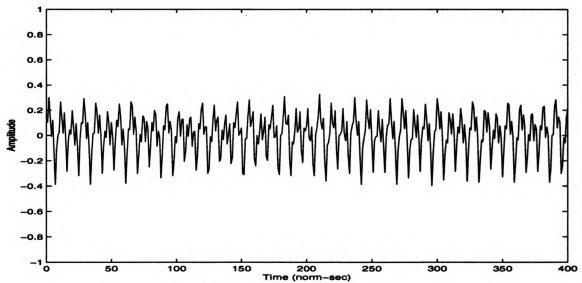
Figure 2.2. (cont'd).



(c) Second stationary signal segment.



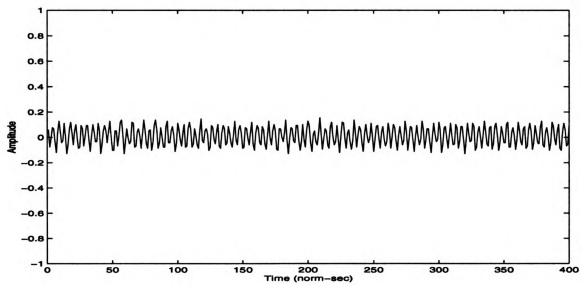
(a) A police siren signal.



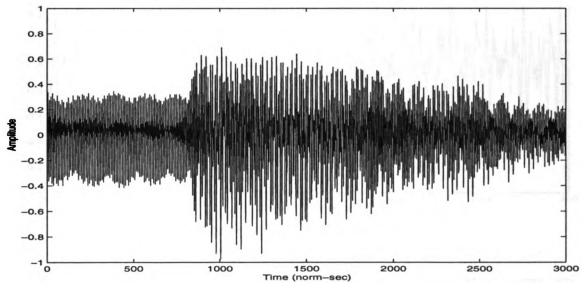
(b) First stationary signal segment.

Figure 2.3. A police siren signal consisting of two stationary signal segments connected by a transition region.

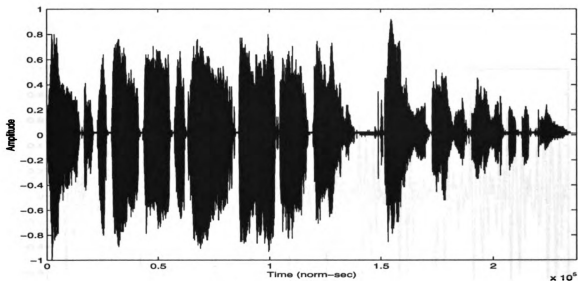
Figure 2.3. (cont'd).



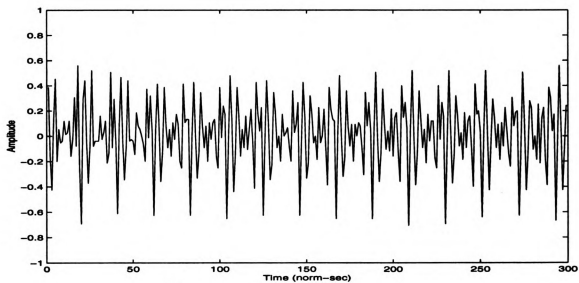
(c) Second stationary signal segment.



(d) Transition region.



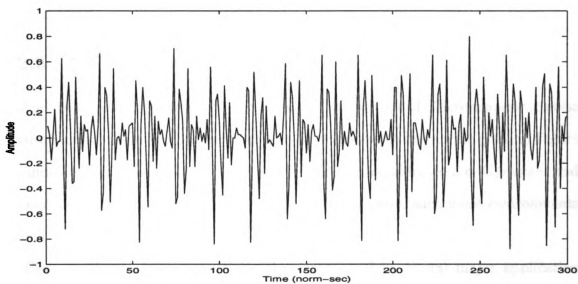
(a) A crying baby signal.



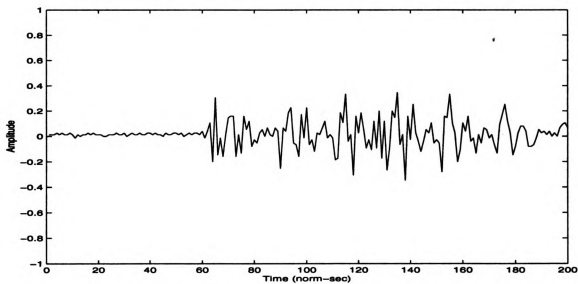
(b) First stationary signal segment.

Figure 2.4. A crying baby signal consisting of two stationary signal segments connected by transition regions.

Figure 2.4. (cont'd).



(c) Second stationary signal segment.



(d) An example transition region.

Next, let us briefly consider the spectral properties of the stationary signal segments. From Figure 2.5, we see that the energy of the stationary signal segments of interest is concentrated in narrow spectral peaks.

In the subsequent sections we exploit the conclusions of the above analysis to derive a detection scheme tailored to the properties of alerting signals.

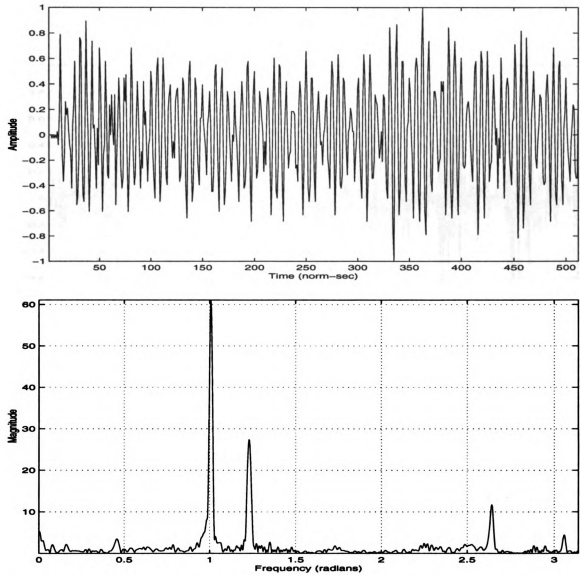
2.2.2 Noise Environment

The adaptive alerting device will operate under a multitude of different acoustic noise environments encountered in everyday life. As the example waveforms given in Figure 2.6 indicate, the noise environment can be described as additive, non-Gaussian, and nonstationary. With the exception of very restricted cases, such noise environments cannot be reliably modeled [9].

Comparable situations are frequently encountered in several major application areas of signal processing, such as in communication and radar systems. In these cases, adaptive noise cancelers are widely used to reduce background noise [6]. However, for reasons explained in Section 2.3, our inability to estimate the instantaneous signal and noise powers bars us from utilizing an adaptive noise canceler.

Due to these difficulties, in the remainder of the thesis we shall not attempt to model the noise environment and develop the detection algorithm under the assumption that the adaptive alerting device operates under independent, identically distributed Gaussian noise. The justification for this assumption will be provided in Section 3.2.

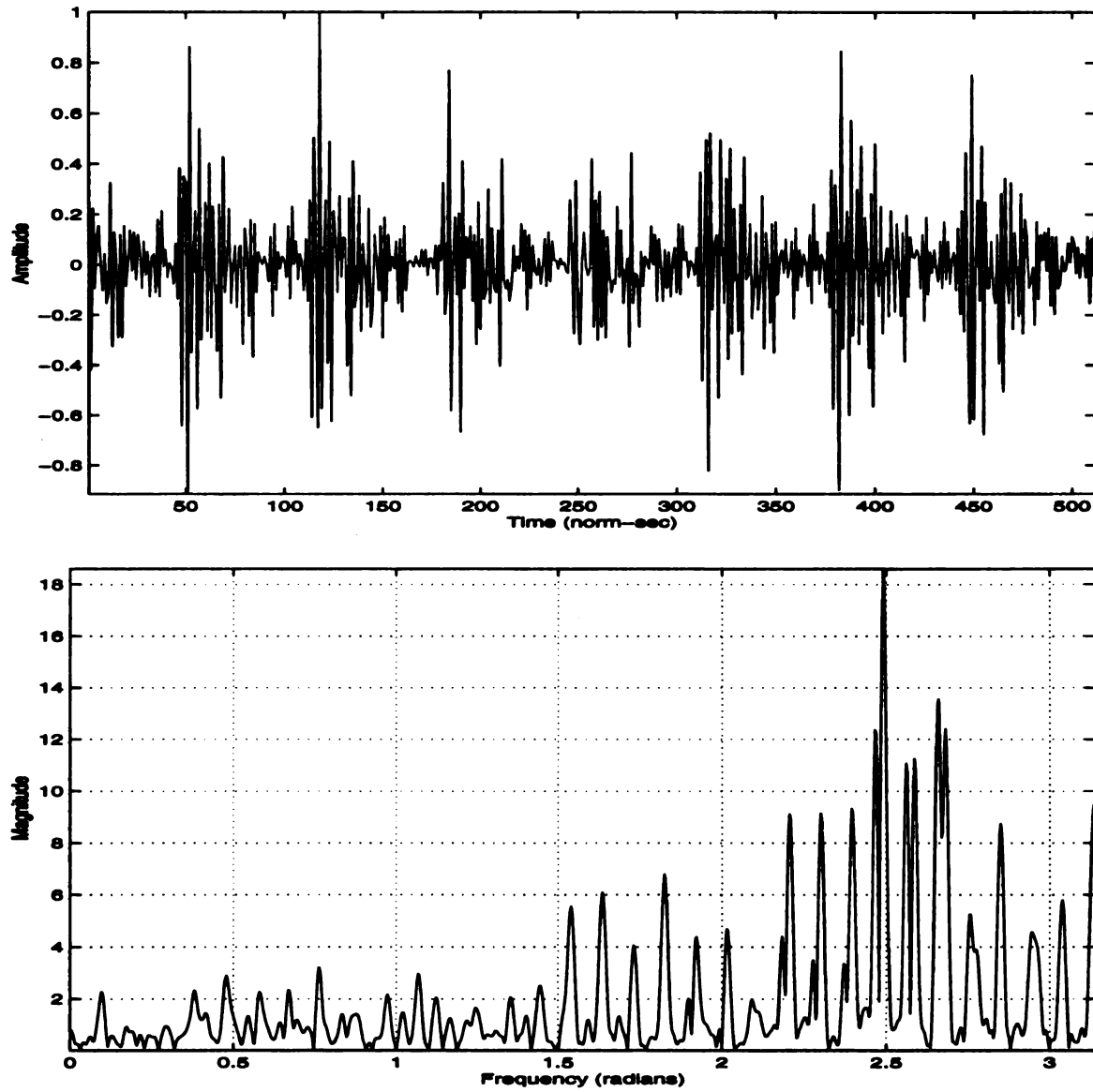
Another important factor in designing the detector is the signal-to-noise ratio (SNR) at which it is expected to operate. Since we do not expect to outperform the human auditory system, we target SNR's higher than -2 dB.



(a) A stationary signal segment from a telephone signal and its frequency spectrum.

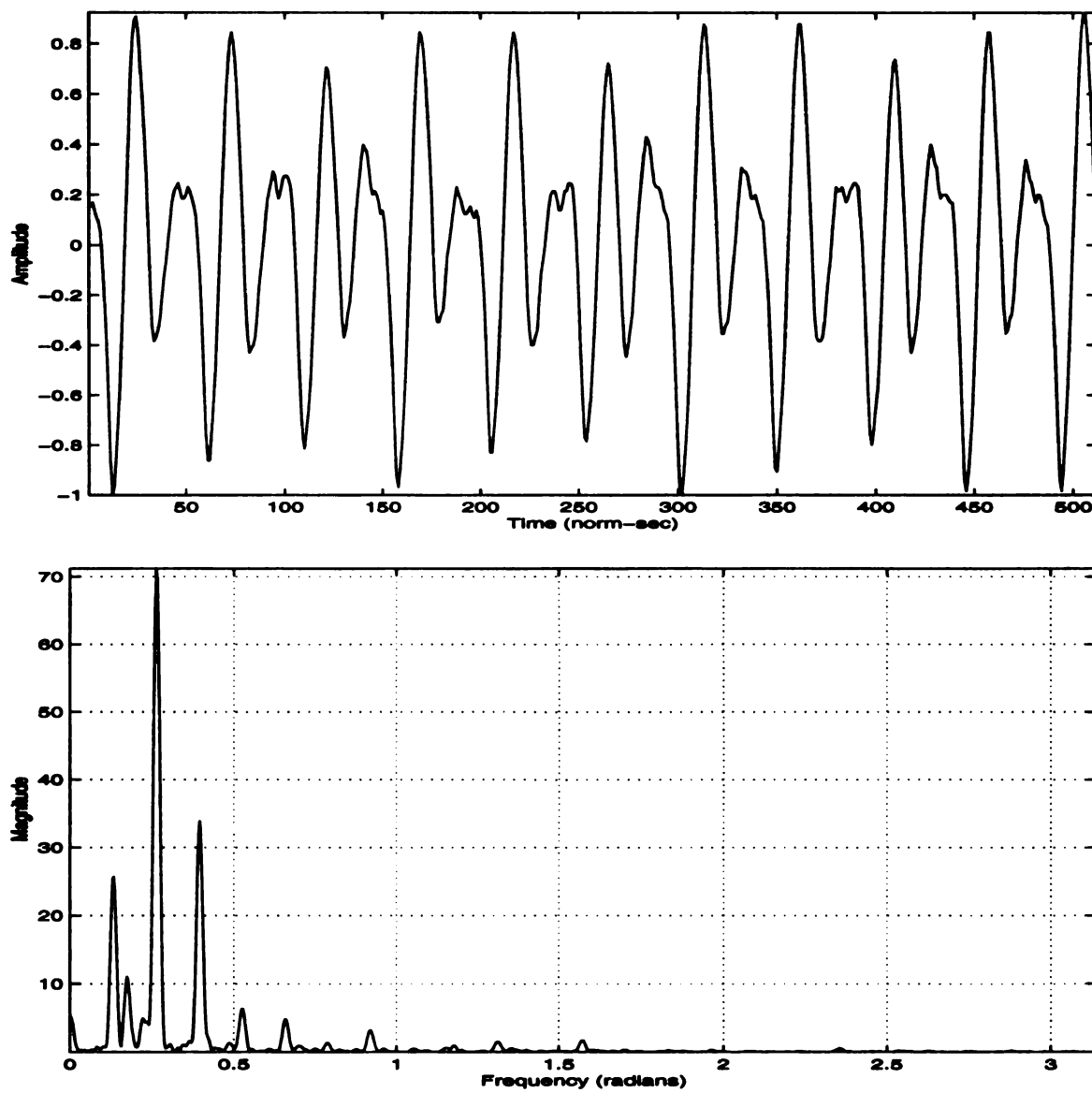
Figure 2.5. Several example stationary signal segments and their frequency spectra.

Figure 2.5. (cont'd).

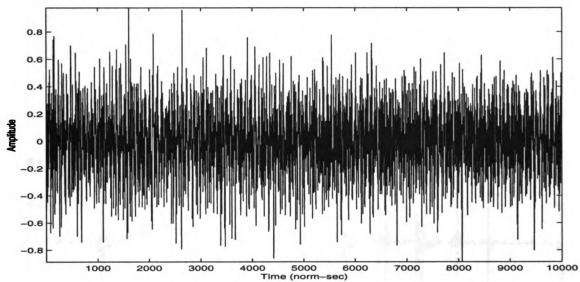


(b) A stationary signal segment from a buzzer signal and its frequency spectrum.

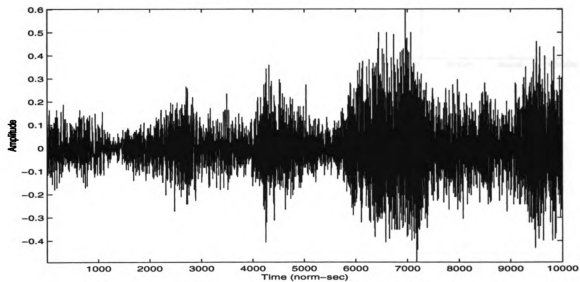
Figure 2.5. (cont'd).



(c) A stationary signal segment from a tornado siren signal and its frequency spectrum.



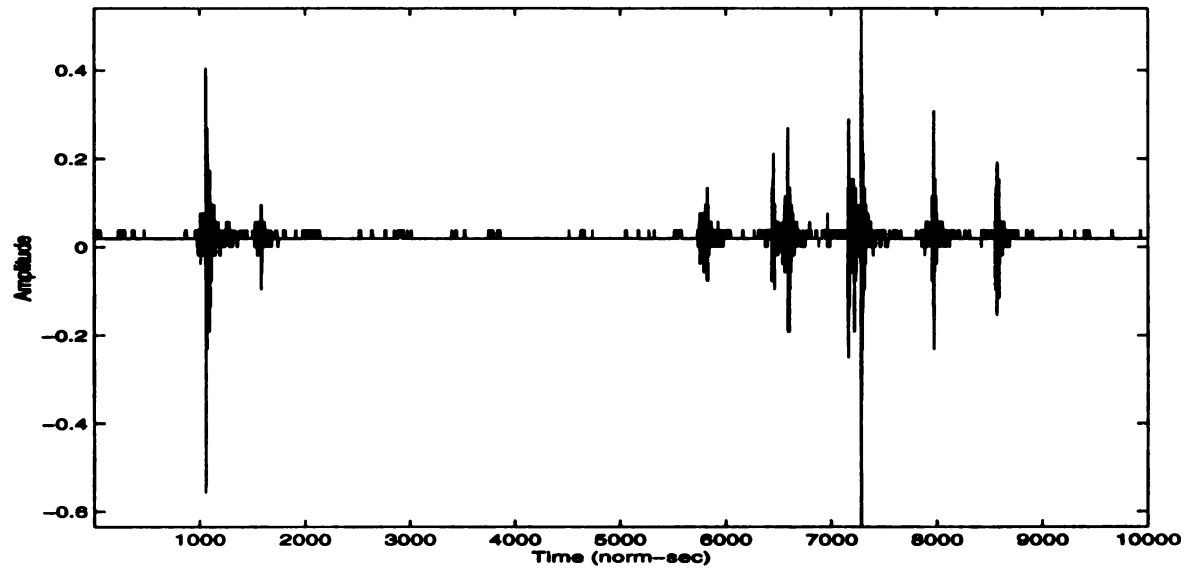
(a) Highway background noise.



(b) People talking.

Figure 2.6. Several example noise waveforms.

Figure 2.6. (cont'd).



(c) Computer keyboard.

2.3 Sampling Stage

Now let us turn our attention to the sampling stage of the adaptive alerting device. This stage consists of four blocks: Microphone, amplifier, sample-and-hold circuit, and analog-to-digital (A/D) converter (see Figure 2.7).

During the amplification stage, the weak signal from the microphone (typically 100-200 mV) is brought to a level suitable for processing by the A/D converter (typically 5-10 V). The desired amplification factor depends on the environment in which the device operates. In a noisy environment (*e.g.*, background highway noise) the amplification should be less than in a quiet environment (*e.g.*, a library). Therefore we use an automatic gain controller (AGC) to adjust the gain of the amplifier. The AGC ensures that the microphone signal is amplified maximally without causing saturation. Thus the resulting sampled signal has approximately constant amplitude. This complicates the detection procedure in the following ways:

- The instantaneous power of some Type I alerting signals (*e.g.* fire alarms, smoke detectors and telephones), when considered as a function of time, is nothing but a rectangular pulse train. The detection of such signals is a well-studied problem in digital communications [3]. However, because of the effects of AGC, the instantaneous power of the signal at the input of the detector stage will be approximately constant; therefore we cannot use the instantaneous power of the received signal as a detection feature.
- The instantaneous signal and noise levels cannot be estimated. Therefore the parameters of the detection algorithm must be independent of instantaneous signal and noise levels. Unfortunately, most standard statistical signal processing and pattern recognition algorithms require *a priori* knowledge of signal and noise powers (see, *e.g.*, [14, 16]). This observation suggests using techniques from robust [10] and non-parametric detection theory [11]; however these schemes are

not widely applicable.

Thus we must develop a detection algorithm which is robust to the scaling introduced by the AGC; the parameters of the detector should not depend on signal and noise powers. We shall model this effect by assuming that the instantaneous power of the signal at the detector input is constant. Strictly speaking, this model is inaccurate. The fact that the input signal has almost constant amplitude does not imply that it has constant power. However, this assumption will still result in an algorithm that is indifferent to the effects of the AGC stage and thus will serve the present purpose.

2.4 Choosing a Detection Strategy

As we mentioned in Chapter 1, we shall pursue a detection-theoretic approach to the development of the alerting signal detector. Of course, this is not the only way of solving this problem; techniques from other fields such as speech recognition and neural networks could also be applied. In this section, we justify our decision to use a detection-theoretic scheme, and discuss the reasons for ruling out other approaches.

As described in Section 2.2.1, alerting signals consist of long stationary regions connected by brief transition regions. Although the transitions play an important role in human perception of auditory signals [1], they are very difficult to model due to their nonstationary nature. On the other hand, for our purposes, the stationary regions contain sufficient information for detection of alerting signals. This observation enables us to reduce the detection of alerting signals to the detection of stationary signal segments.

To form an analogy with communications theory, the stationary signal segments composing the alerting signals to be detected form a “symbol alphabet” [19]. Thus we can use a suitable modification of standard detector structures used in communications receivers for detection of stationary signal segments.

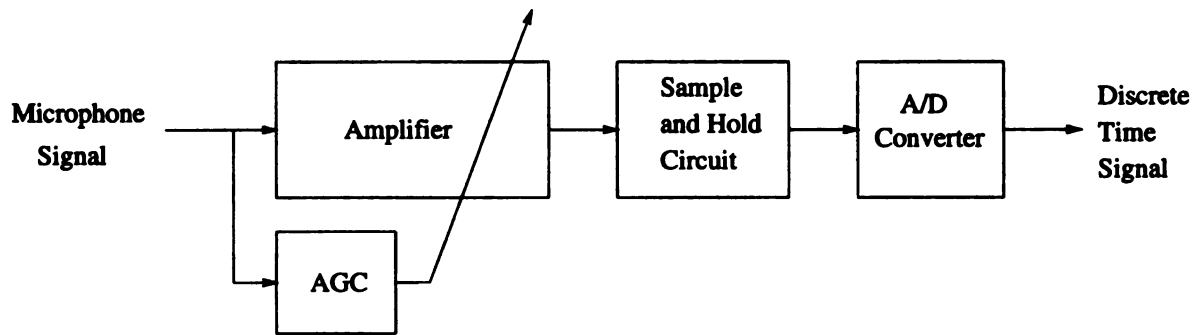


Figure 2.7. The sampling stage.

In addition to this detection-theoretic scheme, we have investigated the feasibility of techniques based on spectral analysis, dynamic time warping (DTW), hidden Markov models (HMM), and artificial neural networks (ANN). The shortcomings and disadvantages of these approaches are discussed in the following paragraphs. However, it should be noted that the primary advantage of pursuing a detection theoretic approach is that — unlike the HMM, DTW or ANN-based approaches discussed below — it promises to result in a computationally efficient solution.

2.4.1 Spectral Thresholding

A very simple detection algorithm could be based on the spectral properties of alerting signal segments. The fact that the energy of stationary signal segments is concentrated in narrow spectral regions, suggests using a detector consisting of a bank of bandpass filters and a comparator. However, under noisy conditions, the spectral peaks are not as dominant as under noiseless conditions; therefore combined with the inability to estimate SNR, reliable threshold selection is not possible. Additionally, signal processing techniques based on spectral thresholding have high false-alarm and miss rates in general (for example, formant estimation based on spectral thresholding [4]).

2.4.2 Dynamic Time Warping and Hidden Markov Models

Another idea is to use techniques from speech recognition, such as *dynamic time warping* (DTW) [4, 17] and *hidden Markov modeling* (HMM) [4, 15]. They are both tailored to handle the variability among different utterances of the same word; therefore they can easily cope with the differences among realizations of Type II and III alerting signals.

However, with decreasing SNR, the performance of these algorithms drops

drastically. This lack of robustness is caused by the distance metrics used in DTW and HMM algorithms. Regardless of the choice of acoustic features, all distance measures are very sensitive to noise. Thus we rule out the possibility of using algorithms from the speech recognition field in the detector stage.

2.4.3 Artificial Neural Networks

Yet another detector structure could be designed using neural networks in combination with other suitable techniques [5]. Neural networks naturally account for commonly encountered properties of real-life data, including nonstationarity and non-Gaussianity [7]. However, their implementation can be very complex and computationally demanding; therefore, in practice, their use is reserved to situations where signal modeling is difficult or impossible.

2.5 A Functional Block Description of Detector Architecture

Before we proceed to the development of the detector architecture, we introduce some notation. We henceforth use the symbol $\mathcal{A}$ to denote an alerting signal and the symbol $\mathcal{S}$ to denote stationary signal segments. Using a set-like notation, we shall describe an alerting signal consisting of n stationary signal segments as $\mathcal{A} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$.

Suppose, we want to develop an algorithm for detection of the signals shown in Figures 2.1–2.4. Using the above introduced conventions, the fire alarm signal can be expressed as $\mathcal{A}_f = \{\mathcal{S}_1\}$, the siren signal as $\mathcal{A}_s = \{\mathcal{S}_2, \mathcal{S}_3\}$, the doorbell signal as $\mathcal{A}_d = \{\mathcal{S}_4, \mathcal{S}_5\}$, and the crying baby signal as $\mathcal{A}_b = \{\mathcal{S}_6, \mathcal{S}_7, \mathcal{S}_8\}$. Thus the “symbol alphabet” is $\mathcal{U} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_8\}$.

This notation leads us to a detector consisting of the following three stages (Figure 2.8):

- Filter Bank Stage
- Decision Block
- Postprocessor

The first two stages represent extensions of similar stages in a standard communications receiver for M-ary symbol sets. They act as a detector for stationary signal segments. In a sense, these two stages take the received signal and convert it to a higher-level representation. The postprocessor takes the output of the decision block, and using the duration and temporal order of stationary signal segments, makes the following decisions:

- Whether an alerting signal is present at the detector input or not.
- If an alerting signal is present, which one.

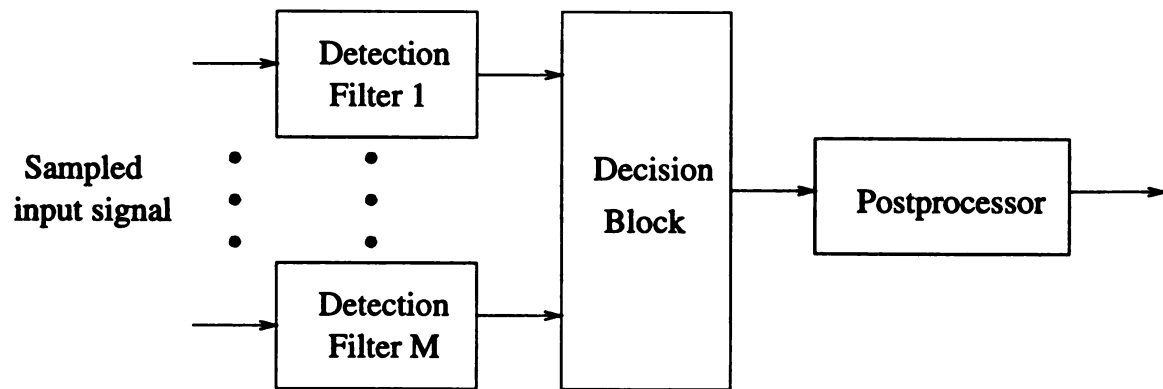


Figure 2.8. The basic detector architecture.

CHAPTER 3

Detector Design

3.1 Design Methodology

The non-Gaussian and nonstationary nature of the noise environment combined with the inability to estimate SNR prevents the pursuit of a strictly analytical procedure for derivation of the detector stages. For the same reasons, we cannot impose standard optimality criteria from statistical signal processing [14] to the solution of this problem. These observations lead to the following design methodology which we shall follow in the remainder of the thesis:

- Using *ad hoc* techniques as well as suitable modifications of standard schemes, we propose several potential solutions for each stage of the detector. Each method will only approximately meet the ideal specifications stated in Section 2.5 and will have different performance, cost, and implementation advantages.
- Because of the lack of analytical models, we cannot meaningfully quantify the performance of different candidates. Therefore, based on simulation studies, we shall weigh the tradeoffs of different approaches and choose the most suitable configuration of the designed alternative blocks for use in the final device.

In this chapter we shall implement the first step of this design strategy, and develop the detector stages. Their performance evaluation will be the subject of the next chapter.

3.2 Detection Filters

Let $\mathcal{S}$ be a signal segment to be detected. In this section, we develop several candidate detection filters for $\mathcal{S}$. The treatment is based on the following assumptions:

- In practice, the realizations of $\mathcal{S}$ are of finite duration. However, we assume that an infinite duration realization $s(n)$, $-\infty < n < \infty$ of $\mathcal{S}$ exists and that a noisy version of this signal is input to the detector. For example, suppose $\mathcal{S}$ is the stationary signal segment corresponding to the fire alarm signal shown in Figure 2.1b. In this case, we can form an “eternal” realization of $\mathcal{S}$ by infinitely concatenating the signal segment with itself. This assumption enables us to design detector filters without being concerned with timing problems. Synchronization issues will be discussed in Section 3.3.
- Another important problem in the design of detector filters is developing appropriate signal models. A careful analysis of the alerting signal waveforms shown in Figures 2.1–2.4 reveals that some stationary signal segments are best modeled as deterministic whereas others as stochastic. For example, let us compare the stationary signal segments shown in Figures 2.1b (fire alarm) and 2.4c (crying baby). Every realization of the first will be identical, *i.e.*, it is deterministic. On the other hand, the second one will be slightly different every time the baby cries. Therefore it should be modeled as a stochastic process.

In the following discussion, we assume that every stationary signal segment can be modeled in both ways. Although incorrect, this will enable us to exploit

both techniques developed for detection of deterministic and stochastic signals.

- Suppose we observe a realization of $\mathcal{S}$ in background noise. Let us denote the SNR in decibels at the input and output of the detection filter as SNR_{in} and SNR_{out} , respectively. The design objective will be to maximize the difference $|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}|$. In other words, the detection filter must accomplish one of the following two tasks:
 - Amplify the signal component while suppressing the noise component;
 - Suppress the signal component while amplifying the noise component.

Because of the absence of analytic signal and noise models, we shall try to accomplish these goals in an heuristic manner. We call the difference $|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}|$ the *SNR improvement factor* * and this design criterion the *SNR improvement criterion*.

Next, let us describe how the detection filters designed according to this criterion will perform. Let $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M$ denote the M stationary signal segments to be detected and let F_i be the detection filter corresponding to $\mathcal{S}_i$ where $1 \leq i \leq M$. Without loss of generality, we can assume that a noisy realization of $\mathcal{S}_1$ is at the input of the detector. Assuming that the detection filters follow the first version of the design criterion given in Section 3.2, F_1 should “resonate” with the input signal and suppress the noise component. On the other hand, from the perspective of all other filters, the input signal should appear to be noise only. Therefore they should suppress the input signal as well as the noise as much as possible. Thus the average power at the output of F_1 is much higher than the output power of the other filters. The decision block should exploit this fact to decide in favor of $\mathcal{S}_1$.

* This notion is analogous to the concept of *demodulation gain* in communications theory.

- For high SNR and large sample size detection problems, improper assumptions regarding the noise environment do not significantly affect the detector performance [9]. It is easy to see that the development of the detection filters falls into this category. Therefore, as we remarked in Section 2.2.2 above, whenever the need to model the noise environment arises, we shall simply assume that it is independent, identically distributed Gaussian noise.

Now we are ready to proceed with the development of detection filters.

3.2.1 Matched Filter

Let $\mathcal{S}$ be a stationary signal segment, and let $s(k)$, $-\infty < k < \infty$ be an “eternal” realization of $\mathcal{S}$, obtained as described above. Assuming that $\mathcal{S}$ is deterministic, $s(k)$ can be regarded as the endless repetition of a representative region consisting of samples $r(0), r(1), \dots, r(N-1)$, i.e. it can be expressed as $s(n) = r(n \bmod N)$. For example, Figures 2.5a, b and c show the representative regions of a buzzer, telephone, and tornado siren signal, respectively.

Now, suppose that we observe a noisy realization $x(k) = s(k) + n(k)$, $-\infty < k < \infty$ of $\mathcal{S}$ where $n(k)$ denotes the noise sequence. An intuitive way of achieving a high SNR improvement factor is based on the cross-correlation sequence between the received signal and representative region. Let

$$y(k) = \sum_{l=0}^{N-1} r(l)x(k+l), \quad -\infty < k < \infty \quad (3.1)$$

be the cross-correlation sequence between $x(k)$, $-\infty < k < \infty$ and $r(0), r(1), \dots, r(N-1)$. $y(k)$ can be expressed as the sum of the sequences $y_s(k) = \sum_{l=0}^{N-1} r(l)s(k+l)$ and $y_n(k) = \sum_{l=0}^{N-1} r(l)n(k+l)$. We expect $y_n(k)$, the cross-correlation between noise signal and representative region to assume values close to zero. On the other hand, $y_s(k)$, the cross-correlation between the signal to be detected and representative

region, should take higher values. Thus (3.1) suppresses the noise component of $x(k)$ while amplifying the signal component and meets the SNR improvement criterion.

It is easy to see that (3.1) can be implemented using a *matched filter* $F(z)$ with impulse response $f(m) = r(N-1-m)$, $k = 0, \dots, N-1$ [14]. However, at a sampling rate of 8 kHz, the representative region is typically 300-500 samples long. Obviously, a filter with such a long impulse response is not suitable for real-time applications. Therefore we shall confine ourselves to approximate implementations of $F(z)$ which are computationally tractable. In particular, since we are only concerned with the average power at the output of $F(z)$, realizing its magnitude response is sufficient. To this end, any of the standard filter design techniques available in the literature can be used [13].

3.2.2 Eigenfilter

Let us model $\mathcal{S}$ as an ergodic wide-sense stationary random process and let $\mathbf{R}$ be its autocorrelation matrix. Under white noise, the impulse response of the optimum linear detection filter maximizing the SNR improvement factor is given by $f(m) = v_{\max}(m)$, $m = 0, \dots, N-1$ where $[v_{\max}(0) \dots v_{\max}(N-1)]^T$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{R}$ and N is the size of $\mathbf{R}$ [6]. This filter is called as an *eigenfilter* and can be regarded as the stochastic counterpart of the matched filter introduced above.

3.2.3 Prediction Error Filter

As we did in the previous section, assume that $\mathcal{S}$ is a stochastic signal segment. Let $\underline{s}(k)$, $-\infty < k < \infty$ be the random process corresponding to $\mathcal{S}$ and assume that it obeys an autoregressive (AR) model [2] of a known order. Let $F(z)$ be the minimum-mean-square-error one-step linear predictor for $\underline{s}(k)$ [2]. Now, let us show how the

one-step linear prediction error filter for $\underline{s}(k)$, $F'(z) = 1 - z^{-1}F(z)$, can be employed as a detection filter.

First, suppose that $\underline{s}(k)$, $-\infty < k < \infty$ is at the input of $F(z)$. The predictions made by $F(z)$ will be close to the actual values. Therefore, when $\underline{s}(k)$ is at the input of $F'(z)$, the average output power will be lower than the average power of $\underline{s}(k)$. Next, assume that, a noise process $\underline{n}(k)$, $-\infty < k < \infty$ is at the input of $F(z)$. $F(z)$ will try to predict future values of $\underline{n}(k)$ as if they belonged to a realization of $\mathcal{S}$. Obviously, except by coincidence, the predictions will be erroneous. Thus, when $\underline{n}(k)$ is at the input of $F'(z)$, the average output power will be much higher than in the previous case. Hence when a noisy realization of $\mathcal{S}$ is at the input of $F'(z)$, the SNR at the output will be much higher than the SNR at the input; thus $F'(z)$ qualifies as a detection filter.

3.2.4 Autocorrelation Detector

Another detection filter can be based on the autocorrelation functions of the signal segment to be detected and received signal. Let $x(k) = s(k) + n(k)$, $-\infty < k < \infty$ be the received signal where $s(k)$ and $n(k)$ are defined as above. Assume that $s(k)$ is a realization of an ergodic wide-sense stationary random process whose autocorrelation function is $\rho_s(h)$, $-\infty < h < \infty$ and $s(k)$ and $n(k)$ are uncorrelated sequences.

Let $\rho_{x,m}(h)$ be the autocorrelation function of a frame of $x(n)$ starting at the time $mN - L$ and ending at time mN , $-\infty < m < \infty$ where N and L satisfy $0 < N \leq L$. In a similar fashion, we can define the functions $\rho_{s,m}(h)$ and $\rho_{n,m}(h)$ for $s(k)$ and $n(k)$. Because $s(k)$ and $n(k)$ are uncorrelated, $\rho_{x,m}(h)$ can be written as $\rho_{x,m}(h) = \rho_{s,m}(h) + \rho_{n,m}(h)$. Noting that $s(k)$ is the realization of a wide-sense stationary random process, we have $\rho_{s,m}(h) \approx \rho_s(h)$. Thus we obtain

$$\rho_{x,m}(h) \approx \rho_s(h) + \rho_{n,m}(h), \quad -\infty < h < \infty. \quad (3.2)$$

Let us define the autocorrelation detector as a filter whose output is given by

$$y(m) = \sum_{h=-P}^P \rho_{x,m}(h) \rho_s(h) \quad (3.3)$$

when $x(n)$ is at its input. We shall call $P > 0$ the order of the autocorrelation detector. Note that the output rate of the autocorrelation detector is N times lower than its input rate.

We see that $y(m)$ is approximately the correlation among the sequences $\rho_{x,m}(h)$ and $\rho_s(h)$. Using (3.2), $y(m)$ can be expressed as

$$y(m) = y_s(m) + y_n(m) \approx \sum_{h=-P}^{h=P} \rho_s^2(h) + \sum_{h=-P}^{h=P} \rho_s(h) \rho_{n,m}(h). \quad (3.4)$$

According to the analysis done in Section 2.2.1, the energy of $s(k)$ is concentrated in narrow spectral peaks. Therefore, $\rho_s(h)$, $h \geq 0$ must be a slowly decaying periodic sequence [2]. No similar statement can be made about $\rho_{n,m}(h)$; however, unless $\rho_s(h)$ and $\rho_{n,m}(h)$ are periodic with the same frequency — which is only possible in pathologic situations — the second term in (3.4) will be close to zero or negative. On the other hand, the first term of (3.4) is always positive. Thus, except under relatively low SNR's we have $y_s(m) \gg y_n(m)$ for $-\infty < m < \infty$. Hence (3.3) meets the SNR improvement criterion and can be used as a detection filter.

Equation (3.3) has an interesting explanation in frequency domain which clarifies the points made above [14]. Let $\phi_{x,m}(\omega)$, $\phi_s(\omega)$ and $\phi_{n,m}(\omega)$, $-\pi \leq \omega \leq \pi$ be the power spectral densities of $x(k)$, $s(k)$ and $n(k)$, $mN - L \leq k < mN$. For $P = \infty$, according Wiener-Kitchine Theorem [12], (3.3) can be written as

$$y(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{x,m}(\omega) \phi_s(\omega) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_s^2(\omega) d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{n,m}(\omega) \phi_s(\omega) d\omega. \quad (3.5)$$

Usually, the energy of $s(k)$ and $n(k)$ are concentrated in non-overlapping regions of

the frequency spectrum. Therefore the second term of (3.5) will be close to zero, and its first term, which is always positive, will dominate $y(m)$, except under low SNR's.

It can be shown that the autocorrelation detector is equivalent to the *locally optimum detector* for stochastic signals with unknown amplitude [14]

3.3 Decision Block

In this section we design the decision block. The decision block must accomplish the following two tasks: Determine the decision instants (*i.e.* the starting and ending points of stationary signal segments) and compare the detection filter outputs to decide which stationary signal segment is present at the detector output.

3.3.1 Synchronization

In developing a synchronization technique, there are two main issues we have to address: The first is the determination of optimal decision instants. The second concerns the removal of the infinite length stationary signal segment assumption made in designing the detection filters: How many samples should be used in making a decision?

In this section, we answer these two questions. First, to get an insight to these two problems, we examine how they are solved in digital communications and explain why these approaches cannot be applied to the decision block of the alerting signal detector. This comparison will lead to a surprisingly simple synchronization scheme.

Let us focus on the the determination of decision instants. In communications receivers, this is usually done using a *clock extraction circuit* which exploits the periodicity of the received signal to obtain the starting and ending instants of the symbols to be detected [3]. However, the structure of alerting signals is quite different than the symbol strings transmitted in a communications channel. Therefore, we cannot

apply synchronization schemes from classical detection theory; we must develop a technique which is compatible with the properties of the given alerting signals.

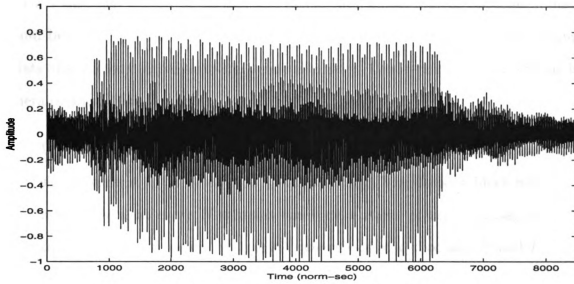
In communications applications the second issue never poses a problem, because the symbols to be detected are of fixed and known length. On the other hand, each stationary signal segment (or “symbol”) we have to detect is of different length; furthermore the duration of some stationary signal segments varies with every realization. We must find a means of overcoming this difficulty.

Having seen why standard synchronization schemes cannot be applied to the decision block of the alerting signal detector, now let us concentrate on a major difference between the stationary signal segments we have to detect and symbols used in digital communications: Duration.

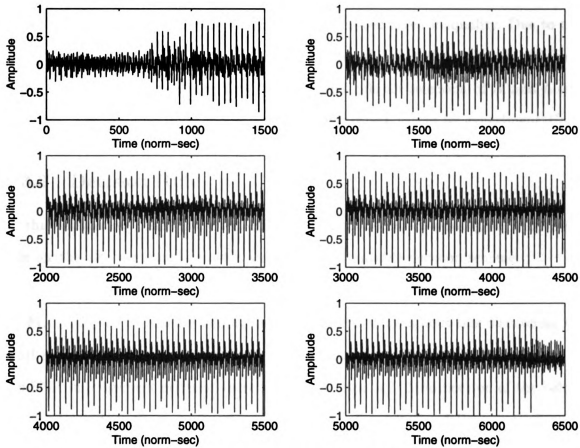
In communications receivers, the importance of synchronization errors on performance is coupled with signaling speed. The faster the bit rate, the shorter the symbols, and therefore the more precise the clock extraction circuit must be. On the other hand, it can be shown that the effects of timing errors on performance diminishes as the signaling rate decreases [3]. Since the duration of a stationary signal segment is at least quarter of a second long, its “signaling rate” is very low. In fact, it is so low that assigning the decision instants arbitrarily would presumably not affect the detection rate.

To illustrate how we can exploit this conclusion to the development of a synchronization scheme, let us consider the detection of the Type I alerting signal consisting of a single stationary signal segment S , shown in Figure 3.1a. As we see, S is ca. 7000 samples long. For comparison, in communications systems, the symbols to be detected are approximately 5-20 samples long.

Now, suppose that we make a decision every 1000 samples, based on the last 1500 samples. In other words, the decision rule processes received signal using overlapping frames of 1500 samples. This point is illustrated in Figure 3.1b. As we see the



(a) A typical stationary signal segment.



(b) Several consecutive overlapping frames.

Figure 3.1. Framewise processing of a stationary signal segment.

decision logic may miss $\mathcal{S}$ in the first and last frames; however all of the other frames would be correctly detected. Thus the exact determination of starting and ending points of stationary signal segments is not necessary. This synchronization scheme is generalized as follows:

Let $x(n)$, $-\infty < n < \infty$ be the received signal. Suppose that the shortest signal segment to be detected is Q samples long. The decision block will make a decision every N samples based on a frame of $x(n)$ starting at time $kN - L$ and ending at time kN where $-\infty < k < \infty$ and N and L are integers satisfying $N \leq L < Q$.

In the above example Q is 7000 samples, N is 1000 samples and L is 1500 samples. Both N and L were chosen arbitrarily within the above given inequality. Due to the heuristic nature of this formulation, it is difficult to tell how different values of N and L will affect the detection rate; therefore no strict rule for selection of N and L can be given. The designer must choose the most convenient values by experimentation.

3.3.2 Decision Rule

In this section, we investigate the difficulties encountered in developing a decision rule for the alerting signal detector and propose a suitable decision scheme.

As before, let $\mathcal{S}_1, \dots, \mathcal{S}_M$ be the stationary signal segments to be detected, and let $F_1, \dots, F_M$ be the corresponding detection filters. Throughout this discussion, we shall assume that the detection filters are designed according to the first version of the SNR improvement criterion introduced in Section 3.2. Modification of the below obtained results to the second version of this optimality criterion follows immediately.

According to the synchronization scheme developed above, the decision block must process the signals at the detector filter outputs framewise. Suppose that $x(0), \dots, x(L - 1)$, an input signal frame, is at the input of the detection filters.

Let us denote the corresponding frames at the output of F_i as $y_i(0), \dots, y_i(L-1)$,[†] $1 \leq i \leq M$. We must compare these frames to make the following decisions:

- Determine whether a stationary signal segment, a transition region, a region of silence, or just noise is present at the detector input.
- If a stationary signal segment is present at the detector input, decide which one.

A simple decision rule can be based on detector filter output frame energies. Suppose that the frame $x(0), x(1), \dots, x(L-1)$ is at the input of the filters and let $E_x = \sum_{n=0}^{L-1} x^2(n)$ be its energy. Let $E_i = \sum_{n=0}^{L-1} y_i^2(n)$ the output energy of F_i , $1 \leq i \leq M$. Further suppose that the detection filters satisfy the following criteria:

- (i) When a noisy realization of S_i is at the input of the detection filters, we have $E_i \gg E_j$ for $j \neq i$, $1 \leq i, j \leq M$.
- (ii) When only noise is at the input of detection filters, we have $E_j \ll E_x$ for $1 \leq j \leq M$.

At first sight, it seems that filters with good SNR improvement factors would meet the above criteria. However, they will not unless their gains are chosen appropriately. Because of the heuristic nature of the detection filter design procedure, no rule for gain selection can be given. The designer must adjust the filter gains based on simulation studies and characteristics of individual filters. Criteria (i) and (ii) immediately suggest the following decision rule which is an extension of the “choose maximum” scheme from classical detection theory [3]:

$$\text{Let } E_i = \max\{E_1, \dots, E_M\} \text{ and } E_j = \max\{E_1, \dots, E_{i-1}, E_{i+1}, \dots, E_M\}.$$

Let $\alpha > 1$ be an experimentally determined parameter which we shall

[†] For the sake of simplicity, we overlook the fact that when a frame of length L is at the input of a filter, the output signal may be of different length. Additionally, the output of one of the detector structures proposed in Section 3.2 is a single number per input frame. The modification of the decision logic explained in this chapter to this detector is straightforward.

call as *sensitivity parameter*. If $\frac{E_i}{E_j} > \alpha$, decide in favor of S_i ; otherwise conclude that no stationary signal segment is present at detector input.

The basis for this technique is easy to understand: At low SNR's when S_i is present at detector input, E_i must be substantially higher than the output energies of other filters. On the other hand, if only noise is present at the detector input, all filter output energies must be close to each other. The ratio $\frac{E_i}{E_j}$ is a measure of this closeness. If α is small, the detector will have a low miss rate but a high false alarm rate. If α is large, the opposite will occur. Thus the term sensitivity parameter for α . At high SNR's, typical of the environments where the adaptive alerting device will to operate, a wide range of α values will result in a low miss and false alarm rate. This point will be verified by simulation in Chapter 4.

3.4 Postprocessor

The duty of the postprocessor is to complete the job of the decision block and translate the detected stationary signal segments into alerting signals. The degree of sophistication required in the postprocessor depends on the reliability of the decision block. To compensate for a high error rate by the decision block, the postprocessor can exploit the length and temporal order of the detected stationary signal segments. For example, algorithms from syntactic pattern recognition could be used to this end.

On the other hand, if our confidence in the decision block is high, the translation procedure can be accomplished relatively easily. In this case, the following algorithm, which is a generalized majority selection rule, can be used:

At every decision instant, consider the last P signal segments detected.

If at least β percent of them belong to a given alerting signal, declare its presence at the detector input. Otherwise conclude no alerting signal present.

As all parameters we introduced in this chapter, P and β must be experimentally determined. This postprocessor algorithm is strikingly simple; therefore we should first evaluate its performance in combination with the techniques developed in the previous pages; and only if we do not obtain satisfactory results we should attempt to develop a more complicated but robust postprocessor block.

CHAPTER 4

Performance Evaluation

4.1 Introduction

In this section, we report the results of simulation studies that evaluate the strengths and weaknesses of the various alternative approaches considered, and draw conclusions about the appropriate detector structure to use in the ultimate device.

There are two levels at which individual detection filter design strategies can be benchmarked: At the lower level, the success of a particular filter is determined by its SNR improvement factor. At the higher level, the performance is specified by the detection rate obtained.

In Section 4.2, we report results of simulations used to obtain the SNR improvement factor achieved by the filter design techniques presented in Section 3.2. In Section 4.3, we quantify the performance of detector strategies based on the foregoing developments.

4.2 Benchmarking with the SNR Improvement Factor

4.2.1 Formal Developments

Let F be a detection filter designed for a stationary signal segment $\mathcal{S}$. Let $s(k)$, $-\infty < k < \infty$ be a realization of the eternal extension of $\mathcal{S}$, and let $n(k)$, $-\infty < k < \infty$ be a noise sequence. Without loss of generality, we assume that the signal and noise are both unit power sequences. Additionally, we assume that the two processes are mutually uncorrelated sequences.

Let $y_s(k)$, $-\infty < k < \infty$ and $y_n(k)$, $-\infty < k < \infty$ be the response of F to the signal and noise, respectively. Assume that the signal

$$x(k) = \sqrt{P_s} s(k) + \sqrt{P_n} n(k), \quad -\infty < k < \infty \quad (4.1)$$

is at the input of F where P_s and P_n denote the power of signal and noise components.

Thus we have

$$\text{SNR}_{\text{in}} = 10 \log \frac{P_s}{P_n}. \quad (4.2)$$

Next, let us derive expressions for SNR_{out} and $|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}|$. The forms of these expressions depend on the characteristics of F . The filters derived in Section 3.2 can be grouped in two categories: Linear filters (matched filter, prediction error filter, and eigenfilter) and the quadratic autocorrelation detector. We treat these cases separately:

Linear Detection Filter. The output of F can be expressed as

$$y(k) = \sqrt{P_s} y_s(k) + \sqrt{P_n} y_n(k), \quad -\infty < k < \infty. \quad (4.3)$$

Thus

$$\text{SNR}_{\text{out}} = 10 \log \frac{P_s \sum_{k=-\infty}^{\infty} y_s^2(k)}{P_n \sum_{k=-\infty}^{\infty} y_n^2(k)} \quad (4.4)$$

and

$$|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}| = \left| 10 \log \frac{\sum_{k=-\infty}^{\infty} y_s^2(k)}{\sum_{k=-\infty}^{\infty} y_n^2(k)} \right|. \quad (4.5)$$

Quadratic Detection Filter. The response of a quadratic filter to input sequence $x(k)$, $-\infty < k < \infty$ can be expressed as

$$y(k) = P_s y_s(k) + P_n y_n(k) + \sqrt{P_s P_n} \varphi(s(k), n(k)), \quad -\infty < k < \infty \quad (4.6)$$

where $\varphi(\cdot)$ is a linear function determined by filter parameters. Thus SNR_{out} is given by

$$\text{SNR}_{\text{out}} = 10 \log \frac{P_s^2 \sum_{k=-\infty}^{\infty} y_s^2(k)}{P_n^2 \sum_{k=-\infty}^{\infty} y_n^2(k) + P_s P_n \sum_{k=-\infty}^{\infty} \varphi(s(k), n(k))}. \quad (4.7)$$

Since $s(\cdot)$ and $n(\cdot)$ are uncorrelated sequences, we have

$$\sum_{k=-\infty}^{\infty} \varphi(s(k), n(k)) = 0 \quad \text{for all } k. \quad (4.8)$$

Hence the SNR improvement factor of a quadratic detection filter is

$$|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}| = \left| 10 \log \frac{\sum_{k=-\infty}^{\infty} y_s^2(k)}{\sum_{k=-\infty}^{\infty} y_n^2(k)} + 10 \log \frac{P_s}{P_n} \right|. \quad (4.9)$$

Unlike the SNR improvement factor of a linear filter, Equation (4.9) is dependent on SNR_{in} .

Adjustments for Finite-Time Analysis. The results above assume infinite length signals and cannot be used in practice without proper modifications. The fact that the synchronization and decision schemes introduced in Section 3.3 process the input signal framewise suggests that evaluating the SNR improvement factor of

detection filters on finite length frames might be a better indication of their performance. This observation leads us to the following algorithm:

Suppose that we observe a stationary signal segment $\mathcal{S}$ in a particular noise environment. Let F be a detection filter designed for $\mathcal{S}$. Let $s_j(k)$, $0 < k < L - 1$ and $n_j(k)$, $0 < k < L - 1$, with $1 \leq j \leq Q$ be frames of a realization of the signal and noise, respectively, assumed, without loss of generality, to each be of unity power. Let $y_{s_j}(k)$, $0 < k < L - 1$ and $y_{n_j}(k)$, $0 < k < L - 1$ be the response of F to the signal and noise frames, respectively. Now fix j and assume that the frame

$$x_j(k) = s_j(k) + n_j(k), \quad 0 \leq k \leq L - 1 \quad (4.10)$$

is input to F . Thus $\text{SNR}_{\text{in},j} = 0$ dB. Modifying (4.5) and (4.9) appropriately gives

$$|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}|_j = \left| 10 \log \frac{\sum_{k=0}^{L-1} y_{s_j}^2(k)}{\sum_{k=0}^{L-1} y_{n_j}^2(k)} \right|. \quad (4.11)$$

Note that since $\text{SNR}_{\text{in},j} = 0$ dB, the expression for SNR improvement factor is same for both linear and quadratic F . For a quadratic detection filter, modification of (4.11) for other SNR_{in} follows immediately from (4.9).

The SNR improvement factor obtained for each value of j will be slightly different. We compensate for this effect by averaging

$$|\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}| = \frac{1}{Q} \sum_{j=1}^Q |\text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}|_j. \quad (4.12)$$

4.2.2 Simulation Results

The algorithm above was used to obtain the SNR improvement factors for the various detection filter designs presented in Section 3.2 for several stationary signal segments

and realistic noise environments. The performance results appear in Tables 4.2-4.5. In all simulations $Q=20$ and $L=1500$ samples. Implementation details are shown in Table 4.1 .

Comparing Tables 4.2-4.5, we conclude that the SNR improvement factors achieved by the autocorrelation detector are superior to all other design techniques. Additionally, we note that the performance of a given approach can show significant variation among different stationary signal segments and noise environments.

4.3 Benchmarking Using Detection Rates

4.3.1 Introduction

Several difficulties had to be overcome in developing a performance evaluation scheme. First, we were unable to assign *a priori* probabilities and decision costs to the stationary signal segments to be detected. Both quantities depend on the specific needs of the device user (pragmatic knowledge). For example, in a car the probability of encountering an ambulance siren is quite high, whereas the probability of encountering a doorbell signal is zero. The opposite is true in a home setting.

Similar situations occur in classical detection problems for communications. In these cases, the performance of the detectors are often evaluated *via* receiver operating characteristics (ROC) analysis [14]. However, ROC curves are most appropriate for problems involving binary hypotheses and are not particularly well-suited to the present application.

The second difficulty concerns the number and choice of the stationary signal segments to be used in simulation. Both of these factors depend on the user's needs and lifestyle.

The performance of the decision block was evaluated on a set of alerting signals

Table 4.1. Design details of the detection filters used in the simulations.

Matched Filter	Representative region of length 512 samples chosen via manual inspection of alerting signal frames. An all-pole filter of order 16 fitted to the representative region using Yule-Walker equations [2]
Eigenfilter	Autocorrelation matrix of alerting signals estimated using a Bartlett estimator [2]. Order of filter 16.
Prediction Error Filter	Autocorrelation matrix computed same as above. An FIR predictor of order 16 designed solving Yule-Walker equations.
Autocorrelation Detector	Autocorrelation computed same as above based on non-overlapping frames of length 1000 samples.

Table 4.2. SNR improvement factors in decibels for matched filter.

	White	Laugh	Highway	Party	Keyboard
Fire	10.796	12.675	12.845	11.476	10.237
Siren	18.716	21.669	19.977	19.839	21.186
Telephone	13.049	12.357	19.571	14.899	11.730
Buzzer	5.726	10.797	18.224	9.408	8.353

Table 4.3. SNR improvement factors in decibels for prediction error filter.

	White	Laugh	Highway	Party	Keyboard
Fire	12.850	5.366	2.286	7.437	7.074
Siren	26.545	19.000	9.517	21.562	23.670
Telephone	11.978	10.670	9.657	10.908	11.515
Buzzer	7.065	7.271	13.932	7.495	7.689

Table 4.4. SNR improvement factors in decibels for eigenfilter.

	White	Laugh	Highway	Party	Keyboard
Fire	7.248	10.443	15.522	8.906	6.746
Siren	9.437	10.662	5.773	9.785	11.582
Telephone	6.950	5.387	19.000	7.663	5.484
Buzzer	6.693	15.007	19.652	12.764	11.052

Table 4.5. SNR improvement factors in decibels for autocorrelation detector.

	White	Laugh	Highway	Party	Keyboard
Fire	50.976	40.449	33.731	53.669	43.789
Siren	67.391	58.135	75.178	55.299	58.156
Telephone	61.893	45.268	34.621	55.933	58.131
Buzzer	55.286	27.242	23.504	33.697	38.979

and in various noise environments that are representative of the diverse needs of the prospective users of the device. We employed stationary signal segments from four Type I signals (fire alarm, telephone, buzzer, and tornado siren) which are likely to be of interest to any device owner. We selected four typical noise environments commonly encountered in daily life (laughing people, party noise, background highway noise, and computer keyboard noise).

The assumption was made that the stationary signal segments and the chosen noise environments are equiprobable and cost of an error is same for all signals to be detected. Obviously, these assumptions are not valid in practice. For example, probability of a telephone signal is much higher than the probability of a fire alarm signal. On the other hand, the cost of missing a fire alarm is much higher than the cost of missing a telephone signal. The results of the simulations performed under these assumptions, however, provide a reasonable estimate of the expected performance of the decision block when deployed in the field. They also indicate how the choices of the sensitivity parameter, α , and frame length, L , are likely to affect the detection rate.

4.3.2 Performance of the Decision Rule

To evaluate the performance of the decision rule the following procedure was used: We sampled 20 frames of each stationary signal segment and noise environment from the test set. Next, we constructed noisy test frames by scaling the powers of the signal and noise frames according to the desired SNR before adding them. To simulate the effects of the AGC, we normalized the power of the generated test frames to unity (see Section 2.2). Using this procedure, we constructed a collection of test frames. We constructed this collection in such a way that all signal and noise environments of interest have the same weight so that they satisfy the equiprobable priors assumption described above. This collection of noisy observation frames was then input to the

various detector structures to assess performance.

The tests were first conducted under noiseless conditions. To simulate a noiseless environment, we set $\text{SNR} = \infty$ in the algorithm. All filter designs (matched filter, eigenfilter, prediction error filter, and autocorrelation detector) yielded 100% detection rate under these conditions.

Next, for the above signal set, we performed simulations to obtain the power curves [9] of the detectors for the SNR range $-5 - 5$ dB. The resulting error rate curves are shown in Figures 4.1–4.9 for several values of frame length, L , and decision threshold (sensitivity parameter), α .

As indicated in Section 2.2.2, the alerting device is not expected to operate at SNRs below 0 dB; however obtaining the performance for unrealistically low SNRs is important for theoretical studies and design of more robust methods. We observe that detection rate initially increases with increasing L . However, if L becomes too large, the adverse effects of arbitrarily assigning the decision instants will reduce the detection rate.

Caution must be taken in interpreting the power curves shown in Figures 4.1–4.9. As presented, these curves represent extremely conservative indications of performance capability. It is important to recall that the final decisions regarding the presence of an alerting signal is given by the postprocessor block. In making this final decision, the postprocessor uses a number of outputs (typically 7-15) from the decision block; therefore, for an incorrect final decision, the decision block must make numerous erroneous decisions in a very short time frame. Accordingly, the error rate of the postprocessor will be much lower than the error rate of the decision block. The majority selection rule used in the postprocessor algorithm effectively corrects the errors made by the decision block.

It is observed that for $1 < \alpha < 1.2$ and for all tested values of L , the total error rate of the eigenfilter and autocorrelation detector is very close to 0% for the range of

practical interest. The matched filter performs slightly worse. Nevertheless, its total error rate is less than 5% for individual frames.

Although the prediction error filter has a good SNR improvement factor, it resulted in rather poor performance on the detection rate benchmarks. An examination of the filter output values showed that the reason for this was the gain scaling problem described in Section 3.3.2.

We note that the test procedure employed in this study is not suitable for determining the effects of synchronization on performance. A meaningful scheme for testing the effects of synchronization requires a large database of alerting signals which is currently not available.

4.3.3 Performance of the Overall Detector

Based on the simulation results above, we can reliably estimate the performance of the overall detector including postprocessing.

Consider the postprocessor algorithm developed in Section 3.4. Let $P = 10$ and $\beta = 0.6$. These parameters were chosen based on the experience gained by analyzing alerting signal waveforms. Next assume that the decisions made by the postprocessor algorithm are independent, and that the error rates for individual stationary signal segments are equal.

Under these assumptions, using the Bernoulli distribution, the overall detector performance for the given alerting signal set can be estimated. Let $P_{e,d}$ be the error rate of the decision block at a given SNR and let x be the nearest integer to βP . Then the overall error probability of the detector at this SNR will be approximately given by

$$P_{e,o} \approx 1 - \sum_{k=x}^P \binom{P}{k} P_{e,d}^{P-k} (1 - P_{e,d})^k. \quad (4.13)$$

The overall performance curves determined in this manner appear in Figures

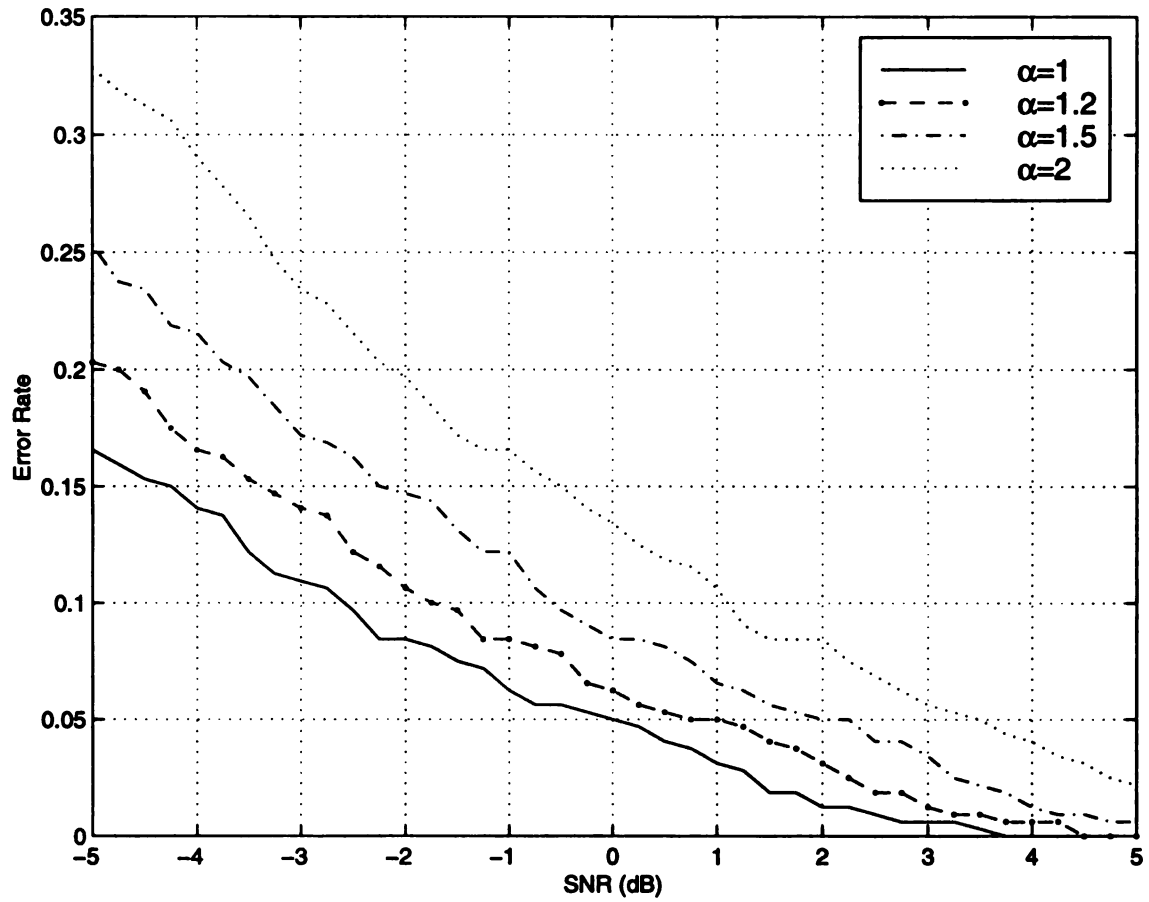


Figure 4.1. Power curves of the matched filter detector for $L=750$ samples and several values of α .

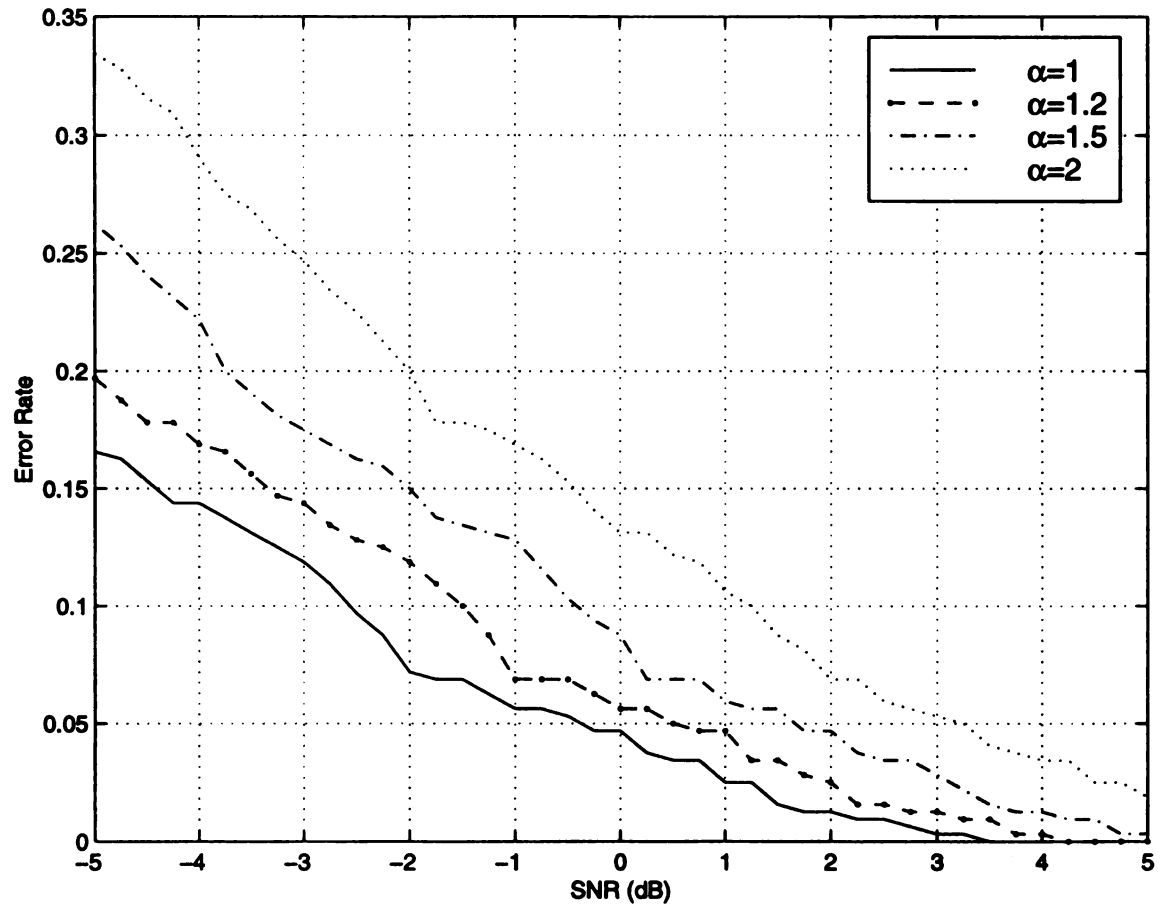


Figure 4.2. Power curves of the matched filter detector for $L=1000$ samples and several values of α .

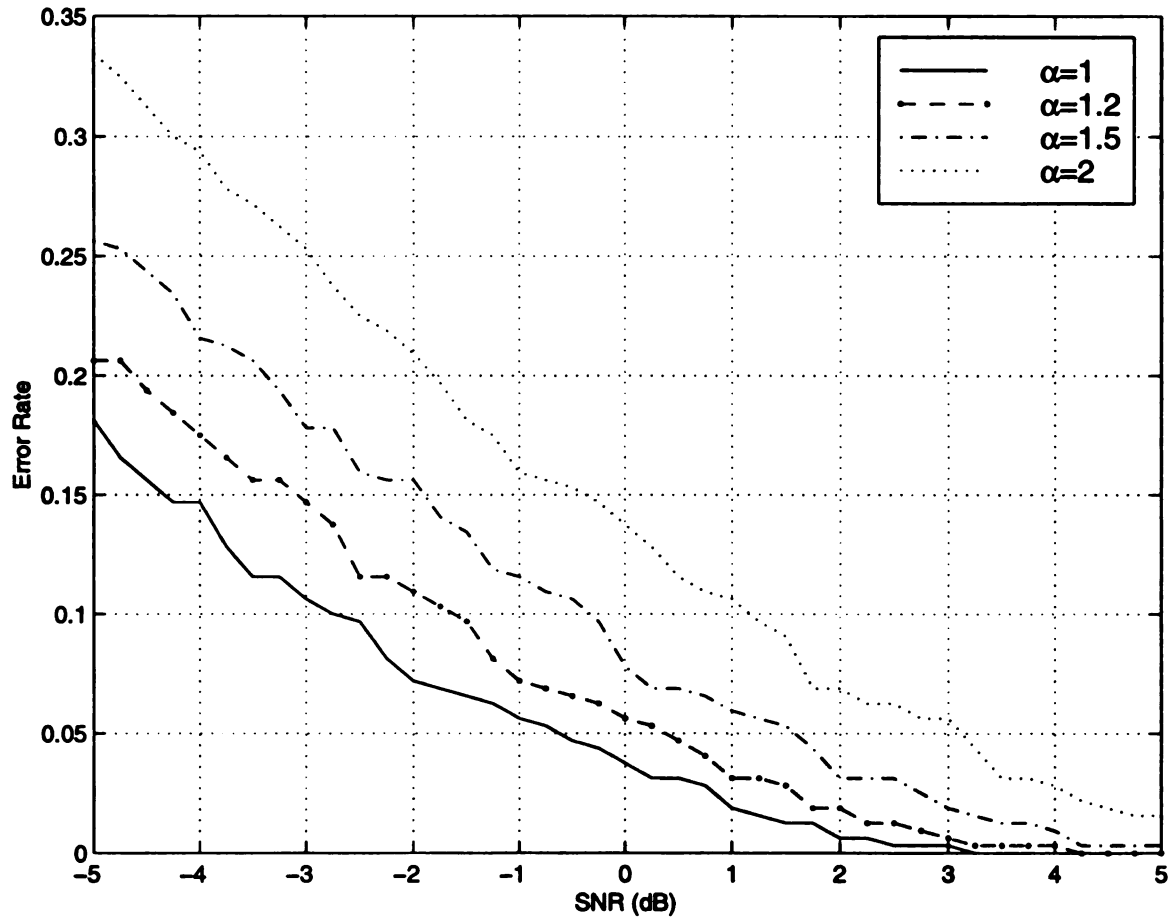


Figure 4.3. Power curves of the matched filter detector for $L=1500$ samples and several values of α .

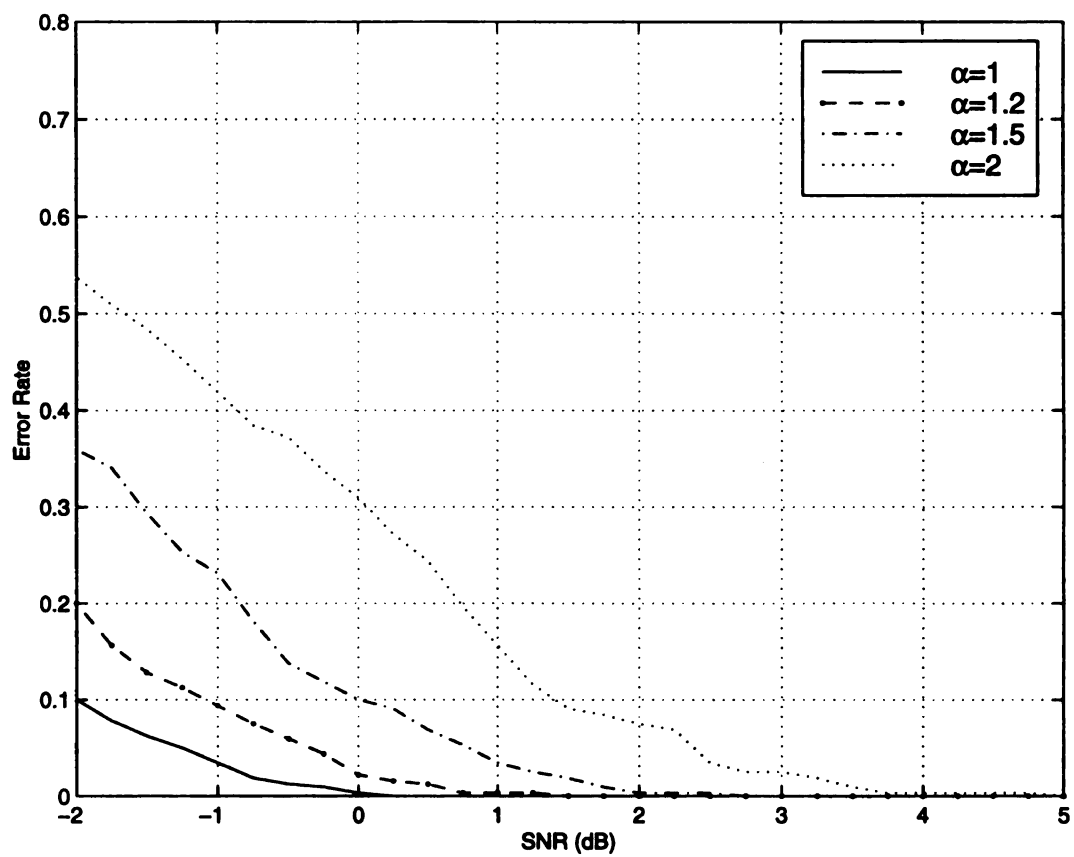


Figure 4.4. Power curves of the eigenfilter detector for $L=750$ samples and several values of α .

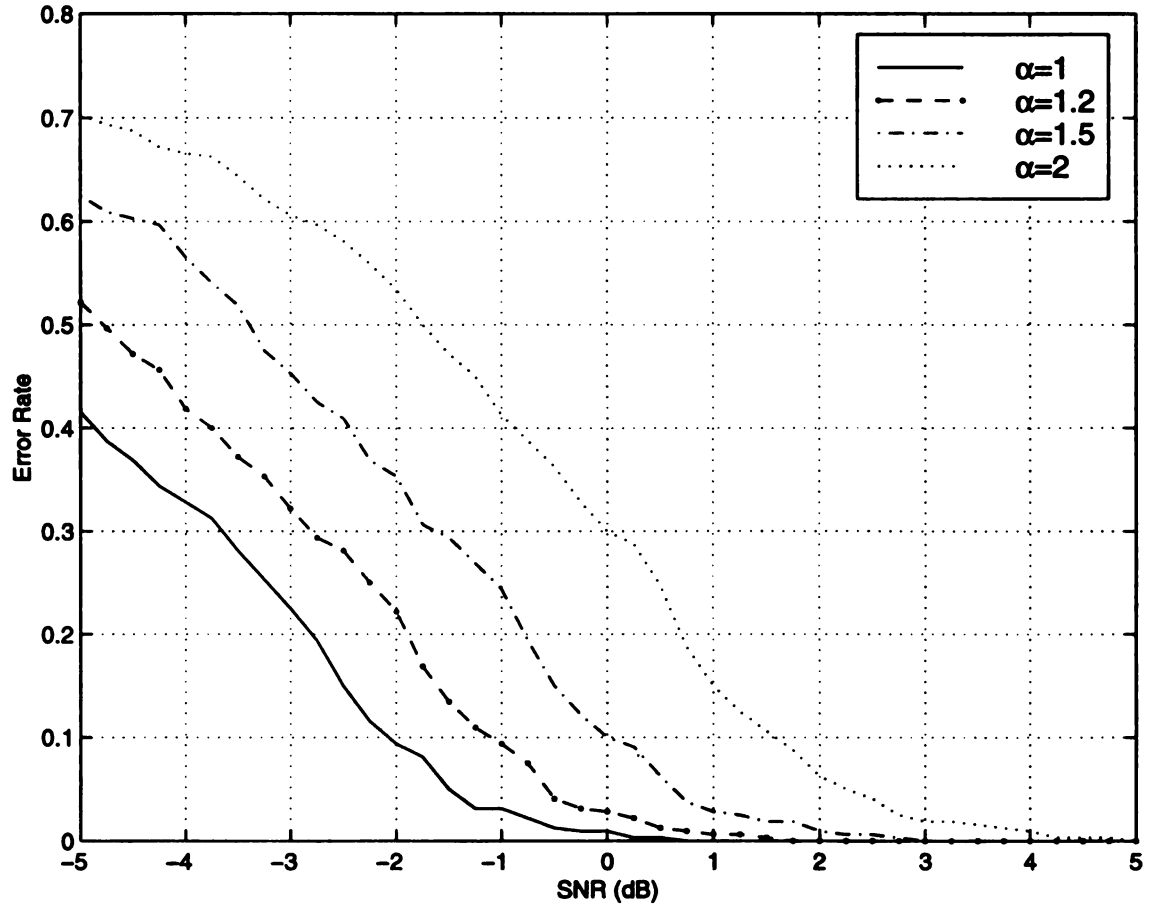


Figure 4.5. Power curves of the eigenfilter detector for $L=1000$ samples and several values of α .

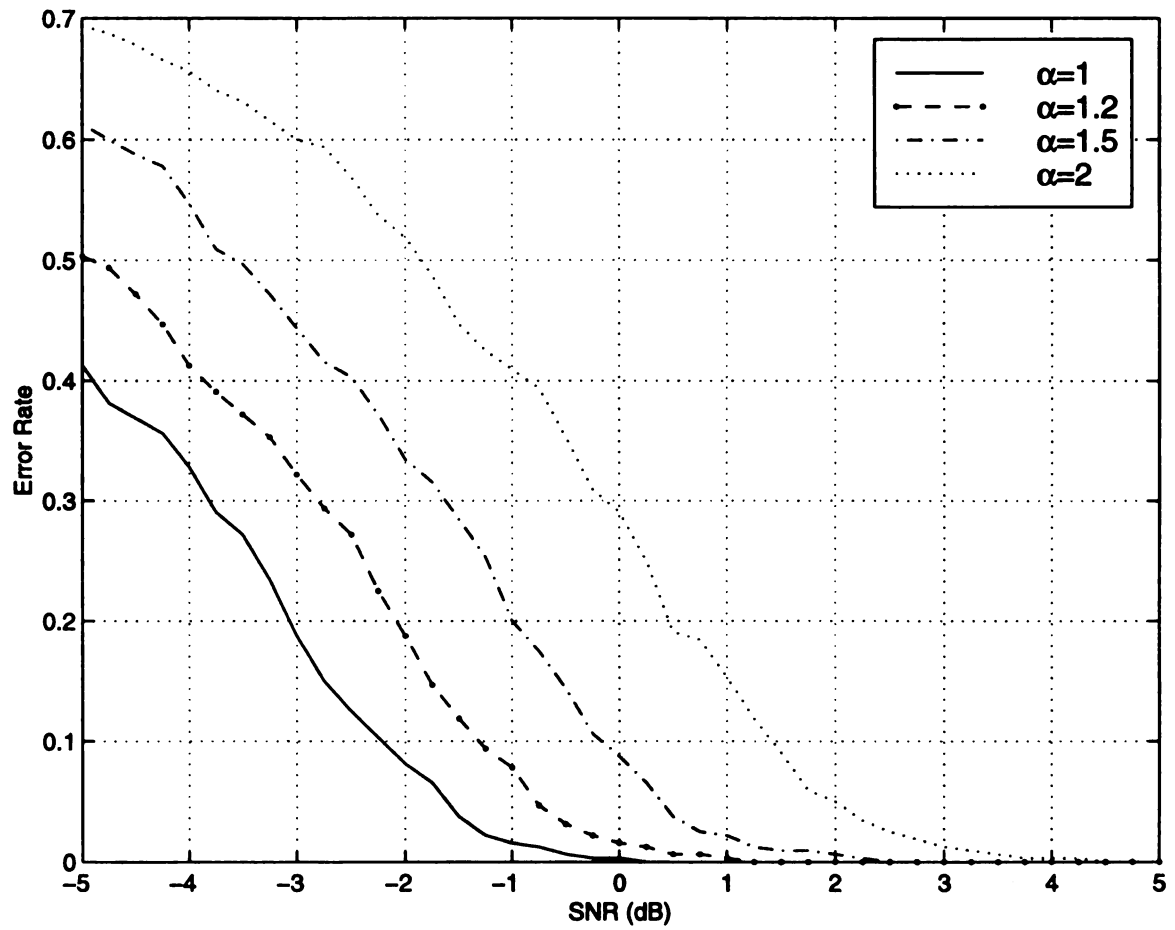


Figure 4.6. Power curves of the eigenfilter detector for $L=1500$ samples and several values of α .

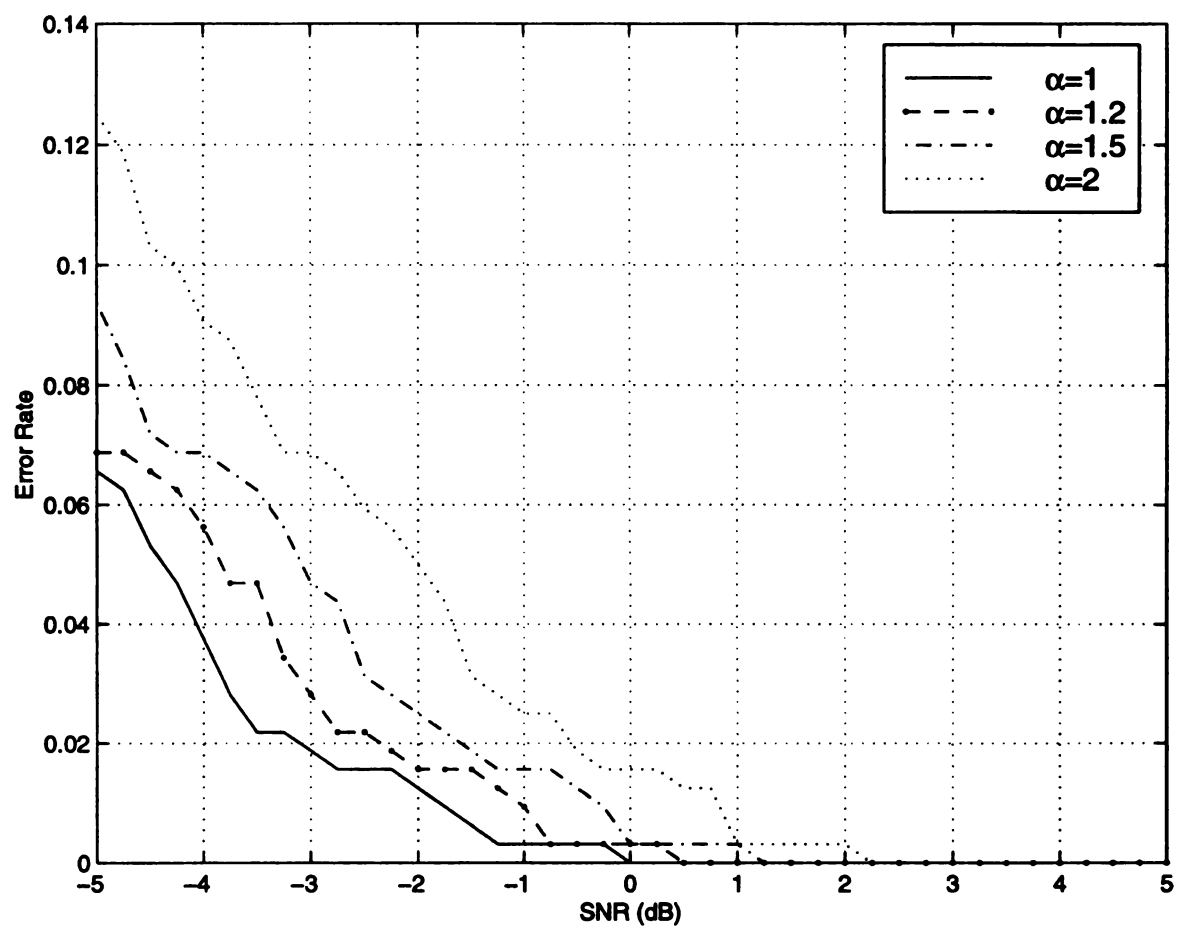


Figure 4.7. Power curves of the autocorrelation detector for $L=750$ samples and several values of α .

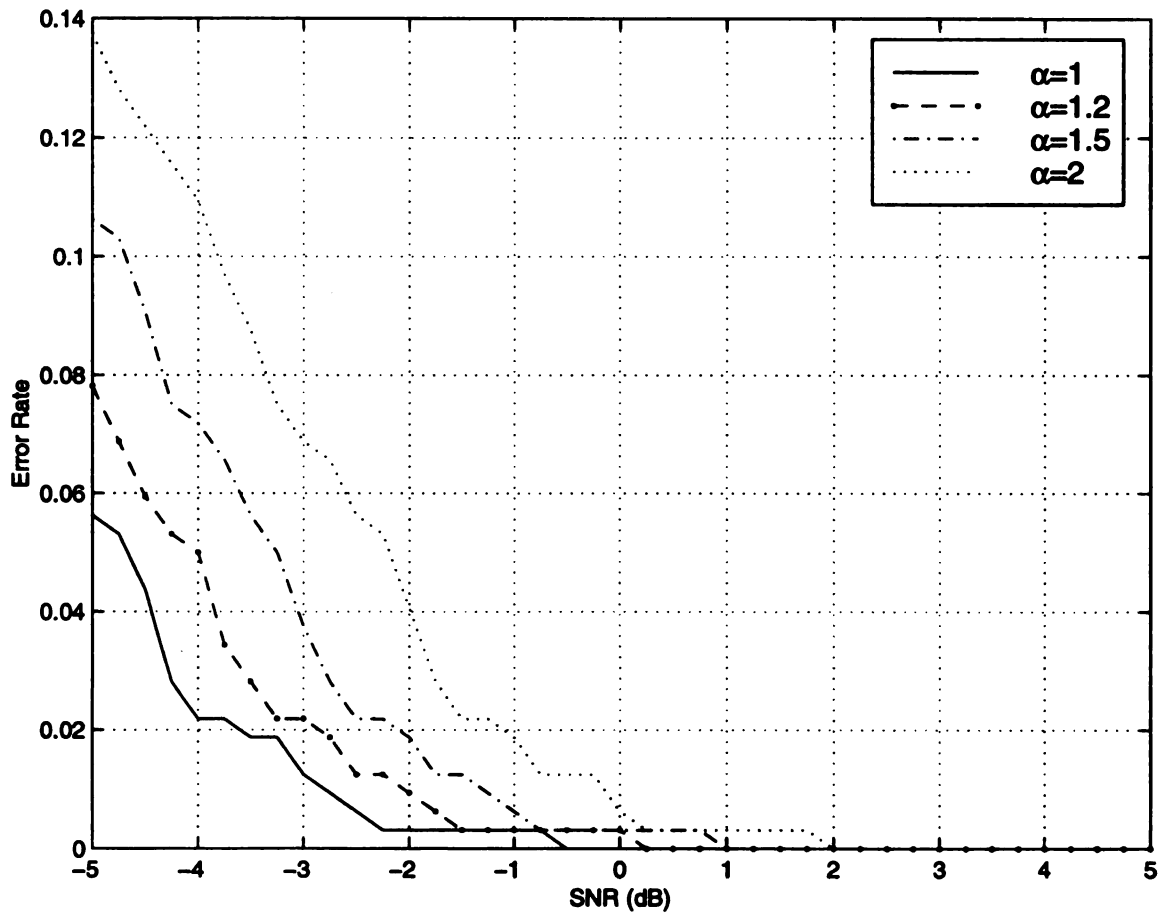


Figure 4.8. Power curves of the autocorrelation detector for $L=1000$ samples and several values of α .

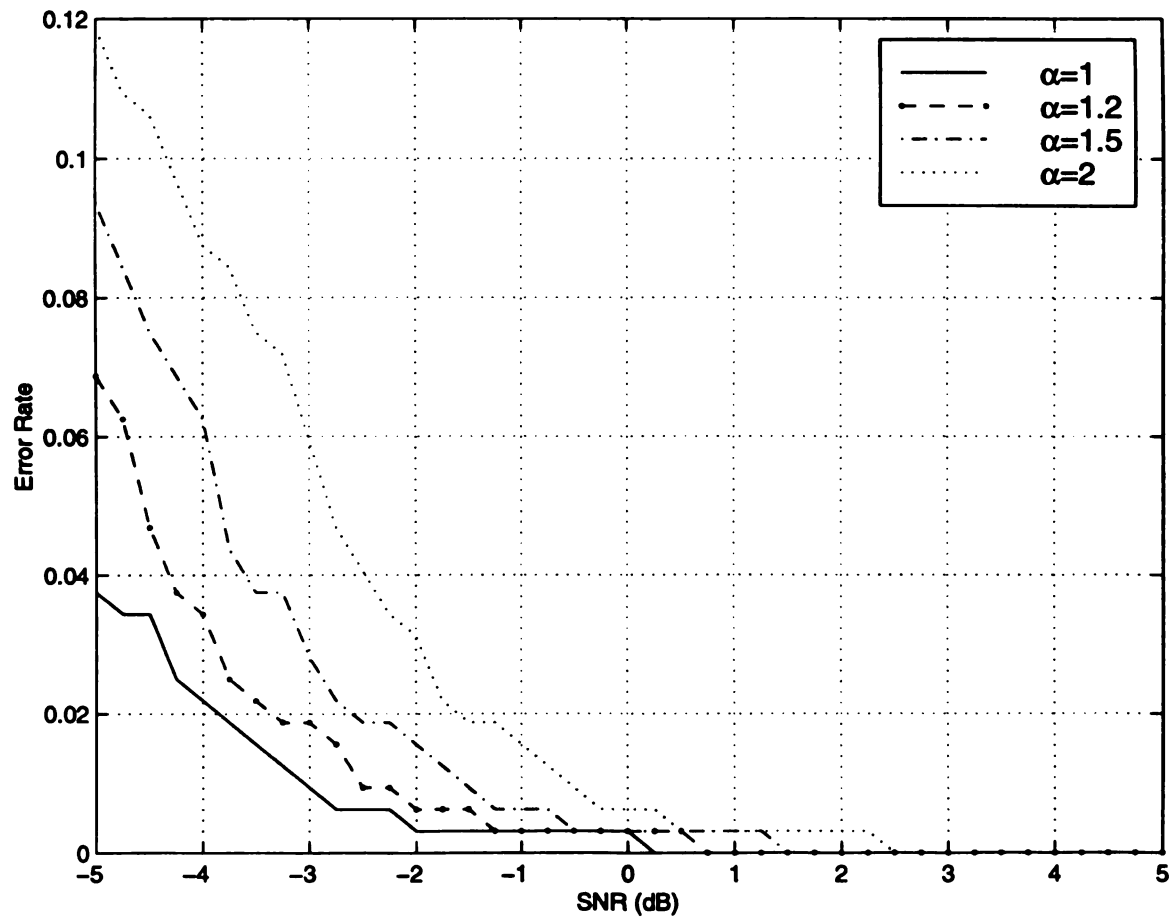


Figure 4.9. Power curves of the autocorrelation detector for $L=1500$ samples and several values of α .

4.10–4.12. We observe that the error rate of the autocorrelation detector is below $2 \cdot 10^{-3}$ even for -3 dB, and the rate is practically 0 for the SNRs at which the device will ordinarily operate.

Other meaningfully chosen values of P and β also give similar performance. The ultimate selection of these parameters should be based on synchronization considerations.

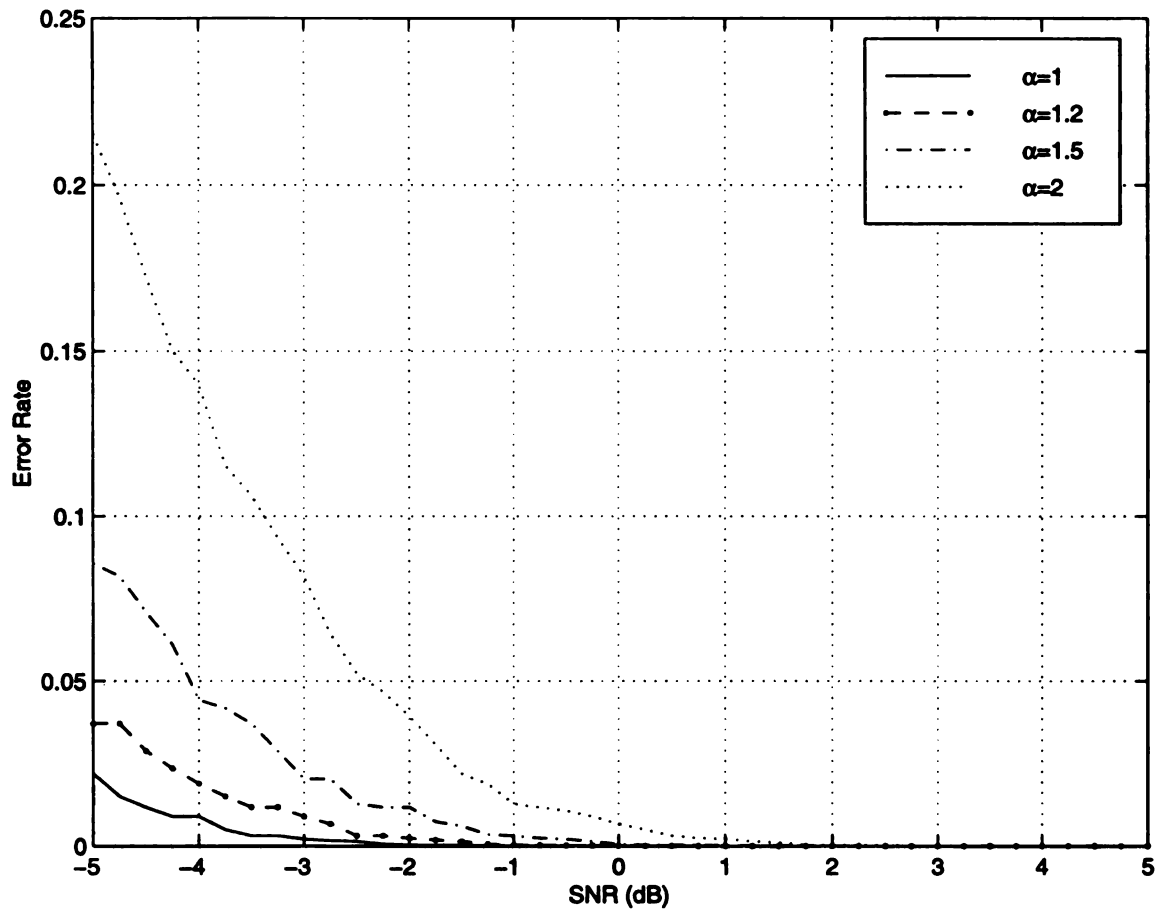


Figure 4.10. Overall error rate of the matched filter detector for $L=1500$ samples and several values of α .

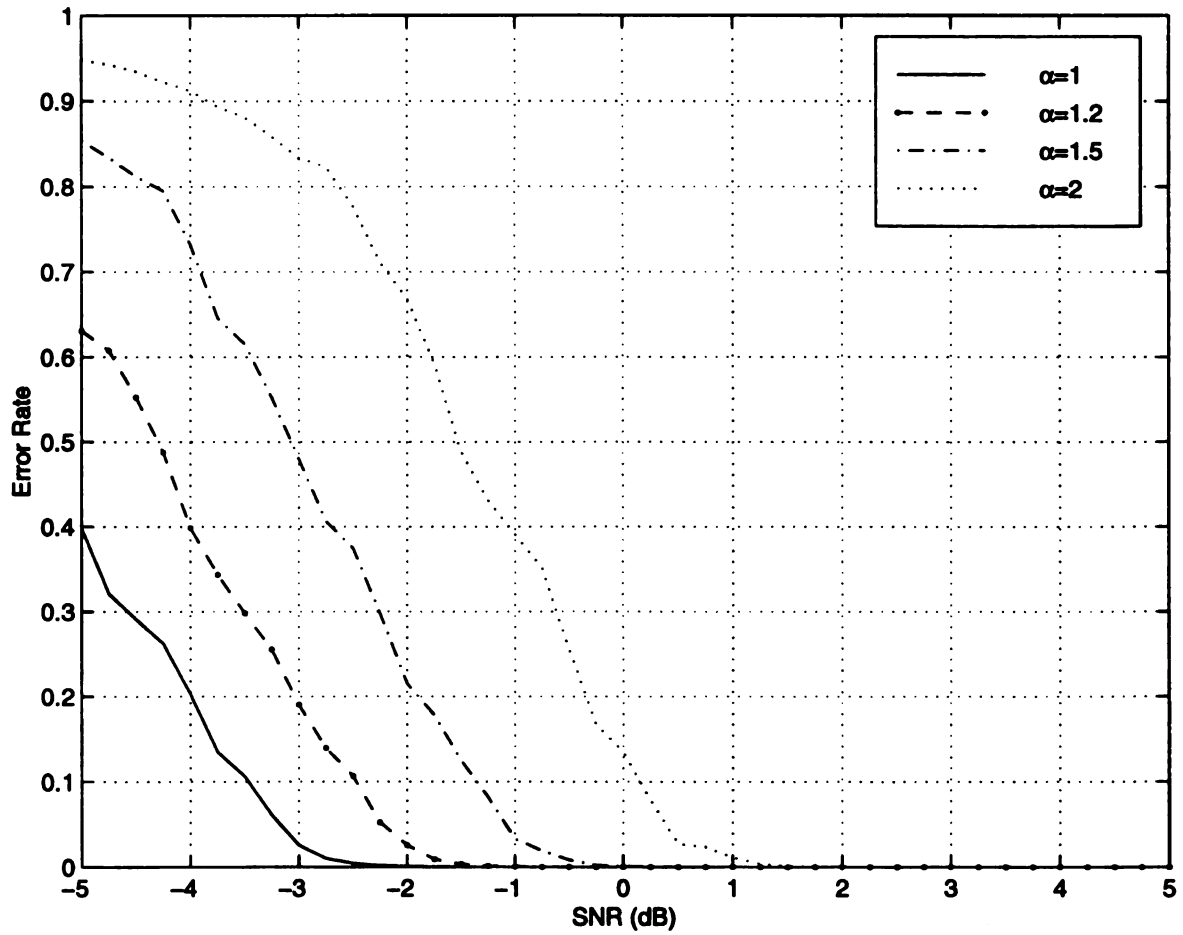


Figure 4.11. Overall error rate of the eigenfilter detector for $L=1500$ samples and several values of α .

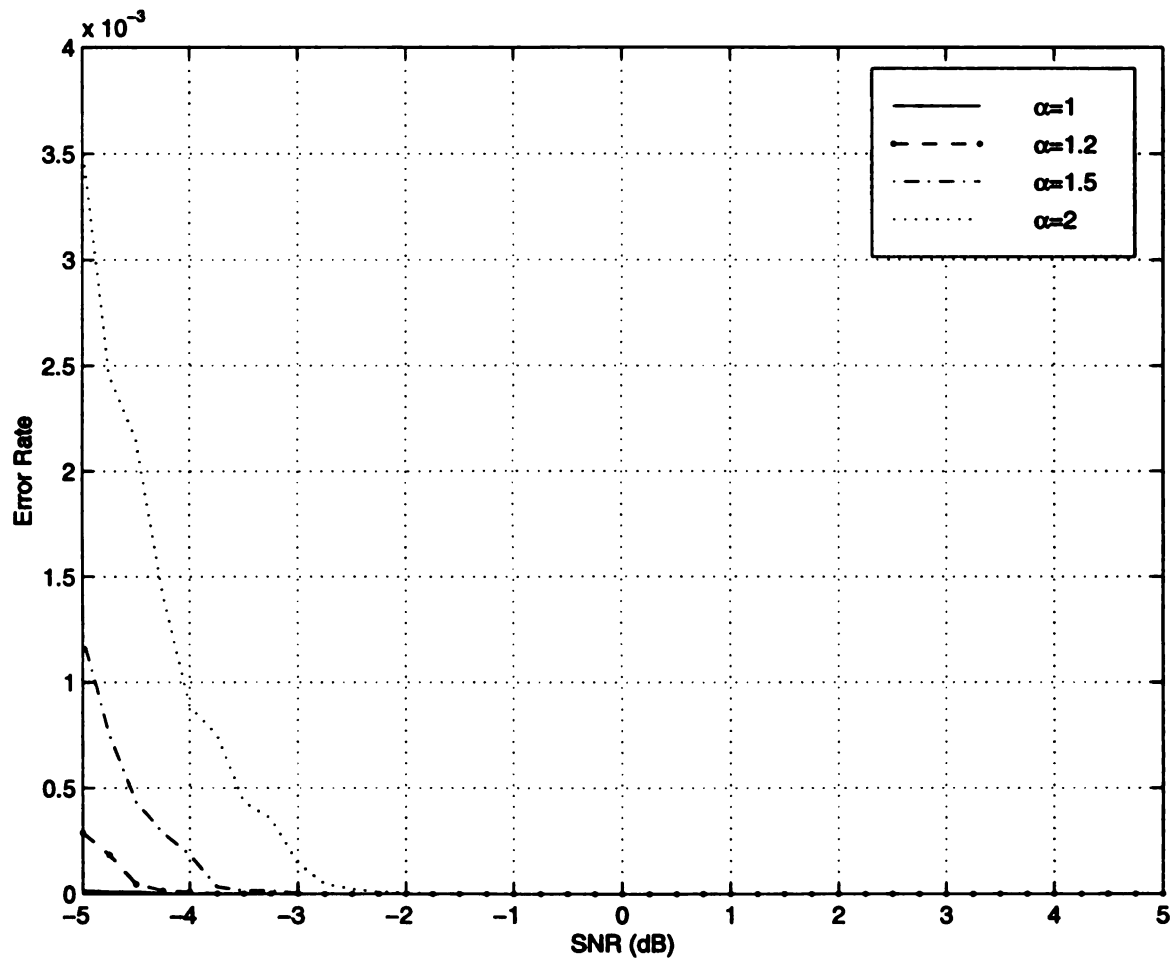


Figure 4.12. Overall error rate of the autocorrelation detector for $L=1500$ samples and several values of α .

CHAPTER 5

Conclusions and Considerations for Future Research

5.1 Conclusions

The simulation results reported here show exceedingly good detection performance of the developed algorithms with excellent false alarm rates over the battery of tests performed in the study. The results also imply very robust performance in rather extreme noise environments for the alerting signals tested. This demonstration of feasibility projects a very viable, useful, and commercially successful product based on the developed technology.

The biggest difficulty in developing the detection algorithm presented in this thesis was caused by the presence of the AGC block. If the SNR could be reliably estimated, we could design a much more reliable decision block and devise a rigorous criterion for adjusting the gains of the detection filters. This would make the error rate of the alerting device practically zero. This study has employed readily available sampling boards on personal computers. These boards do not give access to the instantaneous gain of the AGC. Using custom-designed sampling circuitry which gives access to (and, perhaps control over) the gain of the AGC would permit estimation of the

instantaneous SNR of the operating environment. Even a gross estimate of SNR would enable the design of a vastly improved decision rule, and this would significantly enhance the performance of the developed algorithms. Such a sampling board should serve as the prototype of the sampling circuitry of the final device.

5.2 Future Research

The scope of this work was limited with the development of the detector stage of the adaptive alerting device. The next major phases of this project which will eventually lead to a commercially available alerting device are briefly summarized below:

- **Testing of the developed algorithms:** In this thesis, only interim simulation results pertaining the first two stages of the adaptive alerting device, namely the detection filter bank and decision block were presented. More extensive studies incorporating the performance of the postprocessor block should be performed. Additionally, effects of microphone placement and different microphone types should be addressed.
- **Development of training algorithms:** As indicated in the introductory chapter, the ultimate device will be user-trainable for any alerting signal. Easy to use training algorithms for determining filter bank responses in an unsupervised manner must be developed.
- **Hardware Implementation:** As with any DSP system operating in real time, computationally efficient forms of the suggested algorithms must be developed. Effects of quantization of detection filter parameters must be investigated. Finally, a suitable DSP-processor for use in the ultimate device must be chosen.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] Jont Allen, "How do Humans Process and Recognize Speech?," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 4, October 1994.
- [2] Peter J. Brockwell and Richard A. Davis, *Introduction to Time Series and Forecasting*. New York: Springer, 1996.
- [3] George R. Cooper and Clare D. McGillem, *Modern Communications and Spread Spectrum*. New York: McGraw-Hill, 1986.
- [4] John R. Deller, Jr., John G. Proakis, and John H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New York: Mcmillan, 1993.
- [5] John R. Deller, Jr., *An Adaptive Alerting System for Persons with Hearing Disabilities*, Grant Application submitted to State Reseach Fund (Michigan). Waterford: February 1996
- [6] Simon S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice Hall, 1996.
- [7] Simon S. Haykin, "Neural Networks Expand SP's Horizons," *IEEE Signal Processing Magazine*, vol. 13, no. 2, pp. 24-49, March 1996.
- [8] <http://www.cochlear.com.au>
- [9] Saleem A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer Verlag, 1988.
- [10] Saleem A. Kassam and H. Vincent Poor, "Robust Techniques for Signal Processing: A Survey," *Proceedings of the IEEE*, vol. 73, pp. 433-481, 1985.
- [11] Saleem A. Kassam and John B. Thomas, *Non-parametric Detection: Theory and Applications*, Stroudsburg: Dowden, Hutchinson & Ross, 1980.
- [12] Jae S. Lim and Alan V. Oppenheim (editors), *Advanced Topics in Signal Processing*, New Jersey: Prentice Hall, 1988

- [13] Alan V. Oppenheim and Ronald W. Schaffer, *Discrete-Time Signal Processing*. New Jersey: Prentice Hall, 1989.
- [14] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*, Second Edition. New York: Springer Verlag, 1994.
- [15] Lawrence Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, February 1989.
- [16] Robert J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: John Wiley, 1992.
- [17] Harvey F. Silverman and David P. Morgan, "The Application of Dynamic Programming to Connected Speech Recognition," *IEEE Acoustics, Speech, and Signal Processing Magazine*, vol. 7, pp. 6-25, July 1990.
- [18] Michael Valente, *Strategies for Selecting and Verifying Hearing Aid Fittings*. New York: Thieme Medical Publishers, 1994.
- [19] Andrew J. Viterbi and James K. Omura, *Principles of Digital Communications and Coding*. New York: McGraw-Hill, 1979.

MICHIGAN STATE UNIV. LIBRARIES



31293017102801