IDEA PROJECT FINAL REPORT
Contract ITS- 19

IDEA Program
Transportation Research Board
National Research Council

December 26, 1995

AutoAlert: Automated Acoustic Detection of Incidents

David A. Whitney and Joseph J. Pisano
TASC, Inc., Reading, Massachusetts

Approved by: Robert E. Introne

The ITS-IDEA program is jointly funded by the U.S. Department of Transportation's Federal Highway Administration, National Highway Traffic Safety Administration, and Federal Railroad Administration. For information on the IDEA Program contact Dr. K Thirumalai, IDEA Program Manager, Transportation Research Board, 2101 Constitution Avenue N.W., Washington, DC 20418 (phone 202-334-3568, fax 202-334-3471).

INNOVATIONS DESERVING EXPLORATORY ANALYSIS (IDEA) PROGRAMS MANAGED BY THE TRANSPORTATION RESEARCH BOARD (TRB)

This investigation was completed as part of the ITS-IDEA Program, which is one of three IDEA programs managed by the Transportation Research Board (TRB) to foster innovations in surface transportation. It focuses on products and results for the development and deployment of intelligent transportation systems (ITS), in support of the U.S. Department of Transportation's national ITS program plan, The other two IDEA programs areas are TRANSIT-IDEA, which focuses on products and results for transit practice in support of the Transit Cooperative Research Program (TCRP), and NCHRP-IDEA, which focuses on products and results for highway construction, operation, and maintenance in support of the National Cooperative Highway Research Program (NCHRP). The three IDEA program areas are integrated to achieve the development and testing of nontraditional and innovative concepts, methods, and technologies, including conversion technologies from the defense, aerospace, computer, and communication sectors that are new to highway, transit, intelligent, and intermodal surface transportation systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Project Overview

AutoAlert applies new signal processing algorithms to passive acoustic data to advance the state of practical acoustic incident detection techniques. These techniques, originally developed for national defense applications, will perform reliable, automatic, nearly instantaneous, all-weather incident detection under highly variable traffic conditions. Effective operation of urban high-capacity ITS systems requires speedy detection of incidents at chokepoints, such as tunnels, bridges and other aerial structures, and dense urban arterials. Boston's Central Artery/Tunnel (CA/T) project is an example of a new ITS system where such detection is critical. AutoAlert overcomes shortcomings of loop and video detectors, such as their inability to distinguish between incidents and congestion, and the need for a human-in-the-loop for video detection. The AutoAlert processor "hears" an incident before congestion builds, and can be used either as an independent detector, or its outputs can be combined (data fusion) with other detector outputs for joint improved decisions and incident verification.

The problem of rapid, reliable acoustic incident detection is more complex and difficult than the problems of freeway traffic flow monitoring or vehicle type identification, for which acoustic sensors have already been applied. The AutoAlert processor (see Fig. 1-1) will make use of readily available commercial acoustic *sensors* (e.g., AT&T IVHS NET-2000™). **What is key to AutoAlert is the signal analysis and detection *algorithms* that** we **will employ.** The algorithms will provide a new level of incident detection timeliness and reliability (low false alarms) by applying sophisticated statistical models: Hidden Markov Models (HMM) and Canonical Variates Analysis (CVA). These are used to analyze both short-term and time-varying signals that characterize incidents.

1.2 Principles of Innovation

AutoAlert algorithms provide performance innovations that offer significant improvements over today's most "advanced" acoustic traffic sensor systems:

- Provide *nearly immediate detection,* with no alertment delays, before congestion builds
- React to an incident *directly,* not just the *symptom* of an incident (e.g., congestion)
- Provide *low false alarm rates* via simultaneous analysis of short-, medium-, and long-term acoustic feature patterns

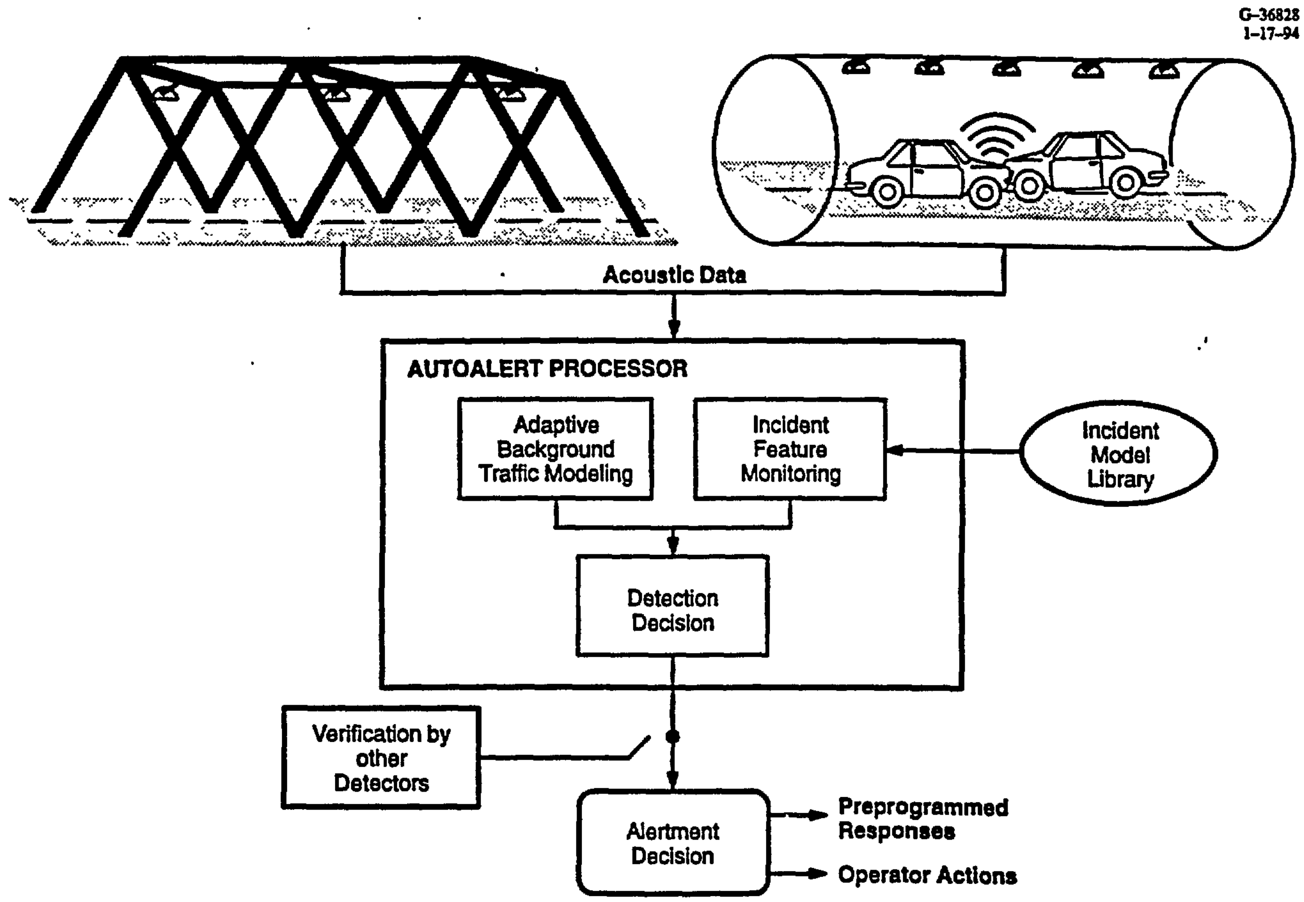


Figure 1-1 AutoAlert Processor Automatically Monitors for Incident Features That Appear Against Changing Background Noise

- Offer *reliable, all weather, day or night detection* under varying traffic conditions through the use of sophisticated data models
- Report a *confidence level* for detections that can be used by AutoAlert as a "self—test" of algorithm performance
- Accommodate minute—to—minute *normal variation* in traffic conditions using *adaptive,* real—time analysis of the background noise
- Capture dynamic characteristics of very short—term *acoustic patterns* for better detection and identification of incidents in noisy traffic conditions

- Account for that fact that no two incident acoustic signatures (e.g., "screech, crunch") are identical by using probability-based modeling
- Apply *unique* new algorithms developed and proven by TASC for defense and other commercial acoustic applications.

1.3 ITS Need

A number of sensor systems have been designed to monitor traffic conditions (pressure sensors, in-ground loop detectors, digital video--based flow estimators and acoustic sensors like the AT&T SmartSonic™ surveillance system). However, they do not distinguish between impeded traffic flow and an incident. They can only provide *inferred* detection of incidents because they only sense and report the resulting congestion. *They react to the symptoms of an incident, rather than to the incident directly, and are therefore* less *effective.* This limitation is particularly prohibitive in the monitoring of traffic chokepoints that have limited or constrained access (e.g., tunnels, aerial structures), where a single incident can cause back-ups to form at rates of hundreds of feet per minute. Immediate detection and identification of incidents in such corridors improves traffic flow and driver safety by reducing the time-lag between the actual event and the emergency response. In addition, loop detectors have been historically prone to high failure rates, while video camera installations are hampered by adverse environments, such as darkness, precipitation, fog or dust, and require human monitoring of many cameras for immediate detections. *The AutoAlert processor provides automated, nearly immediate, direct acoustic incident detection.*

1.4 Technical Challenges

Incidents such as automobile accidents share common characteristics that make them immediately recognizable to human listeners (screech of tires, fender crunch, breaking glass). These characteristics are represented in the very short-term (transient) acoustic pattern of the incident. They provide sufficient information for an automated algorithm to distinguish an incident from *average* traffic background, as well as other similar but benign background events (car-door slamming, car backfiring, sirens, thunder, trucks hitting potholes). The design of a detection processing algorithm is critical to meet such requirements because of

- The high degree of random variation in acoustic signatures from incident to incident (e.g., no two tire-screeches are identical)
- Varying environmental (weather) and traffic conditions (volume, composition)

- The need to keep false alarm rates low in the face of these complexities.

These factors make the performance of classical and many modem detection and identification techniques generally inadequate for this problem.

1.5 New Solutions with Broad Applications

TASC's new Hyperstate algorithms, which are the core of AutoAlert, overcome shortcomings of classical detection and identification techniques by:

- *Adaptively* characterizing time-varying background noise, continually adjusting for varying environmental and traffic conditions.
- Looking for abrupt changes from background noise (such as sudden changes in energy patterns at certain frequencies) that could signal an incident.
- Automatically accounting for expected random variation in incident signatures using built-in stochastic models.
- Using sophisticated simultaneous analysis of both long-term (several seconds) and short--term (milliseconds) acoustic patterns to accurately identify incidents *while lowering false alarm rates.*
- Reporting a probability/likelihood-based *confidence level* with each decision to improve decision quality and allow "self-tests" of AutoAlert performance.

The AutoAlert processor provides a set of computer-based algorithms that can be mated with a variety of existing commercial acoustic sensors which may already be in place for traffic flow monitoring. In addition to serving as a self-contained incident detection product, the AutoAlert technology can be used in several other ways:

- In conjunction with existing detection systems, serving as an automated cue to operators to focus their attention or cameras on a specific area.
- Detection of *acoustic emergency beacons,* preprogrammed with distinctive sound patterns, that are automatically emitted by distressed vehicles.
- Identify and track acoustic *probes* (using a vehicle's unique acoustic signature), thus offering independent estimates of traffic flow that can be used to verify and calibrate other flow sensors.
- Analyze data from a variety of non-acoustic sources, including loop detectors, using the underlying Hyperstate methodology.

The core AutoAlert signal processing technology has already been developed under defense applications and offers a unique capability not found in any existing system. As a result, AutoAlert has the potential to provide a low-risk, high-value capability for ITS.

2. RESEARCH PLAN

2.1 Statement of Work

The AutoAlert program proceeded in two stages: 1) Design and Preliminary Evaluation, and 2) Feasibility Demonstration. Four tasks were accomplished to complete Stage 1, and one task for Stage 2.

Stage 1: Design and Preliminary Evaluation

Task 1.1: Assemble Acoustic Database

Obtain pre-recorded accident data from sources such as the Insurance Institute for Highway Safety, vehicle crash test performed by US. auto manufacturers, NHTSA, and commercially available audio “sound effects”. Collect background traffic noise data under varying operational conditions as required. These data will serve to build the very important database component of “incident-fi-ee” data.

Task 1.2: Acoustic Feature Analysis

Analyze the acoustic database to define accident and incident types, and to determine the optimal feature sets for acoustic discrimination of different incident types (e.g. spectral energy, bispectral energy, transient types). Determine the range of dynamic time scales required to capture both transient and longer duration sounds. For example, a transient burst of acoustic energy around 500 Hz lasting from 0.5 to 3 sec may be one feature of a vehicle impact.

Task 1.3: Optimized Algorithm Architecture

Tailor the AutoAlert algorithms to provide a processing architecture that is optimized for incident detection and classification. Select appropriate Hyperstate model parameters, such as the number of model states, length of time scales, and number of time scales. Identify appropriate metrics for model goodness of fit to data. Select decision procedures and criterion for operator alerts based on AutoAlert processing. Implement a prototype AutoAlert software system on a 486-class PC or low-end workstation.

Task 1.4: Preliminary Performance Testing

Using controlled experiment data extracted from the acoustic database, perform preliminary AutoAlert performance evaluation. The data will be used in a simulation mode to develop test data sets that reflect varying signal-to-noise ratios (SNR).

AutoAlert's false alarm and correct detection performance as a function of SNR for selected incident types will be evaluated.

Stage 2: ***Feasibility Demonstration for Operational Data***

Task 2.1: Final Performance Testing

Demonstrate the overall feasibility of AutoAlert processing in the laboratory. Additional analysis of controlled experiment data begun in Task 1.4 will be completed. Based on lessons learned, the algorithm will be modified as required. In addition, field data collected in "uncontrolled experiments", e.g. from recordings made on-site at a Boston highway, will be evaluated in the laboratory. This will allow important features of the algorithm, such as real-time background noise characterization and response to random variations in incident characteristics.

2.2 Research Milestones

The two major milestones for this project have been completed, Milestone 1 and the work leading up to it was reported in the AutoAlert Interim Report [Ref.l], while Milestone 2 is reported on in this report.

Milestone M1 (Stage 1): Establishment of acoustic database, design of an optimized algorithm architecture, implementation of prototype algorithms based on this architecture, and preliminary performance evaluation using controlled, synthetic data sets

Milestone M2 (Stage 2): Overall feasibility demonstration using available "controlled "and "uncontrolled" field data. Analysis of correct detection and false alarm performance for realistic traffic environments, and operating-aiding functionality

2.3 Documentation of Results

This Final report, in combination with the Interim Report [Ref. 1], provides a complete documentation of the AutoAlert project. Most of the material in the Interim Report is *not* repeated in this report -- references to the Interim document are included in this Final Report at appropriate points. The Interim Report covers the following topics:

- Theoretical background for multiresolution Hyperstate signal processing used in AutoAlert

- Discussion of model development procedures
- Detailed description of sources of acoustic data and the generation of complex hybrid acoustic data sets
- Analysis of incident and background characteristics to identify discriminating features
- Preliminary AutoAlert algorithm analysis architecture
- Appendix listing detailed results of feature analysis.

This Final Report covers the following topics:

- Project overview
- Research plan and milestones
- Description of the acoustic data sets used for performance analysis
- Identification of the key acoustic features for this data
- Description of the optimized AutoAlert algorithm architecture used for performance analysis
- Data processing results
- Description of an AutoAlert implementation for real-time data collection and processing
- Summary of project achievements.

3. ACOUSTIC DATABASE

3.1 The Need for Synthetic, Controlled Signals

The fundamental resource in the prototype AutoAlert system is the acoustic database. The acoustic database contains two types of audio data: background traffic sounds which may be thought of as “noise”; and traffic incident sounds which are the “signal” the system is trying to detect. The input to the detection algorithm described above is the combination of background and incident sounds. In order to assess and refine the prototype detection algorithm the acoustic database must be a controlled input.

3.2 Data Sources and Description

One approach taken to generate an acoustic database of synthetic traffic incident data with which to test against is to employ a hybrid model of two real data sources, a crash sound effects CD and field collected traffic recordings. The crash sound effects were commercially-available digital recordings of specific car crash scenarios. Conditions such as the number of cars involved, microphone location, vehicle direction, and approximate speed about each crash were specified in textual format with each recording. The field recordings are of “normal” traffic and were collected using AT&T’s SmartSonic Sensors and stored in audiovisual (videotape) format. Data was collected on two lanes of a four-lane divided highway under dry daylight conditions. The audio portion of the field recordings was captured with the microphone array and individual cars and the sounds they produce can be correlated using the visual portion of the recordings.

The sound effects recordings include 22 crashes, various horns and siren incidents, and 10 tracks of normal traffic passing. With a total tape time of approximately 45 minutes, the usable incident and background data is about 12 minutes long. The duration of the events vary widely. Some are as short as 3 seconds others as long as 20 seconds. The remainder of the usable data is. in the form of normal traffic passing by the microphone. While the text included with each recording provides a description of the road and vehicles, there is no specification of the recording system used (e.g. microphone type, source to microphone distance, or mastering process). Accordingly some assumptions were made, however the overlap in traffic conditions between the field recordings and these sound effects offered a qualitative and partially quantitative consistency verification.

Four hours of AT&T field recordings, all in VHS format, were previewed. The field recordings did not include any incidents which could have been used as signals.

Twenty-five minutes of divided, four-lane highway traffic sounds were digitized from the field recordings for use in generating the initial acoustic database of traffic background noise. Normal operational sound from all types of vehicles, including motorcycles, cars, trucks, busses, and tractor-trailer rigs, was captured. A single channel of the two-channel VHS acoustic data was processed -- the two channels did not differ significantly. As an additional future pre-processing step, the two channels of data could be combined to reduce instrument or random acoustic noise.

In order to assemble the acoustic database and provide the detection algorithm access to it, the digital recordings from both sources were converted into a common format and stored in magnetic form on a PC. This common format, called WAV, can be read by both Matlab™ and standard sound cards on the PC. This allows the database to be assembled with Matlab routines and then "played" through speakers for qualitative verification.

3.3 Hybrid Data Sets For Analysis

In order to provide a sufficiently rich database and allow for parametric variation of the relative magnitude of crashes relative to background data, a variety of mixtures of operational field traffic recordings and individual crash sounds were developed. Table 3-l summarizes the data sets that were developed.

Table 3-l Data Set Description Summary

Case	Duration	Passing Vehicles	Accidents	Scale Factor	Noise Added	Accident Energy Ratio
1	45 sec	2	1	1	no	-
2	118 sec	15-20	1	1	no	8.8
3	110 sec	15-20	2	1	no	-
2a	118 sec	15-20	1	1/2	no	4.4
2b	118 sec	15-20	1	1/4	no	2.2
2c	118 sec	15-20	1	1/8	no	1.1
2d	118 sec	15-20	1	1/16	no	0.55
2n	118 sec	15-20	1	1	yes	0.9
2an	118 sec	15-20	1	1/2	yes	0.45
2bn	118 sec	15-20	1	1/4	yes	0.23
2cn	118 sec	15-20	1	1/8	yes	0.11
2dn	118 sec	15-20	1	1/16	yes	0.055

Varying Backgrounds and Loud Accidents -- Illustrations of Cases 1 through 3 are presented in Figures 3-1 through 3-3. The amplitude signature for each data case is shown, with passing cars and accidents indicated. The accident amplitudes in these cases reflect an accident occurring close to the microphone.

Accidents of Varying Loudness in Traffic -- Cases 2a through 2d are derived from case 2. They are identical to case 2, except that the acoustic signature of the crash has been scaled by the scale factor indicated in Table 3- 1. For example, the crash used in case 2 has been scaled by a factor is l/2 to generate case 2a. The Accident Energy Ratio given in Table 3- 1 is a measure of the relative intensity of signal and noise/background (SNR), and is the ratio of the variance of the accident signal to the background signal over the interval of the accident. It measures the relative intensity of the signal relative to other traffic sounds. Figures 3-4 through 3-7 show the scaled signature of the accident segment, the background (absent of passing cars or other noise), and the combined accident signature and background for each case.

Accidents of Varying Loudness in Traffic and Noise -- Cases 2n and 2an through 2dn correspond to cases 2 and 2a through 2d. They differ in that, in addition **to** scaling, low-frequency noise with amplitude of approximately half that of the accident has been added to the entire 118 second time series. The frequency range for this noise was approximately 0-500 Hz, which overlaps the first Hyperstate bandpass feature response range. This noise is representative of what might result from rain or the noise from traffic on a distant highway, and was included to evaluate its impact on algorithm performance. The Accident Energy Ratio given in Table 3-1 is the ratio of the variance of the accident signal to the noise signal over the interval of the accident. It measures the relative intensity of the signal relative to other traffic sounds and noise. Figures 3-8 through 3-12 show the scaled signature of the accident segment, the low-pass noise, and the combined accident signature and noise for each case.

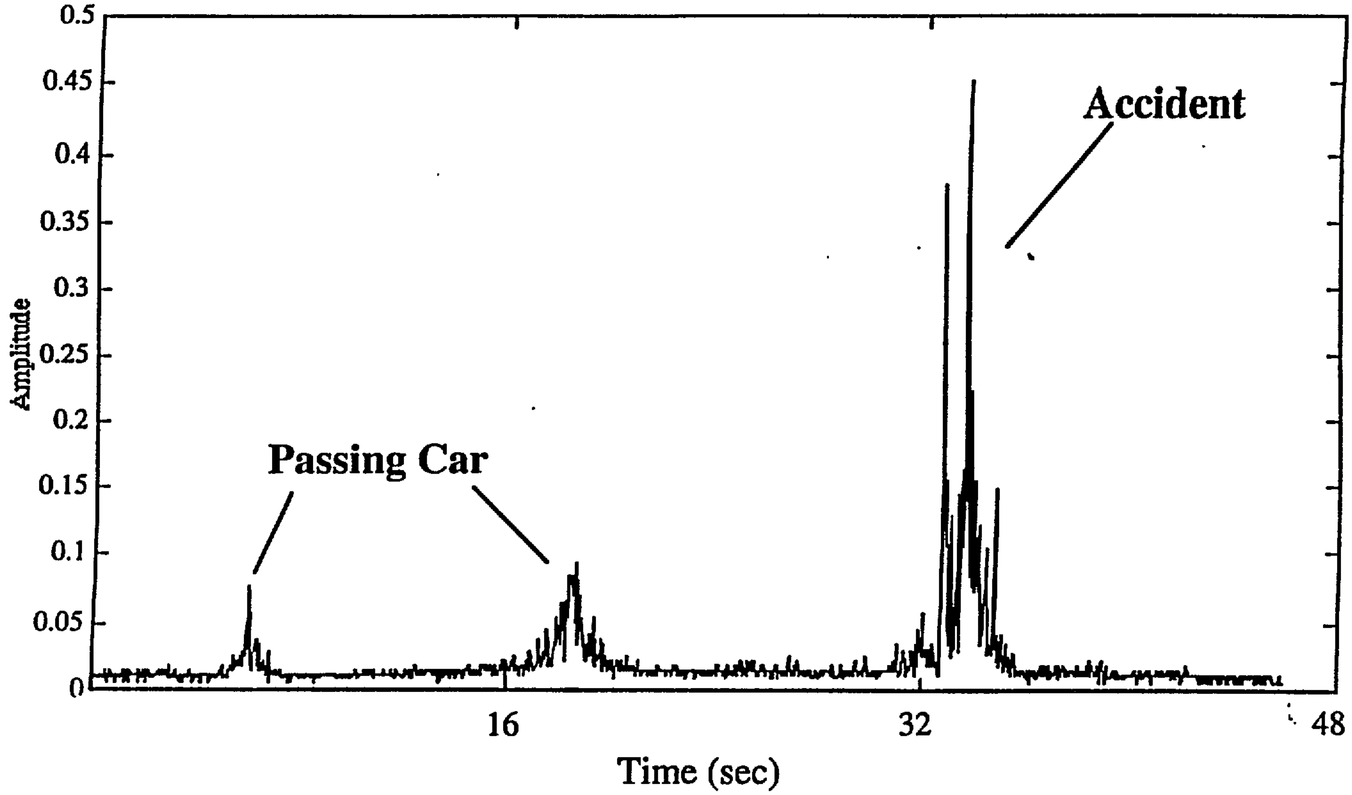


Figure 3-1 Case 1: Amplitude Signature

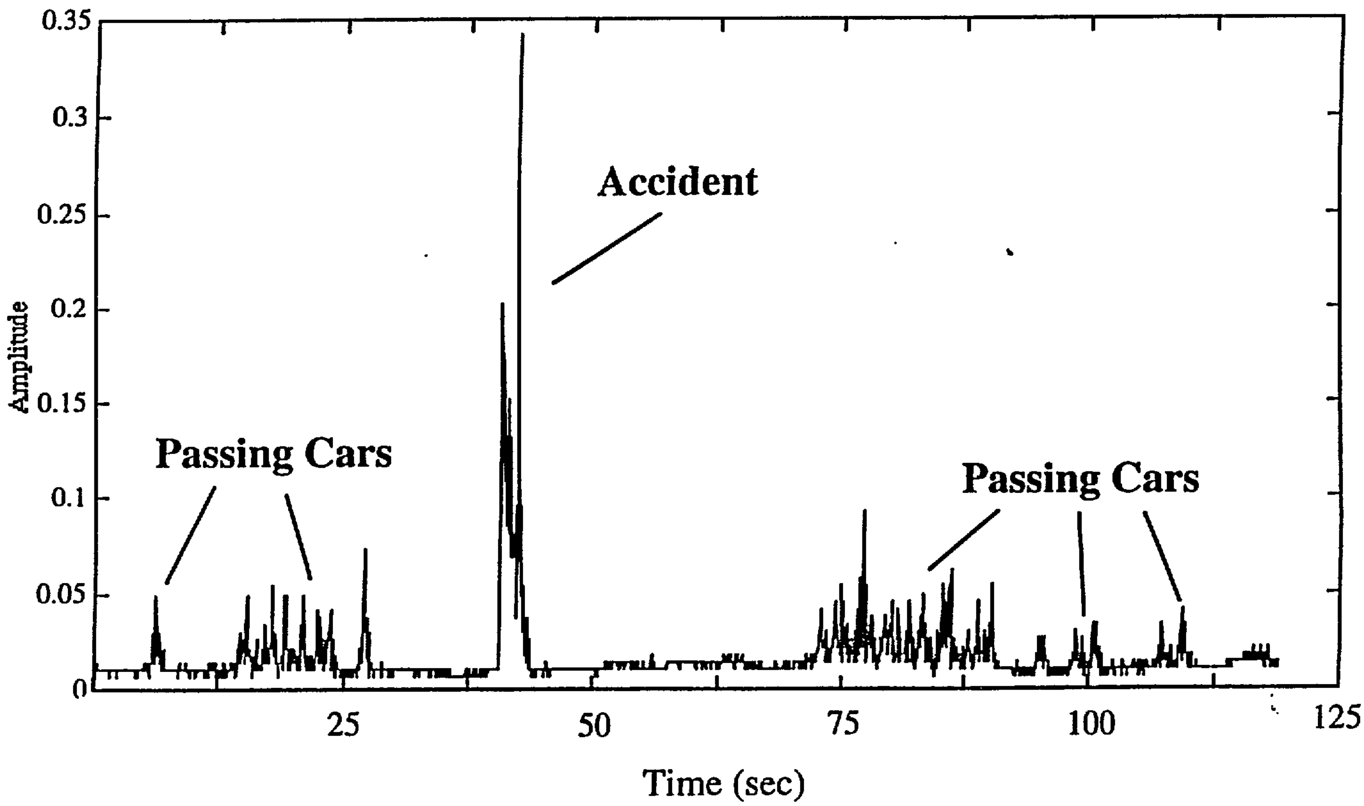


Figure 3-2 Case 2: Amplitude Signature

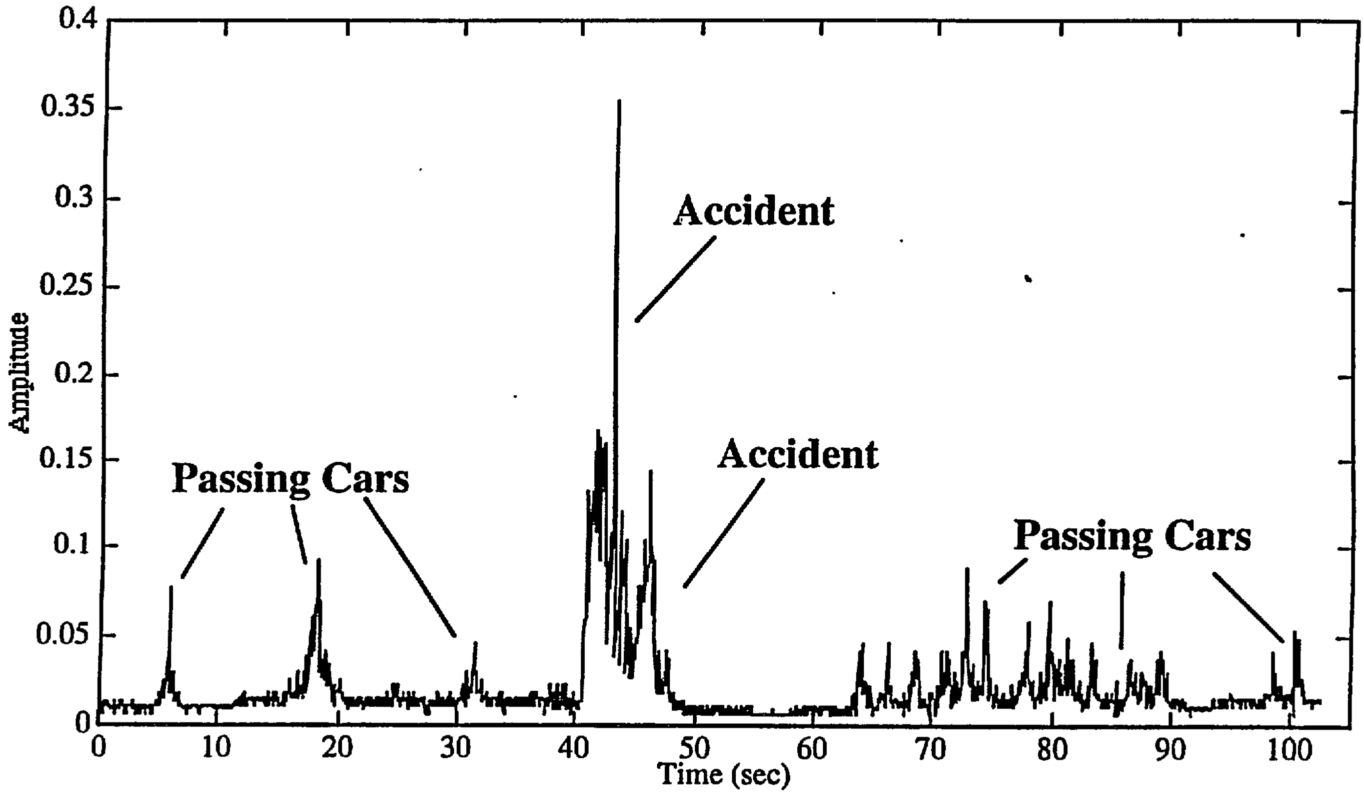


Figure 3-3 Case 3: Amplitude Signature

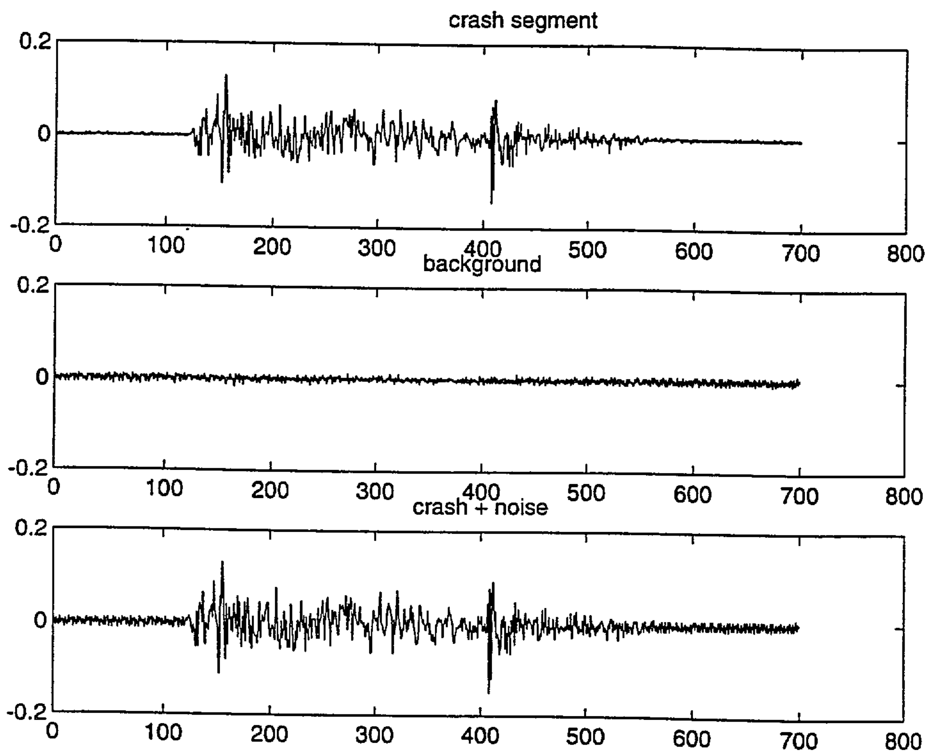


Figure 3-4 Case 2a: Scaled Accident Segment and Traffic Background

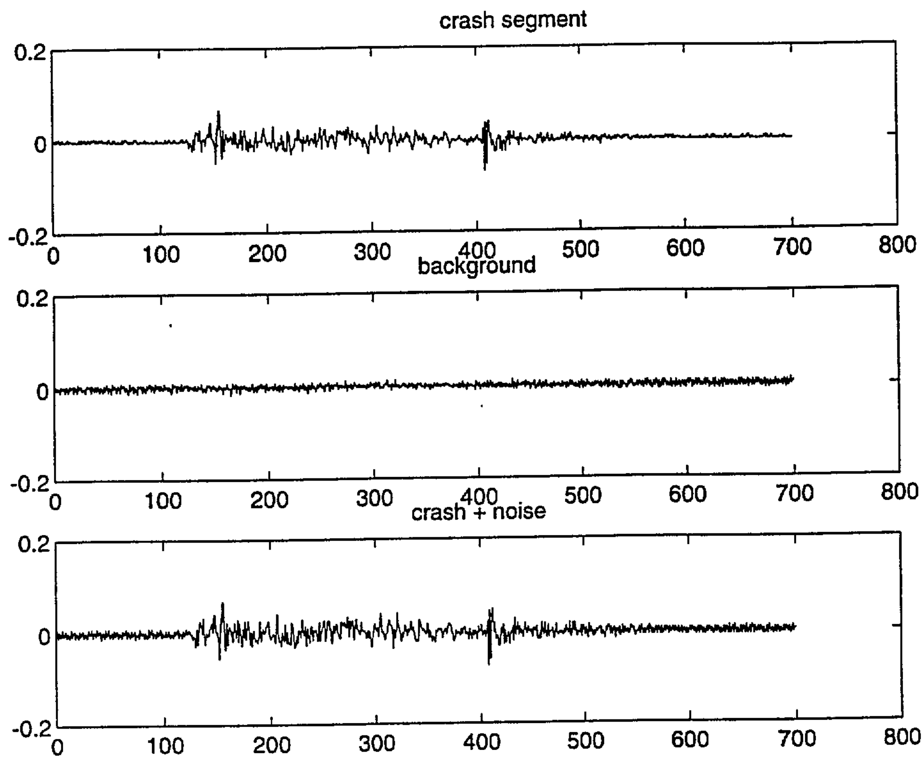


Figure 3-5 Case 2b: Scaled Accident Segment and Traffic Background

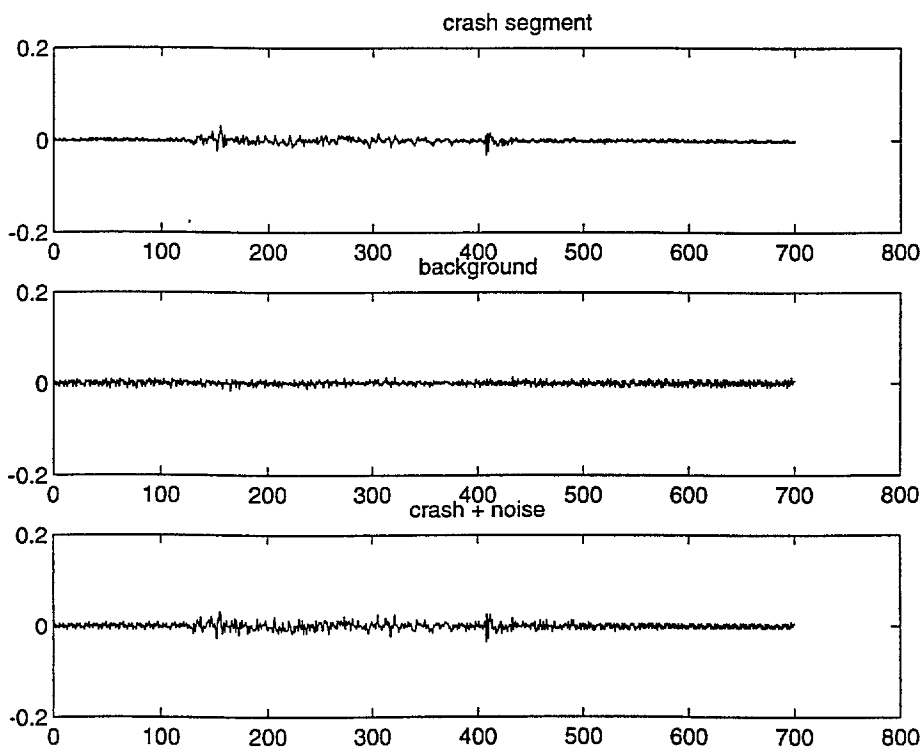


Figure 3-6 Case 2c: Scaled Accident Segment and Traffic Background

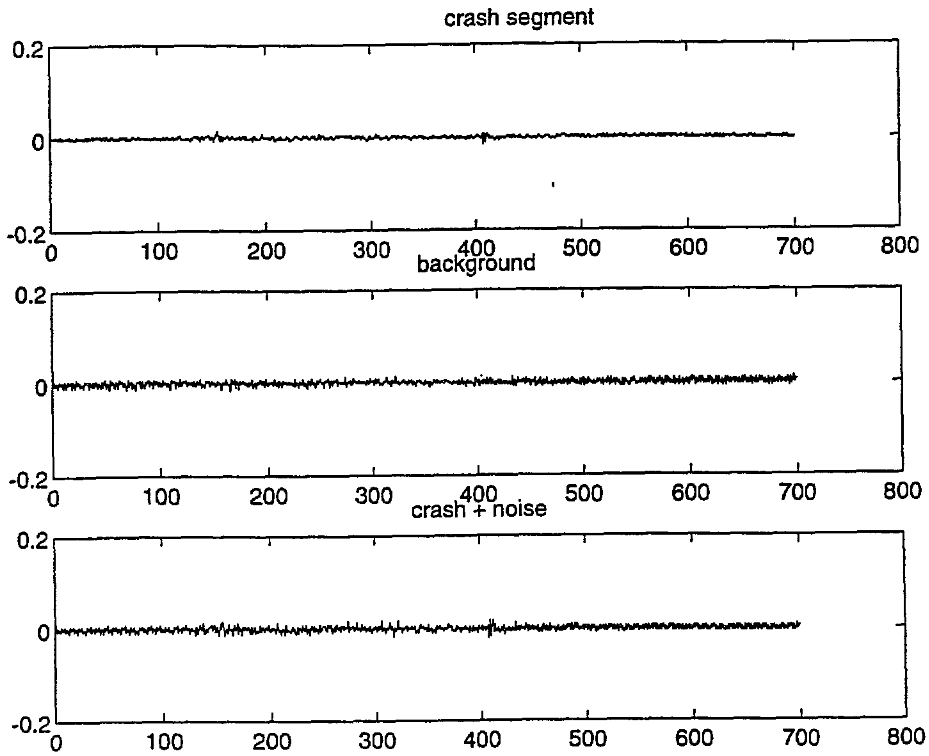


Figure 3-7 Case 2d: Scaled Accident Segment and Traffic Background

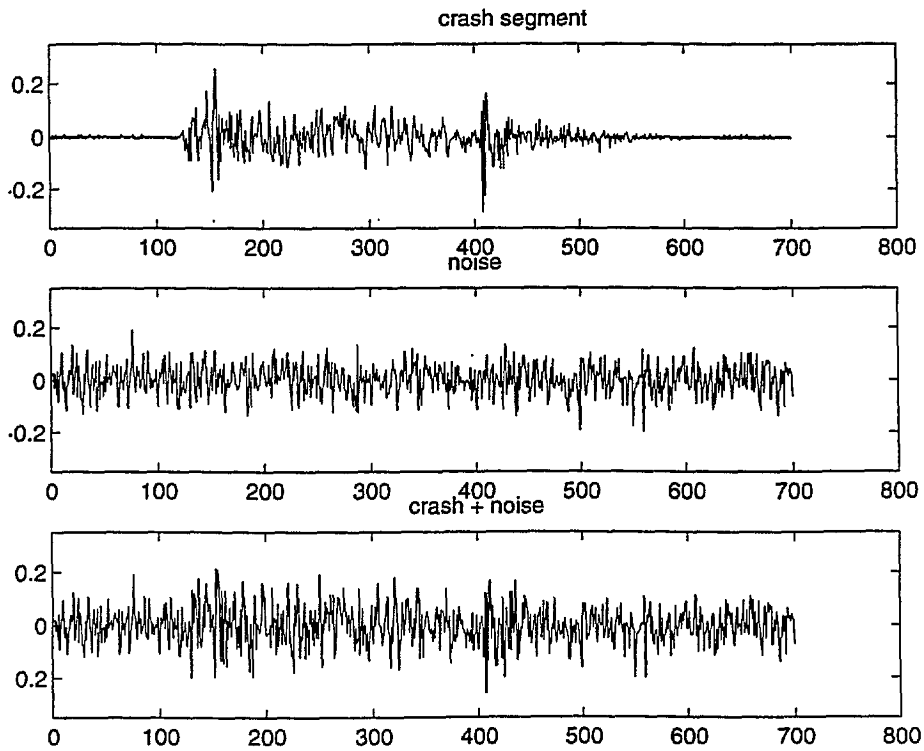


Figure 3-8 Case 2n: Scaled Accident Segment and Added Noise

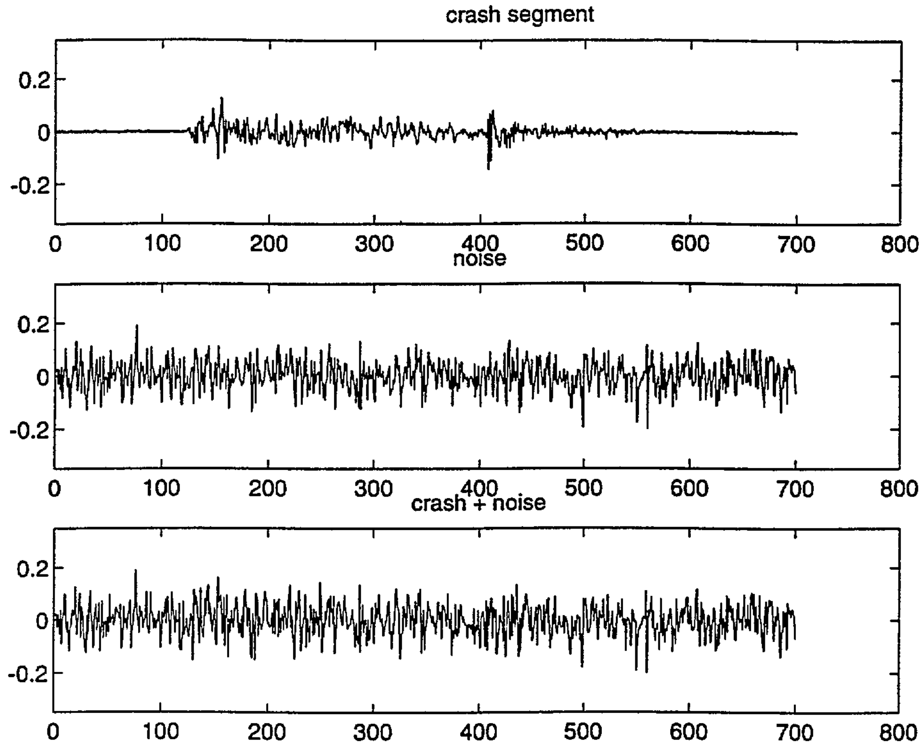


Figure 3-9 Case 2an: Scaled Accident Segment and Added Noise

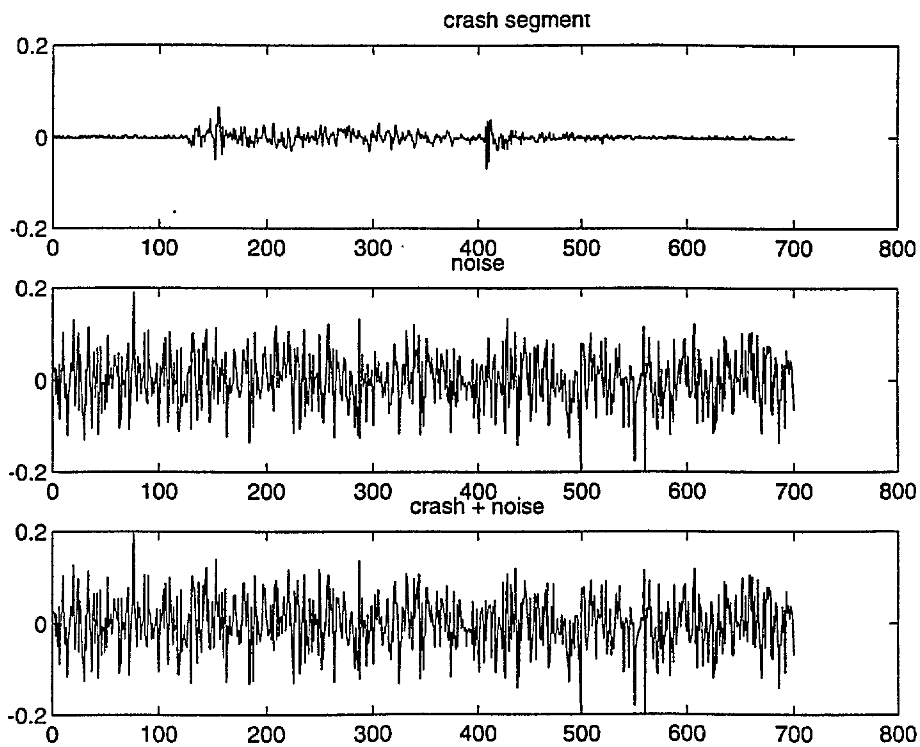


Figure 3-10 Case 2bn: Scaled Accident Segment and Added Noise

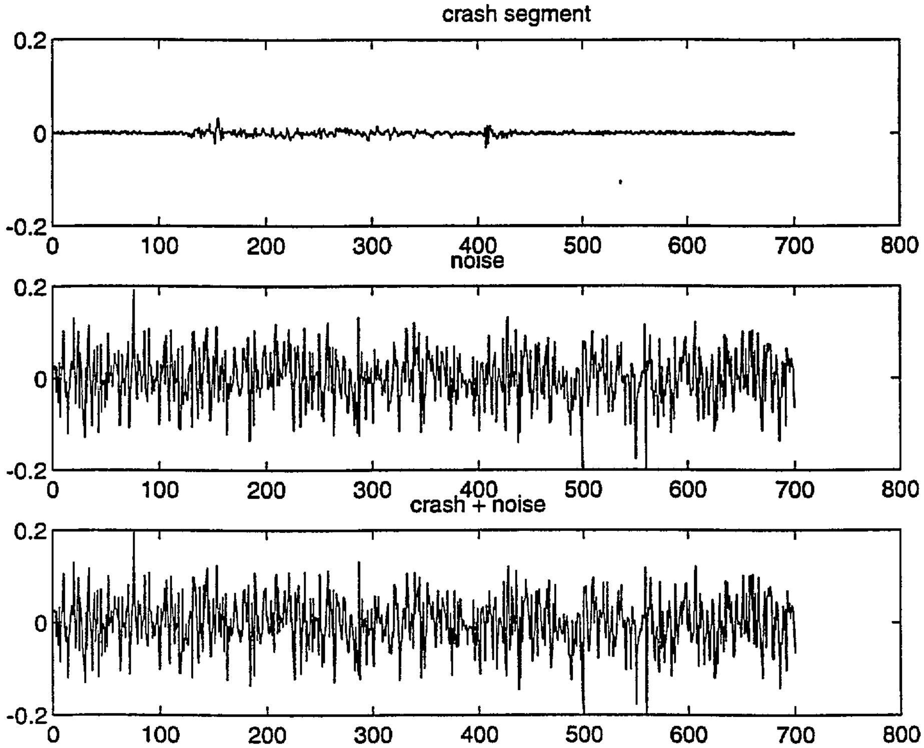


Figure 3-11 Case 2cn: Scaled Accident Segment and Added Noise

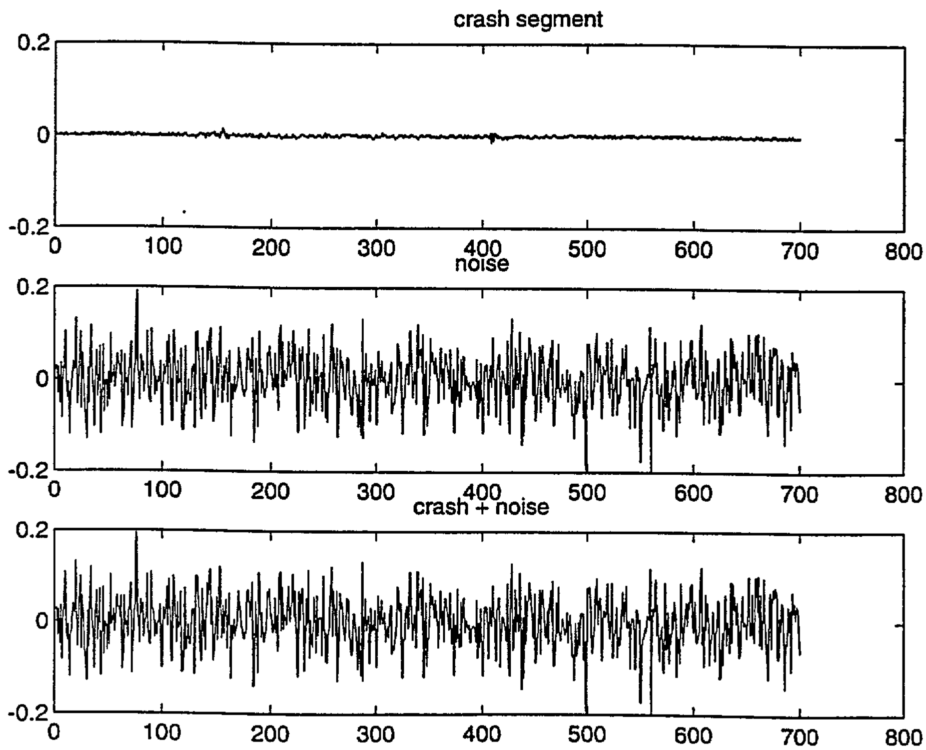


Figure 3-12 Case 2dn: Scaled Accident Segment and Traffic Background

4. ACOUSTIC INPUT DATA CHARACTERIZATION

4.1 Analysis of Incident and Background Features

As a precursor to designing the detection algorithm and the acoustic database, the incident and background acoustic data was characterized according to three general signal features: wave form amplitude; central statistical moments (e.g. mean, kurtosis) ; and power spectral density. The purpose of this characterization was to identify individual features or combinations of features that could be used to distinguish an incident from the background and possibly even one type of incident from another. Identifying the differing features of vehicle crash sounds and passing of vehicle quantitatively is an important step toward building a Hyperstate model to describe the dynamics of those features over time. To do this a combination of features was reviewed. The following section outlines the signal feature characterization and analysis.

The raw digital data from both sources is stored as sixteen bit wave forms sampled at 44.1 kHz. This high resolution data is down sampled to 16 kHz before processing to reduce the computational requirements, and because all of the significant signal features could be observed at frequencies less than 8 kHz (determined by the Nyquist theorem). Individual events are then extracted from the down-sampled data and stored in separate files. The three types of signal features are derived from each event file.

4.2 Crash Data Features

For the first step of Hyperstate analysis, spectral characteristics were chosen as the features to use to define the states of the Hyperstate models. This is because of the strong, consistent incident "signature" produced by the crash events in out database. In future refinements of the baseline algorithm architecture, additional features described in Section X will be considered as well. Figure 4-l shows a plot of frequency intervals in which significant energy was present at some time around a crash, for several representative crashes in the database. Looking down the graph vertically at any interval, we can see that several intervals are common to most crashes. We have selected four of those intervals (spectral bands) for analysis in the prototype. They are:

- 100-400 Hz
- 800-l 140 Hz
- 1560-1760 Hz
- **2200-2700** Hz

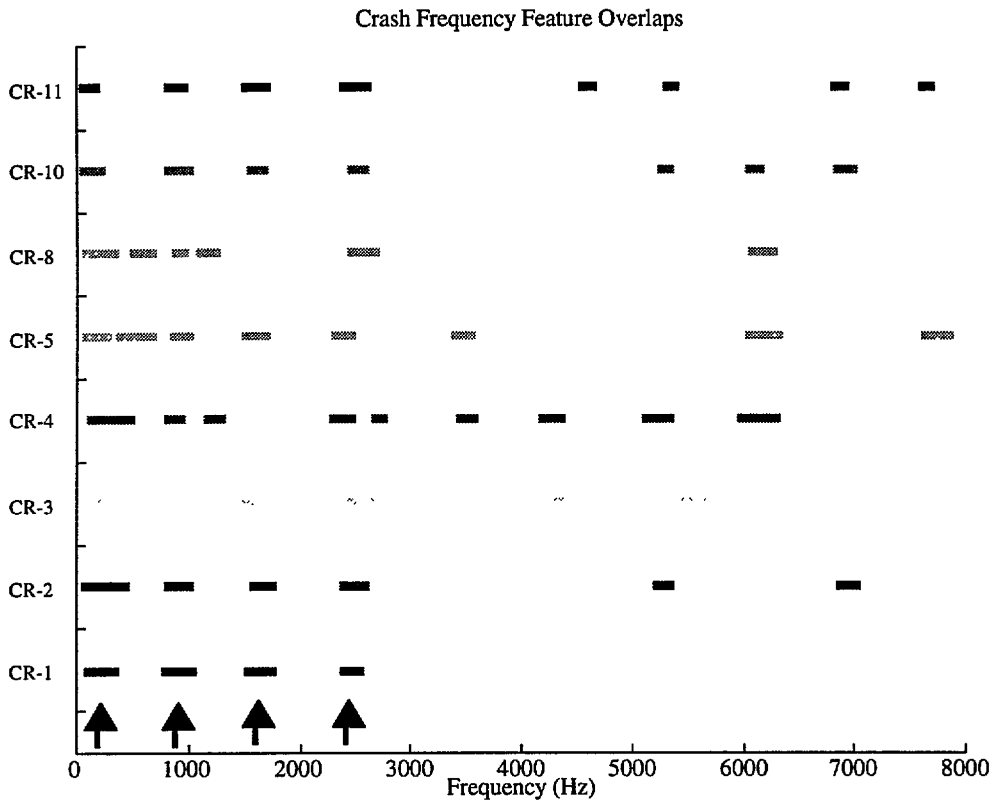


Figure 4-1 Identifying Common Crash Spectral Features

Appendix A of the Interim Report [Ref. 1] contains additional representative plots depicting a wide range of features of illustrative accident and background data sets.

5. OPTIMIZED ALGORITHM ARCHITECTURE

This section describes the optimized AutoAlert algorithm that was implemented for analysis of operational data. For a detailed discussion of the general form and mathematical theory underlying the multiresolution Hyperstate Algorithms used to implement AutoAlert, see the description in the AutoAlert project's Interim Report [Ref. 1].

5.1 Multiresolution Algorithm Overview

The core algorithms for the AutoAlert processor belong to a unique class of *Hyperstate* algorithms *that have already been developed* by TASC. They have been evaluated on acoustic data in other application areas, such as underwater transient detection and vehicle sound quality analysis. The primary objectives of this concept feasibility investigation are to tailor the algorithms for AutoAlert's incident detection for ITS systems, and to rigorously evaluate their performance.

The Hyperstate algorithms to be used in the AutoAlert processor provide new levels of performance in "acoustic fingerprinting." This is achieved by using a stochastic, model-based procedure that can analyze the dynamics of a time-varying process on multiple time scales. Models corresponding to persistent acoustic feature patterns for incidents and non-incidents are developed, stored and used to scan for incidents. Models that characterize the changing traffic background level are computed adaptively, in real time.

A hierarchy of Hidden Markov Models (HMM), in conjunction with adaptive Canonical Variates Analysis (CVA), is the mathematical framework that comprises the Hyperstate procedure. A family of these HMM models, operating on multiple time scales, is used to classify complex patterns of time-varying acoustic features (such as frequency dynamics) that differentiate traffic background noise from sounds associated with incidents. Such models have been applied recently to complex acoustic analysis tasks, such as speech recognition.

5.2 Algorithm Architecture

The general AutoAlert processing architecture is depicted in Figure 5-l. Analysis of operational field test data shows that a particular set of five feature identification filters and two temporal resolutions are appropriate for optimized data processing. Figure 5-2 illustrates the architecture used for data processing.

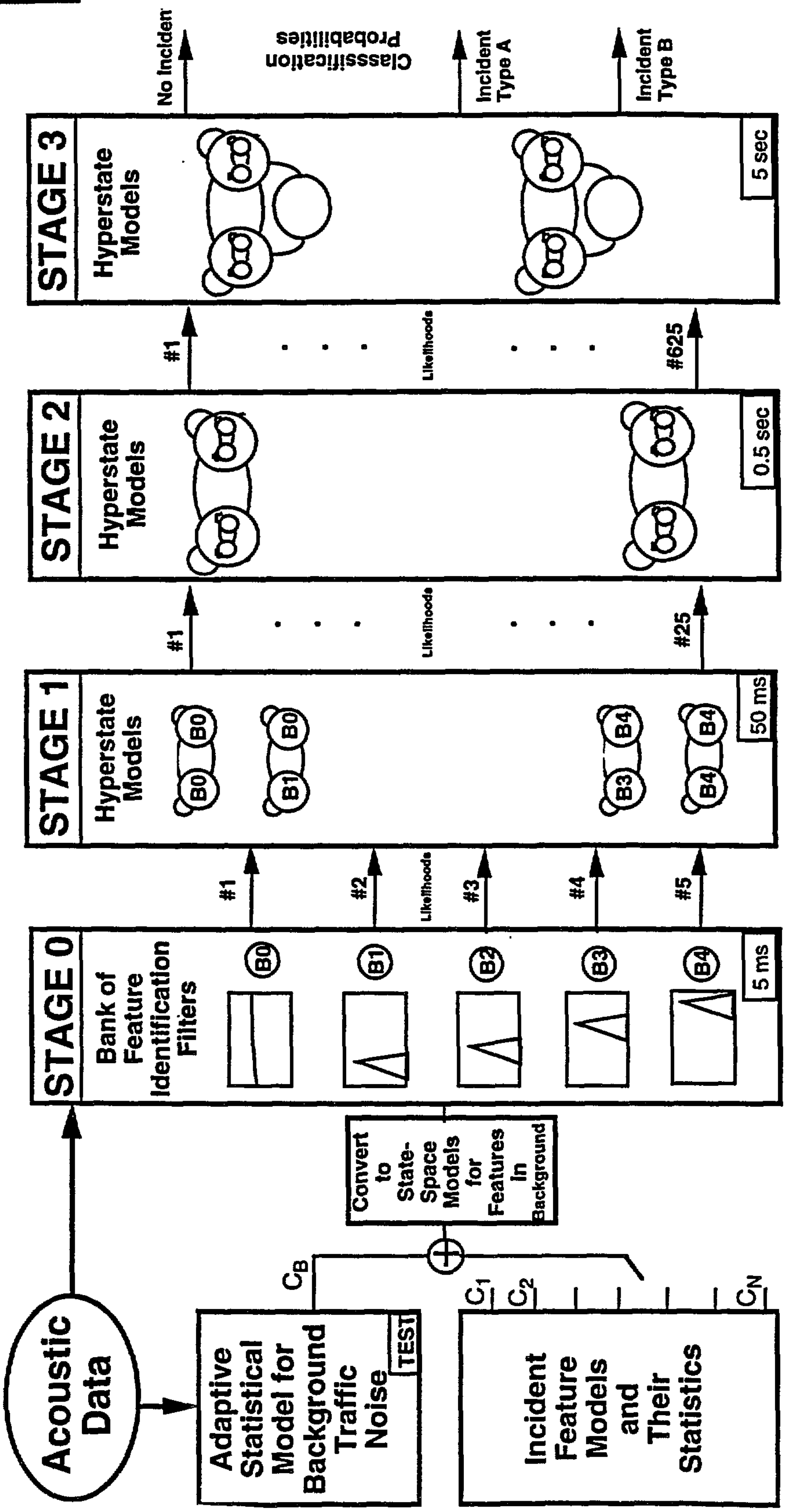


Figure 5-1 General Hyperstate Hierachy

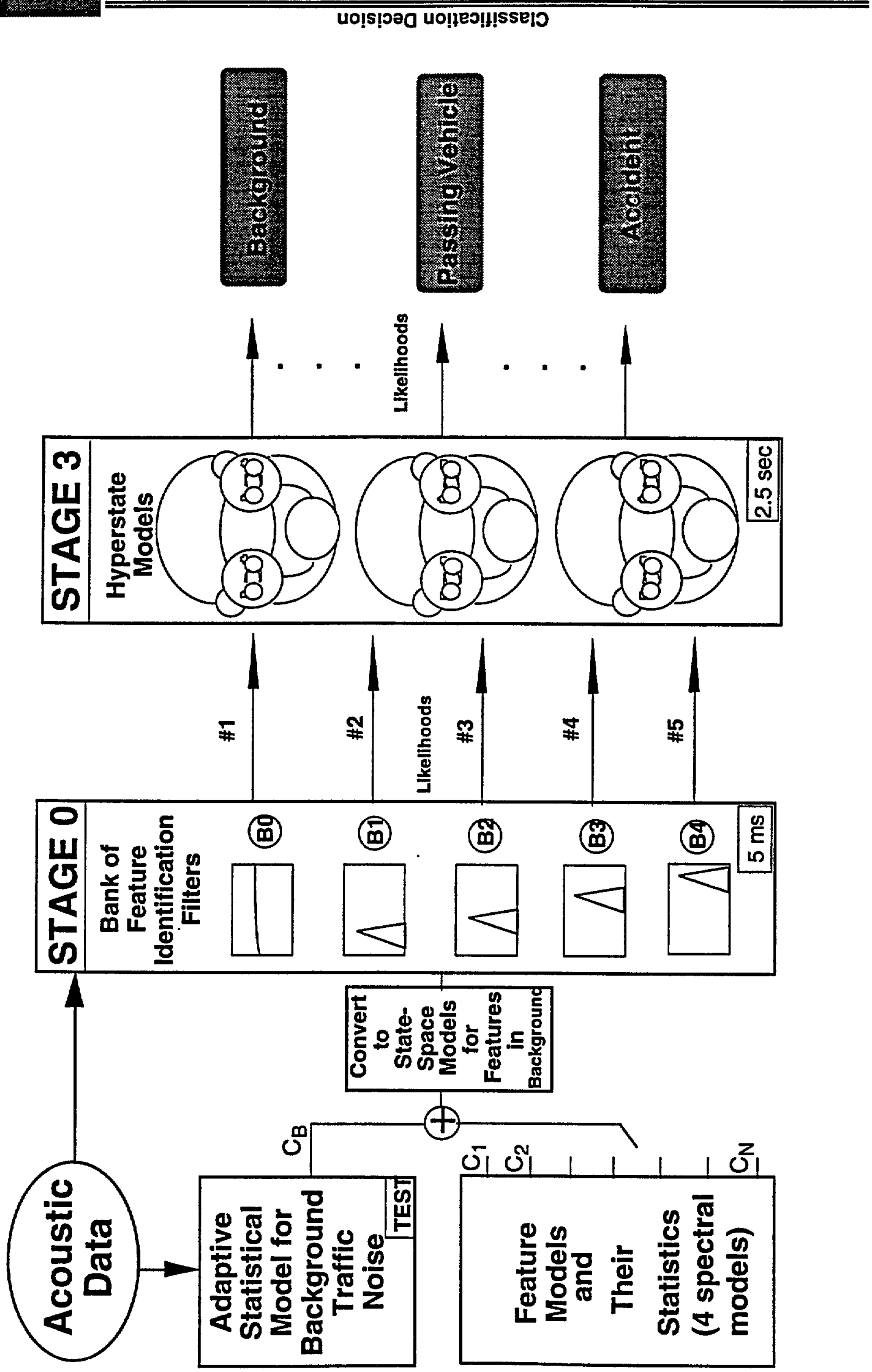


Figure 5-2 Optimized AutoAlert Hierarchy

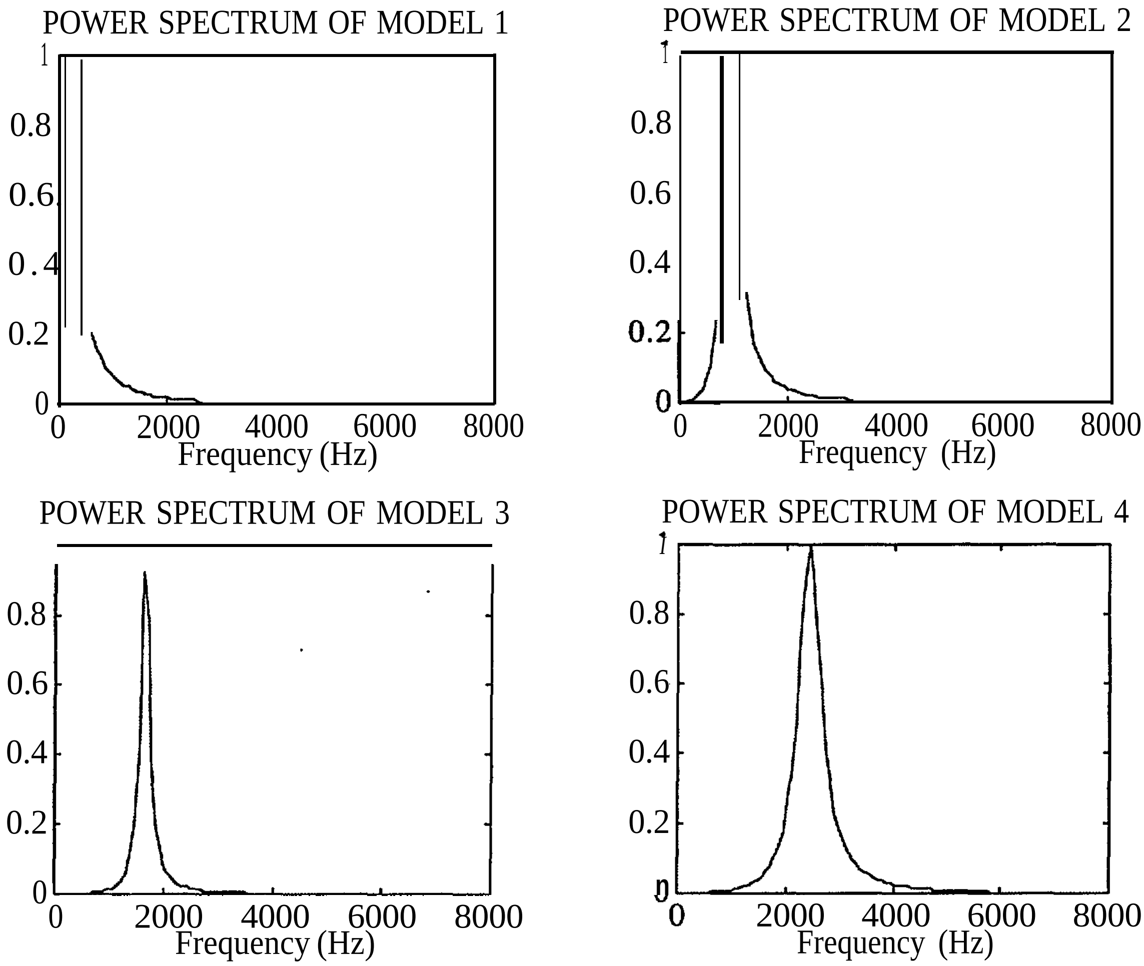
Background Modeling -- Acoustic data is processed along two paths. The first involves adaptive modeling of the background noise in the traffic scene, such as rainfall or traffic sounds from distant roads. Canonical variates state-space modeling is used to compute a new background model every 0.1 sec. A statistical measure of similarity between the current and previously computed background model, called the Discriminant Information, is computed to decide if the background is slowly varying or changing rapidly. If the background is slowly varying, the latest model is adopted. If the background model has changed significantly over the last 0.1 sec, this could indicate that a vehicle has come into range of the sensor, and the background model is not adopted. This allows slow evolution of the background to be tracked, while preventing acoustic information from passing vehicles or crashes from being put into the background model.

Spectral Feature Library -- In order to define the identification filters for the Stage 0 Hyperstate processing illustrated in Figure 5-2, filters tuned to each of the four frequency bands identified in Section 4.2 and Figure 4-l are implemented. The filters are implemented using state-space whitening filters with frequency responses given by the spectrum plots in Figure 5-3. These four filters are combined with an all-pass filter (uniform frequency response) to produce the five labeling filters. These filters are designed to have maximum response when data in their respective pass-bands are processed. Since these pass-bands were selected to match accident features, they are optimized for this detection.

The adaptive background model which is continuously computed by AutoAlert is used by AutoAlert as a “subtraction” filter (‘whitener”) to remove background effects while the spectral features are searched for. This filtering must be accounted for in defining the Feature Model Filters, causing a slight transformation of the original frequency response of the baseline feature library models. This is illustrated in Figure 5-4. The likelihoods for each of the combined feature models are computed every 0.005 sec, in order to allow rapidly varying and transient acoustic effects to be captured.

Hyperstate Modeling -- The final stage of the optimized architecture involves multiresolution feature modeling using Hyperstate models. Three Hyperstate models were developed, one representing the dynamics on a 2.5 second time scale for each of the categories: background (no accidents or vehicles), passing vehicles, and accidents. Five hundred likelihoods from the feature identification stage are processed every 2.5

seconds, and each 2.5 seconds the Hyperstate analysis identifies which of the three classification models is most likely, using a maximum likelihood criterion.



- **Model 5 is an all-pass (white noise) model**

Figure 5-3 Spectral Feature Library for AutoAlert Prototype

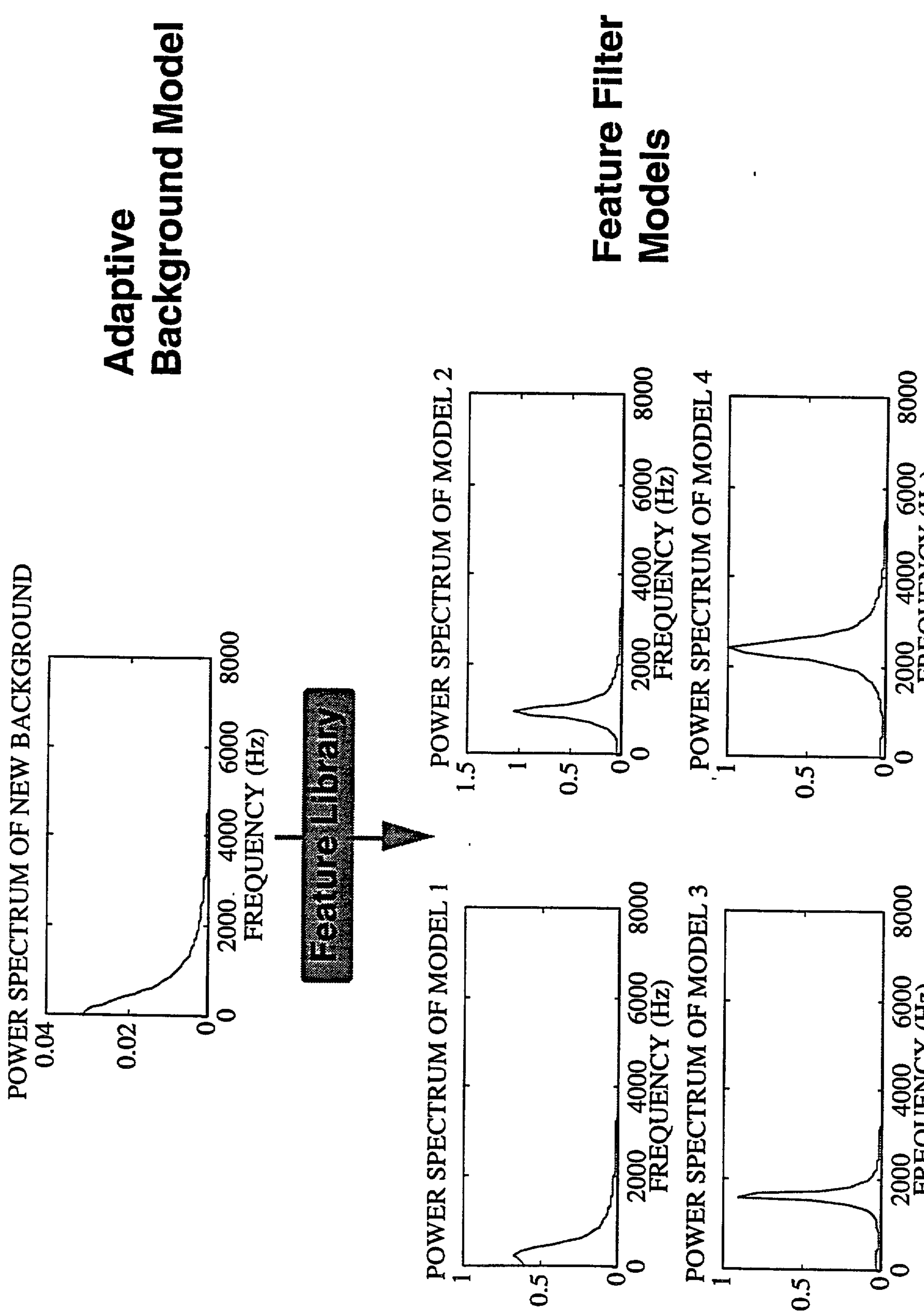


Figure 5-4 Combined Background and Feature Models

6. DATA PROCESSING RESULTS

6.1 Summary

Detecting Accidents – Using the prototype AutoAlert algorithm described above, the data cases described in Table 3-l were analyzed, with the result summarized in Table 6-l. As the table indicates, the AutoAlert algorithms do well at detecting the accidents for a range of accident energy ratios, all the way down to accident scale factors of l/4, for both the cases with and without added noise. This reflects good real-world performance, since scale factors of l/8 and lower correspond to accidents that occur that are quieter than the passing vehicles on the highway. For low signal-to-noise ratios scale factors of l/8 and l/16) in the absence of added low-pass noise, accidents are detected, but are classified as passing vehicles rather than as accidents.

Table 6-l Summary of Classification Performance

Case	Scale Factor	Noise Added	Accident Energy Ratio	Accident Identified	Passing Veh. Identified
1	1	no	-	✓	✓
2	1	no	8.8	✓	✓
3	1	no	-	✓	✓
2a	1/2	no	4.4	✓	✓
2b	1/4	no	2.2	✓	✓
2c	1/8	no	1.1		✓
2d	1/16	no	0.55		✓
2n	1	yes	0.9	✓	
2an	1/2	yes	0.45	✓	
2bn	1/4	yes	0.23	✓	
2cn	l/8	yes	0.11		
2dn	l/l6	yes	0.055		

Detecting Passing Vehicles -- In addition to accident detection, the AutoAlert classifier was also designed to detect passing cars and differentiate them from accidents. The algorithm did this successfully for scale factors down to l/l6 in the absence of lowpass noise. As shown in the next section, both single and multiple passing vehicles were detected and differentiated from accidents and the background.

Effect of Lowpass Noise -- This noise was added to stress the performance of the AutoAlert algorithm and evaluate how it would perform in the presence of other significant distracting sounds, such as rainfall. Table 6-l indicates that the addition of this noise did not hinder accident detection, but it did make passing vehicles more difficult to detect. This is due to the added low-frequency noise (in the frequency range from O-500 Hz) which shares characteristics in common with the features and feature patterns of the passing cars, and effectively “masks” them from the detector. Because the accident signature was composed of a more complex collection and pattern of frequency features, it was detectable even in the presence of the added low-pass noise.

6.2 Detailed Analysis Results

This section presents the detailed data processing results that were used to develop Table 6- 1, including direct analysis of the Hyperstate algorithm processing.

6.2.1 Varying Backgrounds and Loud Accidents – This analysis includes a discussion of how to interpret the Hyperstate/AutoAlert intermediate data displays. Figure 6-l shows the amplitude plot of the accident and passing vehicles for Case 1. Figure 6-2 shows the first stage Hyperstate output data. Each of the five Feature identification filters is listed on the vertical axis. Time is on the horizontal access. The gray arrow indicates where the accident occurred in the data. Every 5 msec, a circle is put on the graph corresponding to the most likely of the five feature models. The distinct pattern associated with the crash, exciting all five feature models with a few seconds in a specific pattern, is clearly visible.

Figure 6-3 shows the output of the AutoAlert classifier. Each of the three possible classification decisions, “Crash”, “Passing Vehicle”, or “Background” is listed on the vertical axis. Every 2.5 seconds, AutoAlert makes a decision as to which of the three event types is most likely. That sequence of decisions is shown by the circles on the plot. AutoAlert correctly identifies the accident and the passing vehicles, distinguishing both from the background.

Detection results from Case 2, involving multiple passing cars, is shown in Figures 6-4 through 6-6. The accident, whose time of occurrence is indicated by the gray arrow in the figures, is correctly detected, as well as the passing vehicles. Periods of vehicle-free background are also identified correctly.

Similar good detection and event differentiation performance is also seen in the Case 3 results presented in Figures 6-7 through 6-9. In this case a two car accident (collision) occurs. AutoAlert correctly identifies the event type and span of the two accidents. Because the acoustic signatures of the two accidents overlap in time (one event is still being completed while the other is beginning) and have a separation that is smaller than the 2.5 sec resolution of the classifier--there is no gap in the detection of the two accidents.

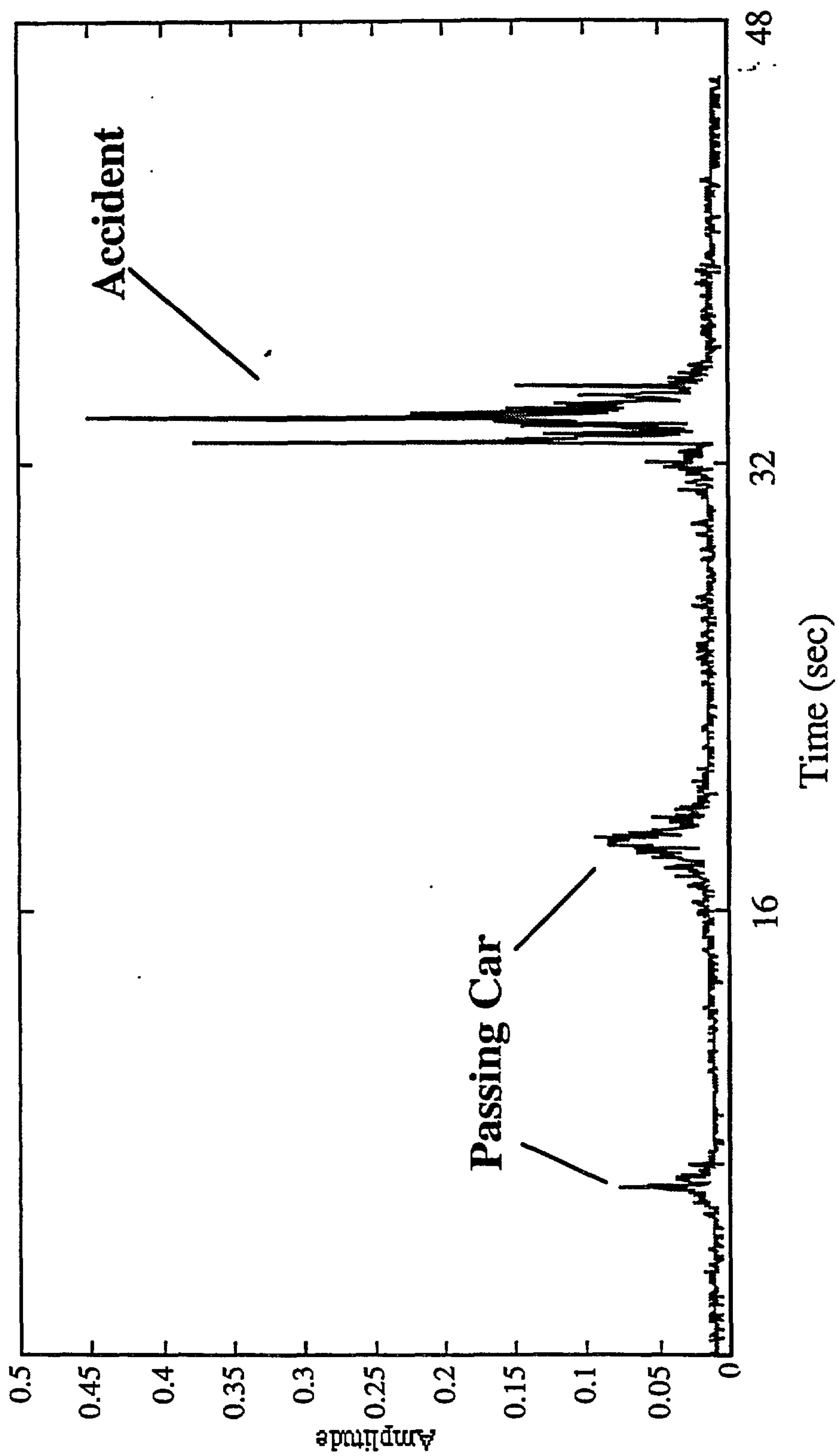


Figure 6-1 Case 1: Amplitude

The ITS-IDEA program is jointly funded by the U.S. Department of Transportation's Federal Highway Administration, National Highway Traffic Safety Administration, and Federal Railroad Administration. For information on the IDEA Program contact Dr. K. Thirumalai, IDEA Program Manager, Transportation Research Board, 2101 Constitution Avenue N.W., Washington, DC 20418 (phone 202-334-3568, fax 202-334-3471).

IDEA PROJECT FINAL REPORT
Contract ITS-19

IDEA Program
Transportation Research Board
National Research Council

December 26, 1995

AutoAlert: Automated Acoustic Detection of Incidents

David A. Whitney and Joseph J. Pisano
TASC, Inc., Reading, Massachusetts

Approved by: Robert E. Introne

INNOVATIONS DESERVING EXPLORATORY ANALYSIS (IDEA) PROGRAMS MANAGED BY THE TRANSPORTATION RESEARCH BOARD (TRB)

This investigation was completed as part of the ITS-IDEA Program, which is one of three IDEA programs managed by the Transportation Research Board (TRB) to foster innovations in surface transportation. It focuses on products and results for the development and deployment of intelligent transportation systems (ITS), in support of the U.S. Department of Transportation's national ITS program plan. The other two IDEA programs areas are TRANSIT-IDEA, which focuses on products and results for transit practice in support of the Transit Cooperative Research Program (TCRP), and NCHRP-IDEA, which focuses on products and results for highway construction, operation, and maintenance in support of the National Cooperative Highway Research Program (NCHRP). The three IDEA program areas are integrated to achieve the development and testing of nontraditional and innovative concepts, methods, and technologies, including conversion technologies from the defense, aerospace, computer, and communication sectors that are new to highway, transit, intelligent, and intermodal surface transportation systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Project Overview

AutoAlert applies new signal processing algorithms to passive acoustic data to advance the state of practical acoustic incident detection techniques. These techniques, originally developed for national defense applications, will perform reliable, automatic, nearly instantaneous, all-weather incident detection under highly variable traffic conditions. Effective operation of urban high-capacity ITS systems requires speedy detection of incidents at chokepoints, such as tunnels, bridges and other aerial structures, and dense urban arterials. Boston's Central Artery/Tunnel (CA/T) project is an example of a new ITS system where such detection is critical. AutoAlert overcomes shortcomings of loop and video detectors, such as their inability to distinguish between incidents and congestion, and the need for a human-in-the-loop for video detection. The AutoAlert processor "hears" an incident *before* congestion builds, and can be used either as an independent detector, or its outputs can be combined (data fusion) with other detector outputs for joint improved decisions and incident verification.

The problem of rapid, reliable acoustic incident detection is more complex and difficult than the problems of freeway traffic flow monitoring or vehicle type identification, for which acoustic sensors have already been applied. The AutoAlert processor (see Fig. 1-1) will make use of readily available commercial acoustic *sensors* (e.g., AT&T IVHS NET—2000™). **What is key to AutoAlert is the signal analysis and detection *algorithms* that we will employ**. The algorithms will provide a new level of incident detection timeliness and reliability (low false alarms) by applying sophisticated statistical models: Hidden Markov Models (HMM) and Canonical Variates Analysis (CVA). These are used to analyze both short-term and time-varying signals that characterize incidents.

1.2 Principles of Innovation

AutoAlert algorithms provide performance innovations that offer significant improvements over today's most "advanced" acoustic traffic sensor systems:

- Provide *nearly immediate detection*, with no alertment delays, before congestion builds
- React to an incident *directly,* not just the *symptom* of an incident (e.g., congestion)
- Provide *low false alarm rates* via simultaneous analysis of short—, medium—, and long—term acoustic feature patterns

6.2.2 Accidents of Varying Loudness in Traffic

Cases 2a through 2d are identical to Case 2, except that the accident signature in Case 2 was progressively reduced by a factor of l/2 while passing vehicle traffic remained unchanged. Figures 3-4 through 3-7 show the segment of each dataset that contains the accident and the relative amplitudes of the scaled signatures, plus the local background with which they are combined to form the input signal to AutoAlert. Figures 6-10 through 6-13 show the results of the AutoAlert feature labeling step in the top half of each figure, and AutoAlert classification results in the bottom half of each figure.

For Cases 2a and 2b (with accident scale factors of l/2 and l/4 respectively), the accident are detected and correctly identified. For Cases 2c and 2d (with accident scale factors of l/8 and l/16 respectively), the accidents are detected, but are classified as passing cars rather than accidents. These missed identifications are due to the low signal strength (accidents are quieter than the passing cars) associated with these cases. For all cases, the passing vehicles are correctly identified.

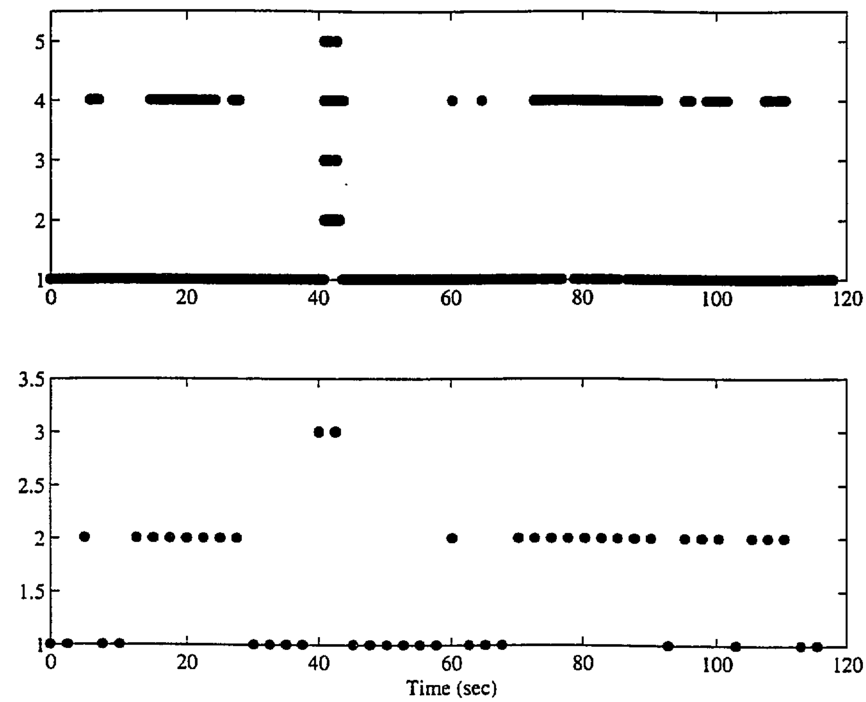


Figure 6-10 Case 2a: Feature Labeling and Classifier Output

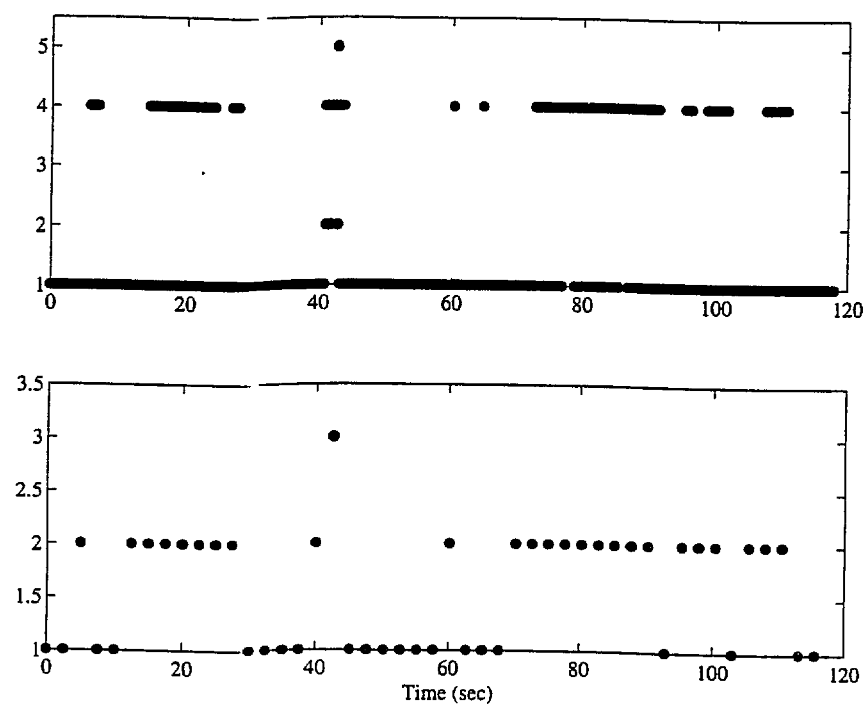


Figure 6-11 Case 2b: Feature Labeling and Classifier Output

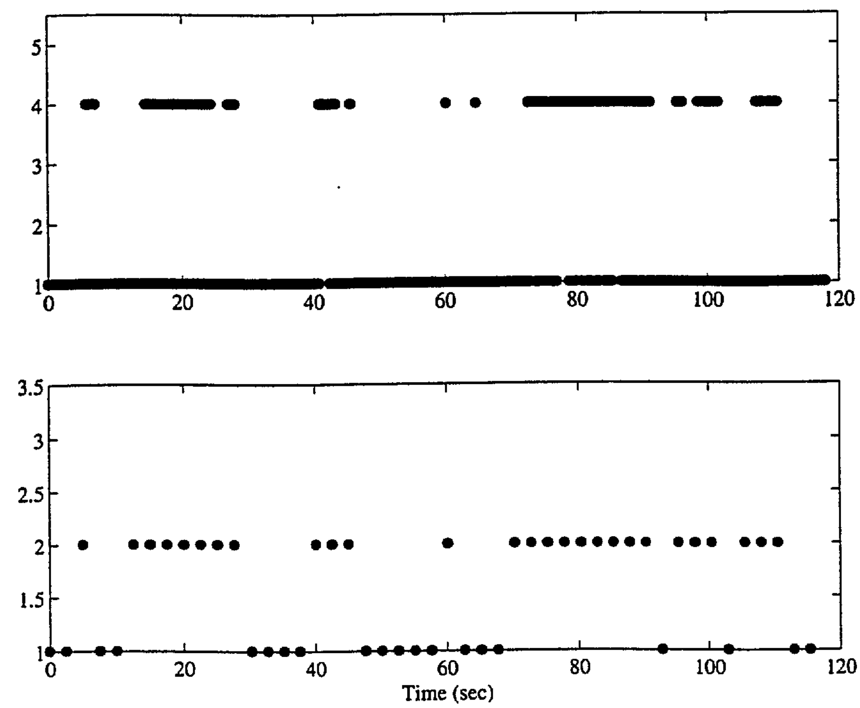


Figure 6-12 Case 2c: Feature Labeling and Classifier Output

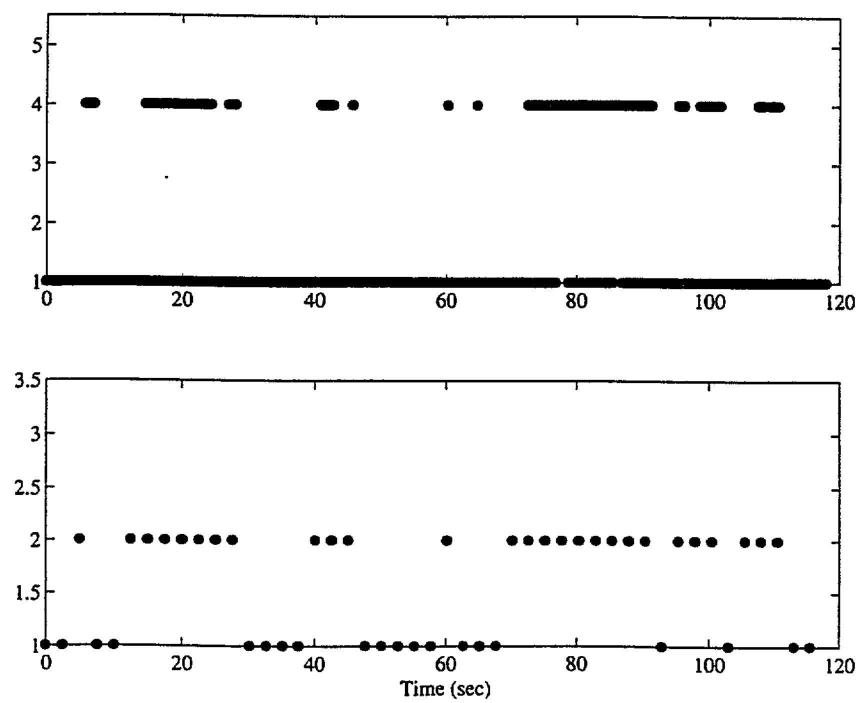


Figure 6-13 Case 2d: Feature Labeling and Classifier Output

6.2.3 Accidents of Varying Loudness in Traffic and Noise

Cases 2n and 2an through 2dn are identical to Cases 2 and 2a through 2d, except that additional low-frequency noise was added uniformly to the acoustic time history, as described in Section 3. Figures 3-9 through 3-l 1 show the segment of each dataset that contains the accident and the relative amplitudes of the scaled signatures, plus the added noise with which they are combined to form the input signal to AutoAlert. The amplitude of the noise is large relative that of the accident, even for the first signature scale factor of l/2 Figures 6-14 through 6- 18 show the results of the AutoAlert feature labeling step in the top half of each figure, and AutoAlert classification results in the bottom half of each figure.

The results parallel those obtained for the case described in Section 6.2.2 where no additional low-frequency noise was added. For Cases 2n, 2an and 2bn (with accident scale factors of l/l, l/2 and l/4 respectively), the accident are detected and correctly identified. For Case 2cn and 2dn (with accident scale factors of l/8 and l/l 6), the accidents are not detected. These missed identifications are due to the low signal strength (accidents are quieter than the passing cars) associated with these cases. For all of these, cases, the passing vehicles are not detected. This is due to the added low-frequency noise which shares characteristics in common with the features and feature patterns of the passing cars, and effectively "masks" them from the detector.

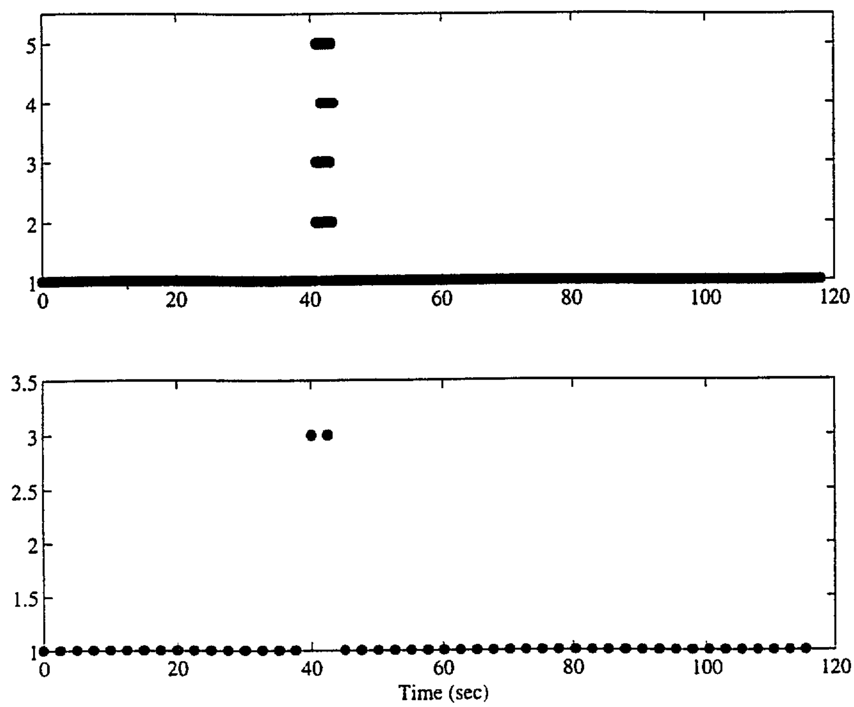


Figure 6-14 Case 2n: Feature Labeling and Classifier Output

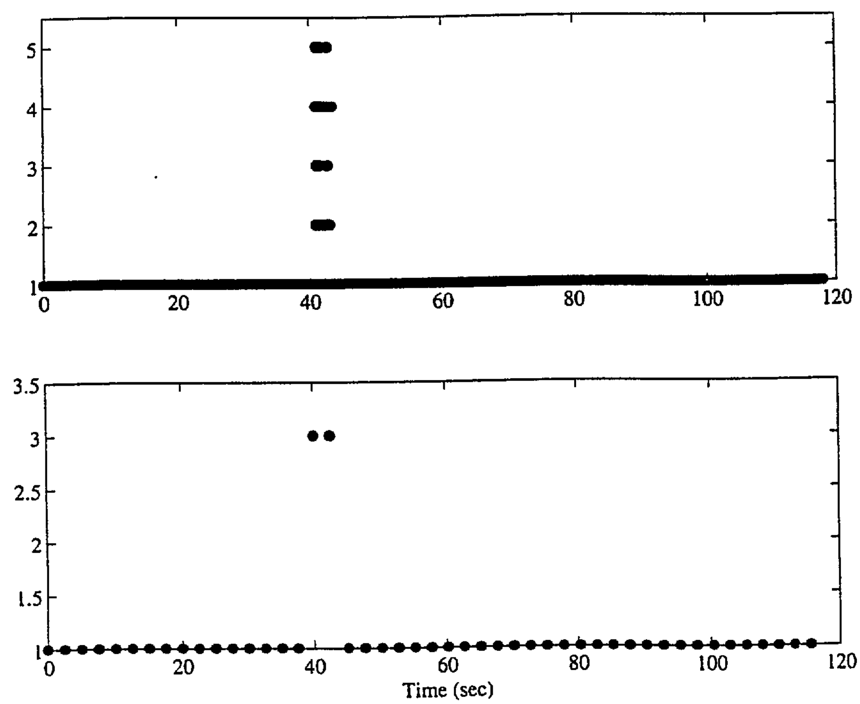


Figure 6-15 Case 2an: Feature Labeling and Classifier Output

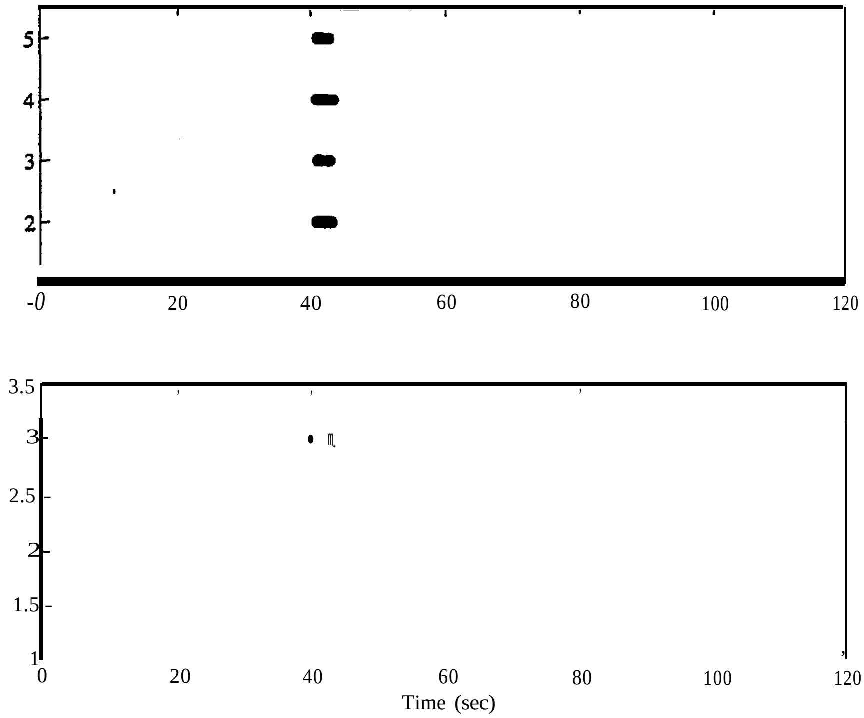


Figure 6-16 Case 2bn: Feature Labeling and Classifier Output

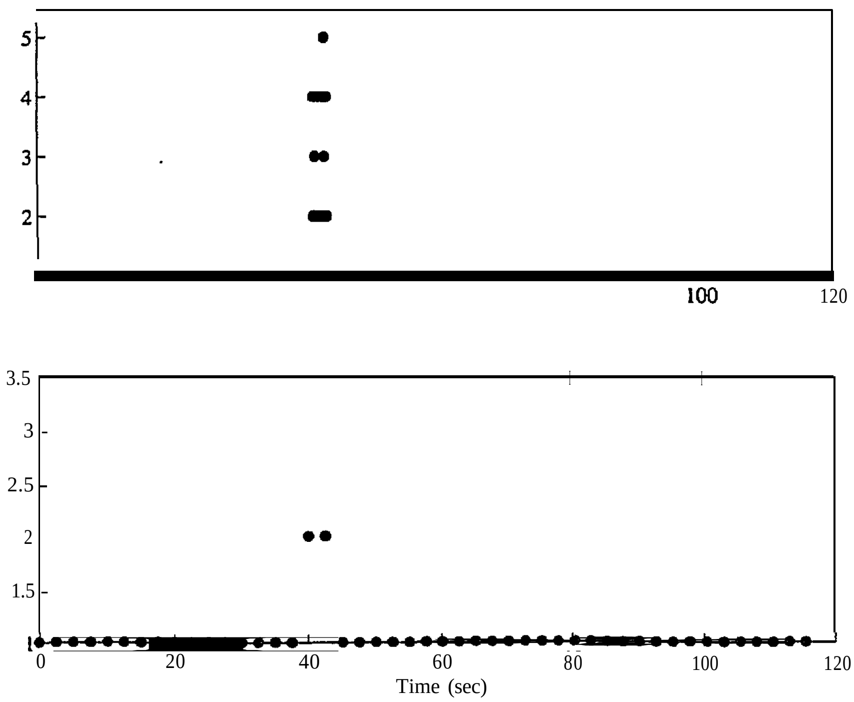


Figure 6-17 Case 2cn: Feature Labeling and Classifier Output

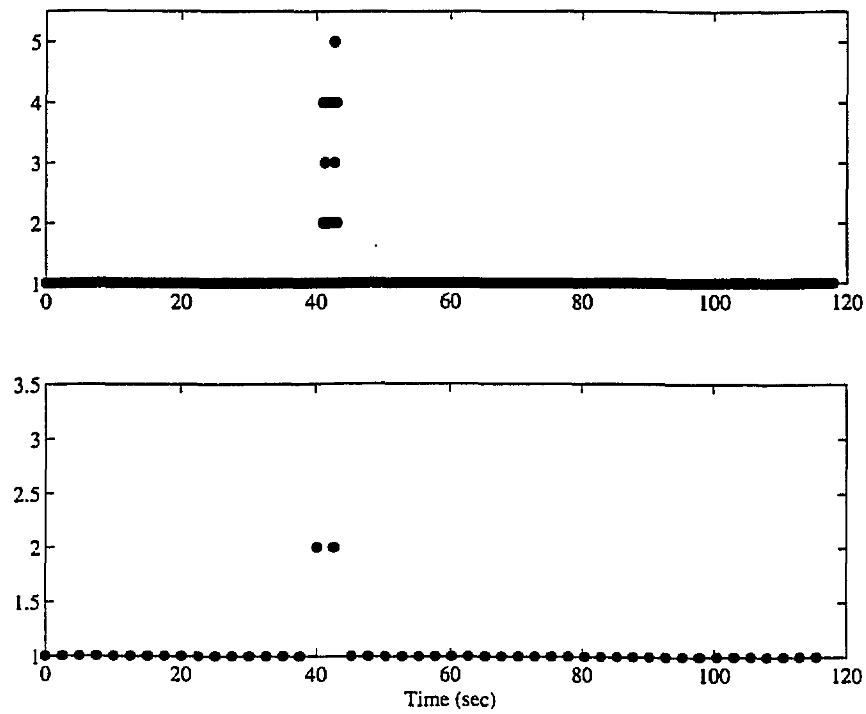


Figure 6-18 Case 2dn: Feature Labeling and Classifier Output

7. REAL-TIME DATA COLLECTION

The objectives of this part of the research were to demonstrate prototype operation of AutoAlert in a real-time data capture mode, once its performance had been demonstrated with controlled, archived operational field data. Resource constraints did not allow for extensive optimization of the algorithm performance in the real-time mode, but the implementation of a prototype algorithm for real-time data processing was completed and analyzed.

7.1 Algorithm Implementation

Using results of related TASC work on restructuring of Hyperstate algorithms to allow them to run in serial, vs. batch, mode, AutoAlert was converted to a serial implementation in Matlab and implemented on a PC hardware platform that supports real-time data capture. This implementation uses a software loop written in the C language to:

- Obtain a block of real-time audio data from the PC analog-to-digital (A/D) converter hardware through the low-level Audio Services of Microsoft Windows 3.1
- Pass the data block to Matlab for processing by the serialized Hyperstate algorithms.
- Display the Hyperstate algorithm output for this data block (i.e., the most likely model: No Accidents/Vehicles Detected, Passing Vehicles Detected, or Accident Detected) and save it to disk.

The software continuously repeats the above three steps and thereby converts a continuous stream of input data from the A/D converter into a continuous stream of model likelihoods.

7.2 Prototype Hardware Suite and Architecture

The architecture for the prototype real-time AutoAlert is shown in Figure 7-l. As shown on the left in the figure, this architecture uses the A/D Converter on the ENSONIQ Soundscape Wavetable Audio card to convert audio signals from the condenser microphone into digital format. The low-level Audio Services of Microsoft Windows 3.1 represent the standard interface between the ENSONIQ card and the C

Language Software Loop. This interface is implemented using a ping-pong direct memory access (DMA) buffer scheme: the Windows operating system fills one of a pair of memory buffers while the C Language Software Loop reads the other. Windows and the C Language Software Loop alternate between writing/reading each DMA buffer- hence the name ping-pong. After the C Language Software Loop reads a DMA buffer, it passes the block of digitized audio data to Matlab, which executes the serialized feature identification and Hyperstate categorization algorithms on the data block and then passes the model likelihoods back to the C Language Software Loop, where it is displayed and stored.

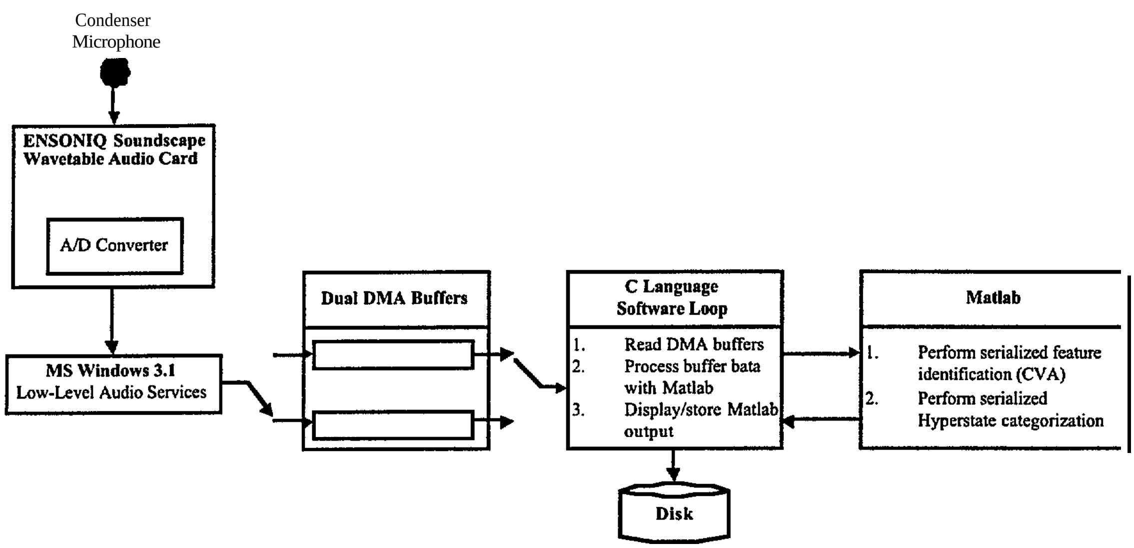


Figure 7-l Prototype AutoAlert Real-time Architecture

7.3 Performance of the Prototype and its Optimization

The basic structure of this implementation proves the concept of real-time Hyperstate processing for AutoAlert. However, the rate at which the software loop repeats directly affects throughput performance and is limited by the performance of Matlab and the Windows 3.1 operating system. Specifically, this implementation uses the Windows 3.1 operating system to simultaneously execute the tasks of input A/D conversion, digital signal processing (DSP), and output display/storage. Windows 3.1 is also used for communications between the tasks. Microsoft Windows 3.1 was not optimized for this type of real-time multitasking operation. As a result, the throughput performance of this implementation would be improved if the same structure were hosted on a dedicated DSP board (e.g., Data Translation's Fulcrum Delta-Sigma or National Instruments' AT-DSP2200). These boards are designed specifically to multitask input

A/D conversion, DSP, and output display and storage. Microsoft Windows was used because it was the operating system under which the laboratory version of AutoAlert was developed – resource constraints did not allow rehosting of the structure onto a DSP board.

The throughput performance of this implementation is also limited because it uses Matlab to perform the DSP functions of feature identification and model classification. Matlab is very flexible for developing algorithms in the laboratory, but it is not optimized for real-time use. As a result, each iteration of the C Language Software Loop is unnecessarily slowed when Matlab executes the serialized feature identification and model classification algorithms.

The following is an analysis that shows:

- The current Matlab implementation is too slow to operate continuously at the throughput rate required for AutoAlert
- The framework for this implementation will execute quickly enough for AutoAlert if it is rehosted in the C language on a commercially-available DSP board.

Figure 7-2 illustrates the number of Matlab floating point operations required per iteration of the C Language Software Loop during real-time AutoAlert execution. Background model processing (every 100 loops) and Hyperstate categorization (every 500 loops) require 2.25 and 0.25 million floating point operations (MFLOP) respectively, but these tasks occur infrequently compared to the regular calculations that require 0.016 MFLOP for each loop. The average MFLOP value is 0.038 per loop.

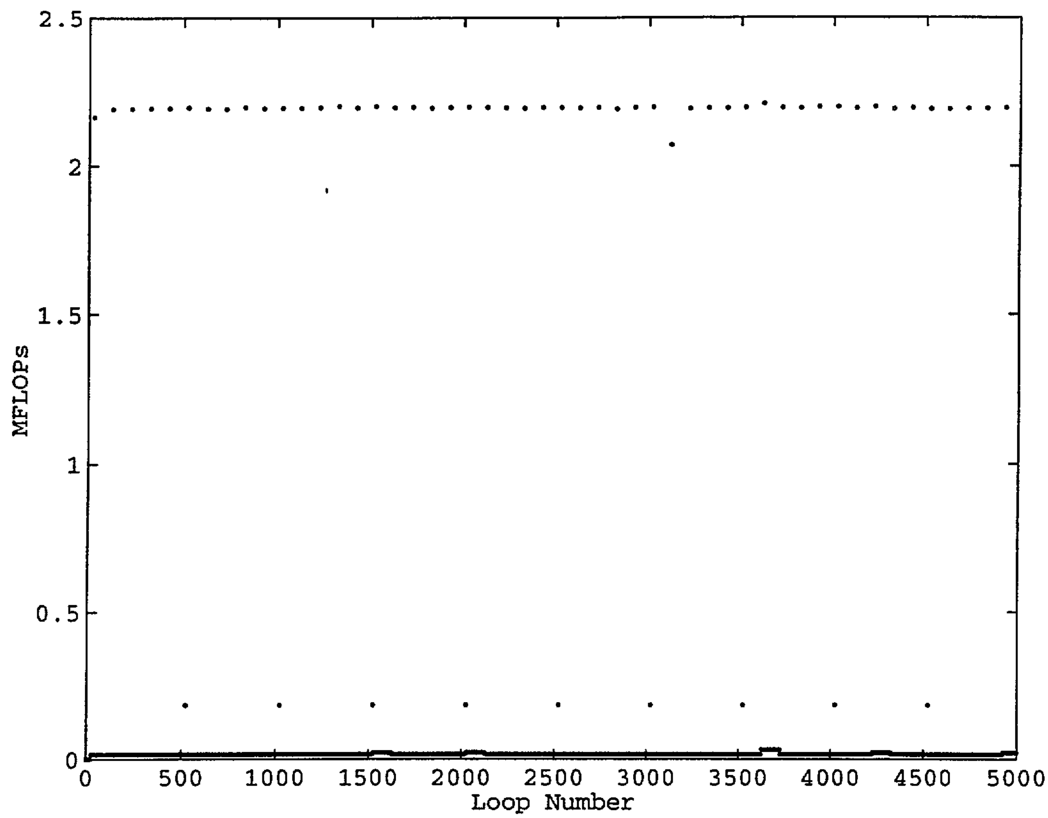


Figure 7–2 Million Floating Point Operations (MFLOP) Required per Loop

The number of seconds Matlab takes to process each loop are plotted in Figure 7–3; the average value is 0.1038 seconds per loop. Each loop processes an 80-point block of audio samples, so the maximum throughput achievable using Matlab is 80 / 0.1038 = 770 Hz (i.e., points per second). Given that AutoAlert requires a 16,000 Hz throughput to address the full frequency content of the input audio data, the 770 Hz rate with Matlab is too slow for AutoAlert. If the data rate exceeds 770 Hz when using Matlab, Windows will write to the ping-pong DMA buffers faster than the C Language Software Loop can read them and the ping-pong operation will jam, thus causing the program to halt. For example, if each buffer is 64 kBytes and the input data rate is 16,000 1-byte samples per second, the program will halt in $2 \times 64{,}000 / 16{,}000 = 8$ seconds.

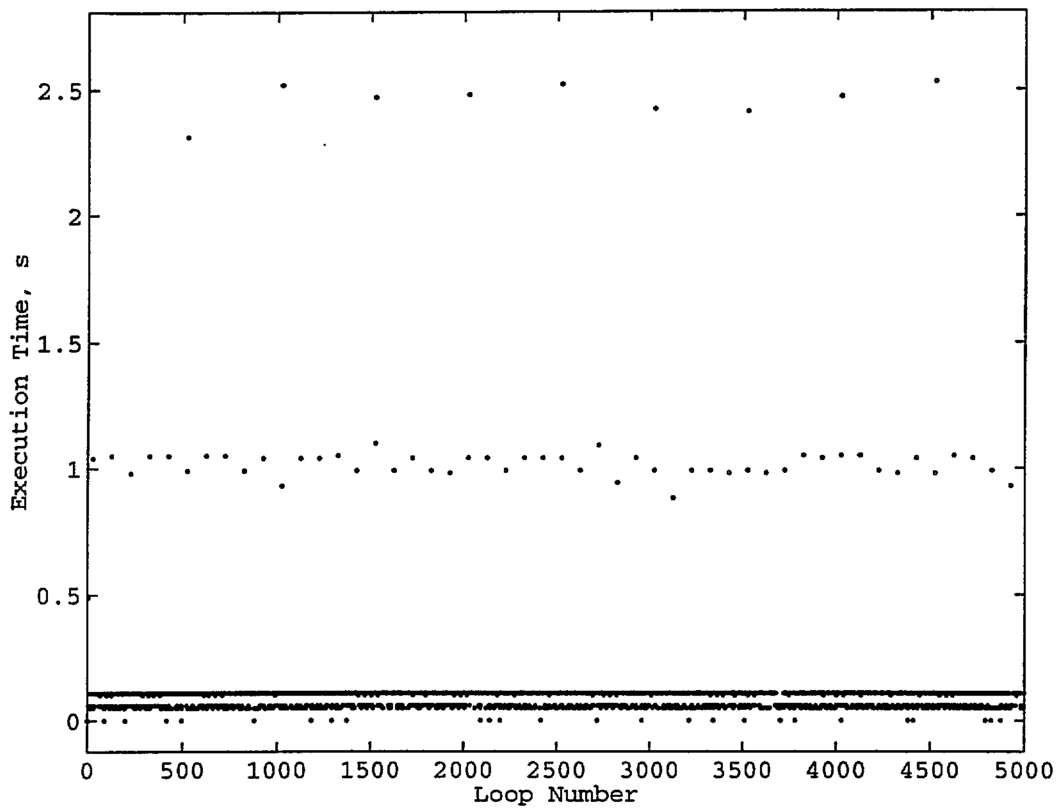


Figure 7–3 Time Required by Matlab to Process Each Loop

To achieve a throughput rate of 16,000 Hz, the average loop processing time must be 80 / 16,000 = 0.005 seconds. With the average computation load of 0.038 MFLOP per loop determined above (Matlab MFLOP counts are valid for any implementation), the average computational speed required is 0.038 / 0.005 = 7.6 MFLOP/second. Commercially-available DSP boards such as the Fulcrum Delta-Sigma and AT-DSP2200 have computational speeds between 25 and 50 MFLOP/second, so it is safe to assume that these boards are fast enough to process data at the rates required for AutoAlert. Furthermore, the memory requirements for AutoAlert processing never exceed 256 kBytes – this amount is well below the 384 to 1,032 kBytes of memory offered on commercially-available DSP boards.

Matlab was used because it was the tool with which the laboratory version of AutoAlert was developed – resource constraints did not allow translating the Matlab code into the C language for a commercially-available DSP board.

8. SUMMARY AND PLANS

The AutoAlert project has achieved its goals. Sources of acoustic data have been identified and obtained. A novel way to generate complex acoustic data sets using a combination of the FRESIM/AHS microscopic traffic simulation and statistical templates for creating "virtual" scenario data was developed and applied. This includes the ability to vary the signal-to-noise ratio of the incident data to test algorithm sensitivity. Several different types of features were analyzed for prototype data sets with and without incidents. Based on this, several spectral features were selected for defining the baseline Hyperstate architecture for incident detection. A detailed baseline AutoAlert architecture was defined, implemented, and evaluated.

Analysis of the prototype system on combined operational field data from both passing background traffic and accidents was completed. The results indicate good accident detection and classification performance for the AutoAlert algorithms for a range of realistic signal-to-noise ratios, even when additional synthetic high-amplitude low-frequency noise is added to mask the input data. This added noise does inhibit accident *classification* and detection of passing vehicles, however. A hardware/software implementation of the system to permit real-time data collection was also protoyped on a Pentium PC, although the prototype has not been optimized for real-time performance.

REFERENCES

1. AutoAlert: Acoustic Detection of Incidents Interim Technical Report, TASC Technical Report TR-07604-0 1, 20 June 1995