University of Southern Queensland

Faculty of Engineering & Surveying

Development of a Noisy Vehicle Detection System

A dissertation submitted by

Benjamin Phillip Curtis

in fulfilment of the requirements of

Courses ENG4111 and ENG4112 Research Project

towards the degree of

Bachelor of Engineering (Electrical & Electronic)

Submitted: November, 2006

ABSTRACT

Traffic noise is an increasing concern in society and is the largest contributor to community noise levels. Besides being a general annoyance and distraction, high noise levels are detrimental to health and wellbeing. Excessively noisy vehicles are a nuisance to communities both with and without high traffic noise levels. Such vehicles consist of cars or motorbikes with illegal exhaust or mechanical configurations and heavy vehicles using engine braking.

A noisy vehicle detection system could be used for the roadside detection of excessively noisy vehicles or those creating unacceptable noise. The presence of a detection system on the roadside would not only aid in gathering information concerning noisy vehicles but would also act as a noisy behaviour deterrent.

The objectives for this project were to investigate the characteristics of vehicle noise on a roadway and current noise measurement techniques and standards. This research then led to the discussion and development of several signal processing techniques for the use in determining the event of a noisy vehicle as well as the vehicle location and noise type.

University of Southern Queensland

Faculty of Engineering and Surveying

ENG4111 & ENG4112 *Research Project*

Limitations of Use

CERTIFICATION

I certify that the ideas, designs and experimental work, results, analyses and conclusions set out in this dissertation are entirely my own effort, except where otherwise indicated and acknowledged.

I further certify that the work is original and has not been previously submitted for assessment in any other course or institution, except where specifically stated.

Benjamin Phillip Curtis

Student Number: 0050009992

____

Signature

____

Date

ACKNOWLEDGMENTS

Thanks go to my project supervisor Chris Snook for guidance and encouragement throughout the year.

I thank God for his involvement in my entire schooling process and for his patience and support throughout this project year.

Benjamin P. Curtis

University of Southern Queensland
November 2006

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 – INTRODUCTION

1.1 Traffic Noise

The importance of managing road traffic noise is demonstrated by increasing population exposure to road traffic noise, and increasing community sensitivity and complaints. The adverse impacts of road traffic noise exposure are well-documented, including lowered property values, productivity losses due to poor concentration, interference with communication, reduction in enjoyment of activities and psychological and physiological health impacts (such as annoyance, sleep disturbance, fatigue and hearing impairment) (Victorian EPA, Road Traffic Noise Strategy Background Paper, 2002).

The largest contributor to community noise levels in terms of the number of people affected is road traffic (Austroads 2004). The noise perceived by a community emanating from a roadway is therefore generally unwanted and many measures exist to mitigate such noise. There are a number of factors that influence the noise level created on roads. These are:-

- The volume of traffic
- The speed of the traffic
- The type of vehicles using the road

Combinations of these three factors can result in high amounts of noise affecting the areas around a roadway. Traffic volume may be defined as the number of vehicles that pass a point on a road in a certain amount of time. An increase in traffic volume results in a greater average sound level on a carriageway. The amount of cars in a group of passing cars will also increase the instantaneous sound level perceived in the surrounding areas. An increase in traffic speed also results in greater road noise emissions. Traffic noise is therefore a problem encountered with high speed, high traffic flow roadways.

The noise level contributing to traffic noise on a road depends upon individual vehicles travelling on the carriageway. Different vehicle types exhibit various noise emission levels and noise characteristics. A roadway used often by heavy vehicles will produce much greater noise levels than corridors with infrequent heavy vehicle use.

1.2 Noise Reduction

The owner of a roadway (usually local, state or federal government) has a responsibility to limit the noise levels reaching surrounding areas whether they be commercial or residential. Each state or district may have differing average traffic noise levels above which noise reduction measures will be considered. Noise reduction measures on operating roads can exist in the form of:-

- Noise Barriers
- Vegetation
- Traffic Management

Noise barriers are commonly used where a noisy highway runs adjacent to residential areas. They are in the form of solid walls or earth mounds and can reduce noise levels reaching the areas surrounding roads by 10 to 15 decibels (cutting loudness in half).

Vegetation, if high enough, wide enough and dense enough can attenuate traffic noise when placed along a roadway. However a large amount of area is required for any significant noise reduction.

The management of traffic on a road can influence its noise emissions and can sometimes be used as a measure to address noise problems. Techniques of traffic management that reduce traffic noise consist of using signage to implement restrictions on certain vehicles or vehicle behaviour. Some residential areas roads for example are off limits to heavy vehicles outside daylight hours or restrict the use of noisy compression braking. Traffic lights can be used to smooth out the flow of traffic and eliminate the noise produced by frequent stops and starts. Speed limits can also be used to reduce overall traffic noise.

The three noise reduction measures outlined above may be implemented on already existing roads that develop noise issues over time. Many of the factors leading to noisy roads can be addressed however in road design before a road is built. Design improvements involve locating heavily used roads away from residential areas or in cuttings below ground level. Likewise the problem of vehicle noise emission can be tackled in vehicle design.

1.3 Noisy Vehicles

The installation of noise reduction measures can be an expensive operation for transport authorities and must therefore be justified by a consistently high noise level from a roadway. An issue faced on many roads is that the overall noise level may not be high enough to justify installation of noise reduction infrastructure but the road still experiences occasional high noise levels. These high levels are caused by excessively noisy vehicles which can often still be heard despite measures to reduce general traffic flow noise. Excessively noisy vehicles can be in the form of:

- Trucks and heavy vehicles (particularly when using compression braking)
- Motorbikes and cars at high revs
- Motorbikes and cars with noisy exhaust or engine configurations

Most countries employ a set of standards which limit the noise emissions allowable from moving vehicles. Australian design rules 28/01 (Appendix A), 39/00 and 56/00 define the maximum allowable sound pressure levels for moving and stationary passenger vehicles, motorbikes and mopeds respectively.

The occurrence of excessive noise levels on roadways is not the only cause of community annoyance. The nature of the sound can also have an impact the severity of its disturbance. Vehicle operations such as high rev’s, squealing tyres, heavy vehicle compression braking and even loud stereos are examples of traffic noise that are not only excessive but also characteristically annoying. The effect of noisy compression braking in residential areas has attracted national attention in recent years and spurred on the reformation of current noise emission legislation. The vehicle recognition research presented in this report has been specifically aimed towards the recognition of heavy vehicles using compression braking.

Compression Braking

Stopping or slowing heavy vehicles tends to cause a great deal of wear on brakes which consequently need frequent replacement. Compression braking causes no wear and can be used to help slow the vehicle before brakes are applied. Application also reduces the occurrence of heat build up in brake drums called brake “fade”. The primary use of compression braking is on downhill grades where frequent braking would otherwise be required. This is an acceptable action in most scenarios however where the roadway lies in residential areas, the sound produced can be a serious disturbance. Engine compression brakes can double the noise emitted by heavy vehicles and this particular noise is not only louder than that usually associated with heavy vehicles, but is sharp, harsh, and disturbing (*Engine Compression Braking* (2004), Queensland Main Roads)

1.4 Concept of a NVDS

The issue presented to communities and traffic authorities by excessively noisy and offensive vehicles has given rise to the concept of a roadside noisy vehicle detection system (NVDS). Such a system could potentially identify vehicles on a roadway that cause a noise level or operation offence. This information could be used by local councils, state departments or developers to monitor average traffic noise levels, monitor occurrence of noisy vehicles and also identify noisy vehicles for possible prosecution. The presence of such a device could deter offensive vehicle behaviour much in the same way as the presence of a speed camera controls vehicle speeds. The installation (either temporary or permanent) of traffic control measures also has a psychological effect in reducing community complaints.

In practice, a NVDS would be mounted somewhere along an affected roadside and would use image or video capture to provide identification of the offending vehicles. The image capture would be triggered by signal processing algorithms fed with the audio information from one or a number of microphones. The system would operate in real time being accurate enough to withstand legal scrutiny.

1.5 Project Aim

The aim of this research project has been to investigate the primary concepts behind a noisy vehicle detection system. The primary investigation of the report introduces the theory relative to outdoor traffic noise such as sound pressure levels, propagation and attenuation, the available acoustic measurement and processing equipment and related research and technology.

The latter section of this report details the application of noise measurement technology and theory to the determination of the following properties of vehicles moving on a roadway:

Noise level – whether a vehicle exceeds its acceptable noise level

Noise type – whether a vehicle is operating in an unacceptable mode

Source location – where the offending vehicle lies on the roadway

The specific aims of the project are given in the project specification in appendix A.

1.6 Project Methodology

The methodology that has been used to achieve a satisfactory project outcome is as follows:

- Investigate the problem posed by traffic noise, its causes and current methods in place to reduce the problem
- Investigate the sources and acoustical properties of excessively high noise levels on roadways
- Identify the characteristics and behaviour of sound propagation in an outdoor setting such as a roadway
- Examine the available technology and procedures for roadside noise measurement and analysis
- Determine which acoustical properties observed on a roadway are suitable for analysis in the hope of isolating noisy vehicles
- Gather field data using audio recording instruments deemed suitable. Use video, image or description to assign recorded data to vehicle types

- Analyse the captured data in the time and frequency domain to identify characteristics that may be used in signal processing for source recognition and source localization
- Develop an algorithm for noise level analysis that can take into account the layout of a specific recording site
- Process recordings of a variety of vehicle types and activities (in particular compression braking) and use characteristic feature extraction for source recognition
- Develop an algorithm for vehicle type and behaviour classification
- Gather field data using multi-point recording.
- Analyse the captured audio and identify characteristics suitable for vehicle position approximation
- Develop an algorithm that uses the multi-point recordings to pinpoint a noisy vehicle's location
- Combine the algorithms to provide a system that is triggered by a high noise level or by an unacceptable noise characteristic to approximate the location of the source on the roadway

CHAPTER 2 - GENERAL THEORY

2.1 Sound Theory

Sound may be described as small pressure variations in air, a property often measured in Pascals (Pa). The human ear responds to a wide range of sound pressures (from 0.00002 Pa to 20Pa), the measurement of which when dealing with the human hearing range can produce an inconvenient range of numbers. The human ear responds to sound in a manner that is approximately logarithmic rather than linear. It is convenient therefore to measure and express the magnitude of sound pressure in a logarithmic scale. Such a measurement is known as a sound pressure level (SPL). Sound pressure levels are measured by electronic devices called sound level meters.

A sound source can take the form of anything from a whistling bird to a resonating panel on a washing machine. The sound received at any point in the surrounding area of a noise source may differ depending on the physical properties of the source and its orientation. In most cases however, it is sufficient to consider the sound from a source as radiating spherically outwards with equal magnitude in all directions. This simple form of propagation is described by what is called spherical wave theory. While it is obvious that motor vehicles as sound generators can be broken into a number of contributing sources, each radiating sound spherically, a single point assumption has been adopted for simplification in the initial stages of this project. A uniform wavefront magnitude will also be assumed.

Considering then a spherical wavefront, the acoustic power radiated by a point source will have the value

$$W = 4\pi r^2 I \tag{2.1}$$

where I is the acoustic intensity measured a distance r from the source. If then the acoustic power of a sound source at a certain moment in time has a value of W, the acoustic intensity observed at any point surrounding the source will be proportional to the acoustic power divided by the distance from the source squared. This relationship of magnitude reflecting the inverse of the distance squared is an important property of a sound field.

The acoustic intensity is related to the mean squared pressure by the equation (Snook)

$$I = \frac{\langle p^2 \rangle}{\rho c} \tag{2.2}$$

where ρc represents the acoustic impedance of the propagating medium.

As was mentioned above, sound levels are dealt with in a logarithmic scale in the form of sound pressure levels, the unit of which is the decibel (dB). The sound pressure level is also derived from the root mean square value of sound pressure as (Snook)

$$L_p = 10 \log\left(\frac{\langle p^2 \rangle}{p^2{}_{ref}} \right) \tag{2.3}$$

where $p^2{}_{\mathrm{ref}}$ constitutes a reference level being the lowest audible sound pressure of 0.00002 Pa. When this value is substituted into the above equation, the expression for sound pressure level simplifies to (Snook)

$$L_p = 10 \log\langle p^2 \rangle + 94 \tag{2.4}$$

2.1.1 Addition

The sound level observed at any point in an uncontrolled sound field will most often be made up of a number of contributing sound levels from different sources. As will be explained later, overall sound pressure is a measure of the combination of sound pressures from a contributing frequency range. The addition of sound levels therefore depends on the frequency content of the contributing signals. In addition, sound pressure levels are logarithmic values and cannot be added linearly. There are two general techniques used for the addition of sound pressure levels.

Coherent sound pressure levels are those produced by sources having identical frequencies that are not phase. Their combined mean squared pressure value may be found as (Bies & Hansen)

$$\langle \rho^2 \rangle = \langle {\rho_1}^2 \rangle + \langle {\rho_2}^2 \rangle + 2\langle \rho_1 \rho_2 \rangle \cos(\beta_1 - \beta_2) \qquad (2.5)$$

where β_1 and β_2 are the signal phase angles relative to some datum. The sound pressure level may then be found using equation 2.4 above. The occurrence of coherent sources is a rare occurrence in real life and the detection of such is difficult in an uncontrolled sound environment.

Incoherent sounds on the other hand are very common. These are sounds from sources with phases at random and quite possibly different frequencies. The expression for the addition of incoherent sound pressure levels is (Bies & Hansen)

$$L_P = 10\log(10^{LpA/10} + 10^{LpB/10} + \ldots + 10^{LpX/10}) \qquad (2.6)$$

This relationship is applicable regardless of the amount of sound sources present. Motor vehicles are generally considered to be incoherent noise sources because of the highly complex and differing frequency and phase patterns they generate. This relationship also applies where reflected sound waves are received at an observation point along with the original incident wave.

The speed at which a sound wave propagates through the atmosphere is called the speed of sound and is defined by the approximation below (Snook)

$$c \approx 332 + 0.6T \qquad (2.7)$$

where T is the temperature of the air. This approximation is accurate to about 0.1% at temperatures near 20ºC.

2.1.2 Doppler Shift

For a noise source moving through space with velocity v_S and producing a sound travelling at velocity c with frequency f_S, the perceived frequency at a stationary point is given by (Wolfram Research)

$$f = f_S\left(\frac{c}{c + v_S}\right) \tag{2.8}$$

This phenomenon is called Doppler shifting and is observed in all cases of electromagnetic and sound wave propagation where there is a velocity difference between the source and the receiver.

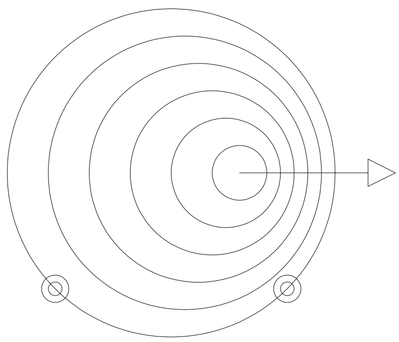


Figure 2.1: Wavefront propagation of moving source

Figure 2.1 above gives a simplistic view of the effect that a moving source has on the perceived periodicity of a signal wavefront. It is seen that the observation point behind the moving source receives wave fronts further apart than they were generated. The observation point in front of the moving source sees wave fronts at a higher frequency than they were produced.

2.2 Vehicles as sound sources

As mentioned previously, motor vehicles are made up of many components that act as individual noise sources. The sound produced by these components combines to produce the perceived vehicle noise. The main sound sources on a motor vehicle are the engine and exhaust. This is dependant however on the condition and make of the vehicle. Many new cars for example are designed so that engine and exhaust noise are minimal. Here the externally created noise will dominate the vehicle's acoustic emission. External vehicle noise consists of the noise created through the interaction of the vehicle with its surrounding environment. The most common external noise source is the tyre on road noise produced by a moving vehicle.

2.2.1 Tyre on road noise

There are many factors that can affect tyre noise generation. The majority of these factors may be separated into three categories; tyre properties, vehicle characteristics and road surface.

Tyre properties include tread patterns, tyre width, and state of wear of the tyre and rubber hardness. The vehicle characteristics that affect tyre noise generation are the number of tyres, vehicle weight and inflation pressure. These properties are uncontrollable on a public roadway and will therefore have little effect on the specific design considerations of a roadside monitoring system. Properties that affect tyre noise of a vehicle which are site specific are the road surface and the speed limit on the road section.

The two primary causes of tyre/road noise related to road surface texture are the air-pumping effect and tyre vibration (Austroads 2005). There are two components to the air-pumping effect. These are tread enclosed air pumping and road surface enclosed air pumping. Variations in the road surface texture lead to deformation of vehicle tyres and the generation of noise through tyre vibration.

Figure 2.2 below illustrates the effect of macrotexture (small) and megatexture (large) variation in road surfaces. It is seen that where the surface variation wavelength is relatively small (compared to a tyre footprint), little tyre deformation occurs. Here there is little air pumping or tyre vibration. The open air voids beneath the tyre also allow for the escape of air. A surface variation wavelength comparable to the tyre footprint size will cause much tyre deflection and vibration. An even larger surface variation (often due to bumps or potholes) may cause little tyre deflection but considerable air pumping.

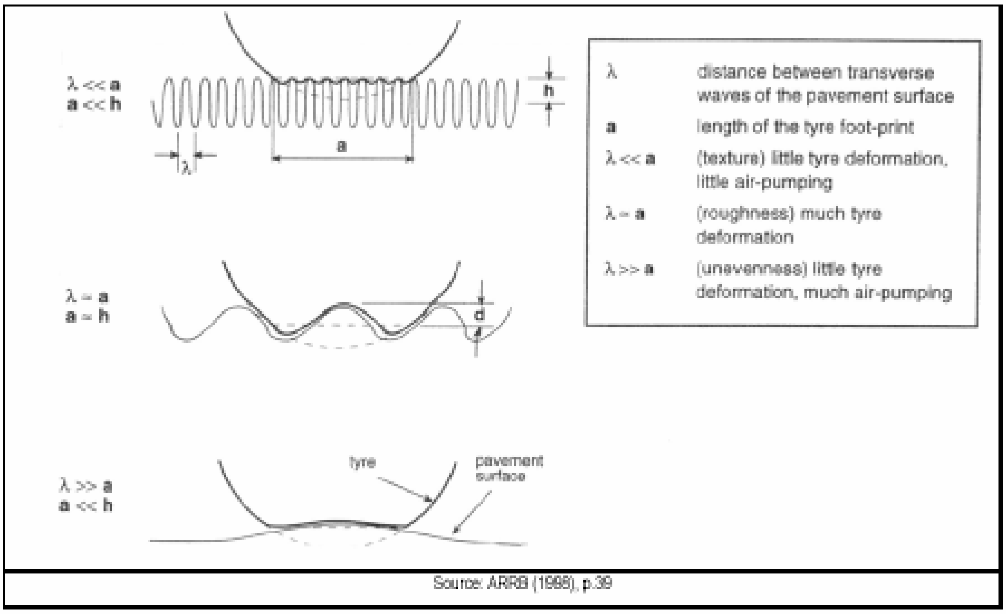


Figure 2.2: Tyre deformation (Austroads Inc (2005) AP-R277-05)

Differences in road surfaces can result in tyre/road noise emission differences of up to 12dB(A). Low noise roads generally aim to minimize the road surface macrotexture while avoiding excessive air pumping due to a flat pavement. The condition of the site specific road surface should then be considered in the operation of a roadside noise analysis device. In addition to road surface texture, the presence of wet patches on the road and the immediate speed limit are also factors worthy of consideration.

2.2.2 Vehicle size

The size and type of a vehicle can play a large role in the number and type of individual noise sources. Heavy vehicle for example usually have an exhaust outlet above the drivers cabin where passenger cars have exhaust outlets at the rear underside of the vehicle. The combination of these noise sources therefore behaves differently in both cases producing varying sound patterns at surrounding observation points. The decision for a single point source assumption in the initial developments of this project has been made to both simplify the sound processing theory required and the source location process.

2.2.3 Vehicle Operation

As defined in Chapter 1, vehicle operation as well as noise level is of interest when identifying offending vehicles. Some of the current vehicle operations and sound types that could warrant vehicle identification are high rev's, squealing tyres, loud incessant horns, modified exhausts and heavy vehicle compression braking.

Compression Braking

Engine compression braking (otherwise known as exhaust braking or Jake braking) refers to an add-on braking mechanism for diesel engines. The principle behind compression braking is simple. Its application changes the action of the engine exhaust valves, turning the engine into a large scale air compressor. The mechanism converts a power producing diesel engine into a power-absorbing retarding mechanism (Engine Compression Braking (2004))

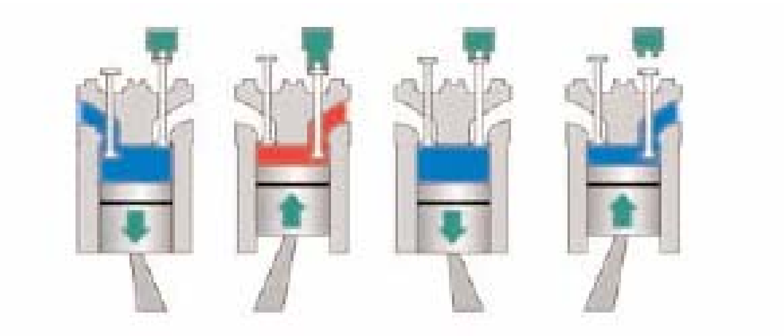


Figure 2.3: Operation of a compression or Jake™ braking (Queensland Main Roads)

The operation of the retarding braking system (for a 4 stroke engine) is portrayed in Figure 2.3 and can be explained simply as:

1. The intake valve opens and air is forced into the cylinder by boost pressure from the turbocharger
2. Air is compressed to approximately 3500kPa by the piston with energy provided by the driving wheels. Near top stroke, the exhaust valve is opened allowing the compressed air to dissipate.
3. On the downward stroke, no energy is returned to the piston (and therefore the wheels) so there is a loss of energy.
4. A normal exhaust stroke follows

2.3 Fourier analysis

In the signal processing environment, many signal analysis techniques and transforms are utilized in gaining information. One of the most common of these techniques is called Fourier analysis. Fourier analysis states that any periodic waveform may be divided into an infinite number of contributing sine and cosine waveforms. These waveforms are considered as harmonic components of the original waveform. The Fourier series is derived using the well known relationship (Wolfram Research)

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n \cos(n x) + \sum_{n=1}^{\infty} b_n \sin(n x), \tag{2.9}$$

Use of the Fourier series however requires that the analysed signal be periodic. To find the frequency contents of any waveform (periodic or not), the continuous Fourier transform is used. The continuous Fourier transform is defined as (Leis)

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \tag{2.10}$$

Where x(t) is a continuous time domain signal .

In continuous signal processing applications, the Fourier transform is used to convert an expression in the continuous time domain x(t) into the continuous frequency domain X(f). The advent of the digital computer has facilitated the development of a discrete Fourier Transform (DFT). The DFT is a mathematical procedure used to determine the harmonic (or frequency) content of a discrete signal sequence. The DFT is defined as (Leis)

$$X(m) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi nm/N} \tag{2.11}$$

Where x(n) is a discrete sequence of time-domain sampled values of the continuous variable x(t).

2.4 Sound Level Meters

Sound level meters are real time analysis instruments that use omni-directional condenser microphones to measure and record the sound level at a certain point.

Either a ½ or 1 inch condenser microphone is used to record acoustic pressure fluctuations that act to deflect its thin diaphragm.

The measurement of a sound level meter will most commonly incorporate the levels from frequencies within the audible sound range (typically 20Hz to 20,000Hz). Many mid to high range instruments will also provide the individual levels of component octave or 1/3 octave frequency bands within the measurement range. The separation of frequency bands as achieved either through the use of analogue filters or digital processing (Discrete Fourier Transform).

Many sound level meters will also apply a weighting filter to each individual frequency band to mimic the sensitivity of a certain environment or observation subject. The most commonly applied filter in all sound measurement fields including traffic studies is called the 'A' weighting. The 'A' weighting scale closely approximates the response of the human ear, thus providing a measure of the subjective loudness of noise and enabling the overall noise level of noises with different frequency characteristics to be compared. Figure 2.4 shows the actual relative response of the human ear over the acoustic range (blue) as well as a commonly applied approximate (red) 'A' filter.

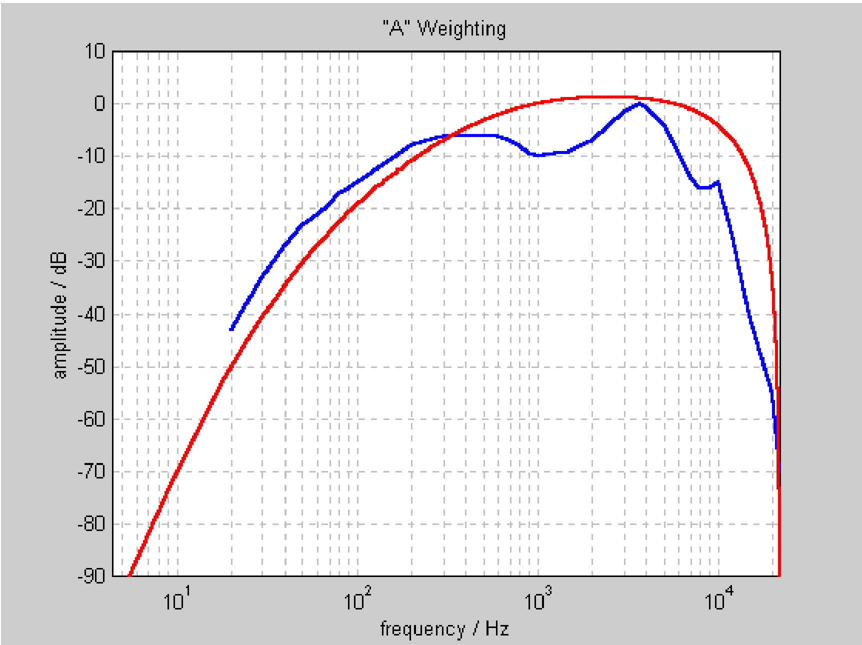


Figure 2.4: Actual (blue) and simulated (red) 'A' weightings

The instantaneous sound pressure level at a point cannot be detected or calculated for an independent moment in time but is found rather using a very short time constant. The time constant for a sound level meter is an RMS averaging time that may be adjusted to suit the nature of the sound being measured (Snook). Typical time constants used are 125ms for a fast response and 1 second for a slow response.

In the case of road traffic noise, noise is not constant and varies over time. In an effort to accurately quantify time variant noise, a number of noise indexes have been developed which portray certain characteristics of the noise. Some of these indices are:-

LAeq - Equivalent continuous sound level, commonly referred to as the average noise level.
LAmax - The maximum noise level. This is defined as the highest noise level which occurs during any noise event occurring during a particular time period.
LA10 - The noise level exceeded for 10% of the measurement period, commonly referred to as the average maximum level.
LA90 - The noise level exceeded for 90% of the measurement period, commonly referred to as the background level.

2.5 The Roadway

The Queensland road authority, The Department of Main Roads specifies the design rules for road parameters. The document, *Road Planning and Design Manual,* 2004, states that the desired lane width for Queensland roads is 3.5m. This is enforced on marked multi-lane roads subject to heavy vehicle traffic. The legal limit for commercial vehicle width on Australian roads is 2.5m. Manufacturers tend to make the most out of this limit as many heavy vehicles and buses are exactly 2.5m wide.

This chapter has provided a brief introduction to some of the many sound field properties existent on a roadway. An introduction to the specific contributors to road noise and the standard noise measurement equipment was also considered necessary.

CHAPTER 3 – RESEARCH REVIEW

3.1 Existing Traffic Noise Technology

3.1.1 The Noise Camera

The “Noise Camera” is a device designed by Acoustic research laboratories Australia for the intent purpose of identifying noisy vehicles on a carriageway. The system consists of two primary components, a sound analysis module and a video camera processor module.

The sound analysis module monitors a roadway using a single omni-directional microphone configured as a type 1 sound level meter by AS-1259.1 and AS-1259.2 (Australian standards for precision sound level meters). The road noise is recorded using a high end sound card as a “.wav” file and is then sent wirelessly to the camera processor module.

The camera processor consists of two video cameras, video capture hardware and signal processing hardware. One of the cameras surveys the entire road where the other is zoomed and focused on a particular lane to be monitored. The device can be instructed to capture video and audio only for instances of road noise exceeding a given level or for continuous capture. The two video streams and the audio can then be wirelessly accessed by a user and analysed to identify any noisy vehicles.

Some of the shortcomings of this system in regards to an automated noisy vehicle detection system are:-

- The omni-directional microphone is not shielded and therefore records noise from all surrounding areas. The device is mounted in an elevated position thus increasing the noise input area.
- The system is completely user discerned. The video and audio must be manually assessed by an operator to identify the type and whereabouts of noisy vehicles.

- The single lane zoom camera means that only one lane of a roadway may be monitored for vehicle identification.
- As there is only one recording point, spatial information concerning noisy vehicles cannot be gathered. Rather, if there is more than one vehicle on the road section capable of creating the recorded noise, the information must be discarded. The system cannot discern between directions of travel.
- The noise camera system is a highly elaborate piece of equipment. The technology involved such as video cameras, video capture hardware, wireless communications and large hard drives result in an expensive and elaborate product.

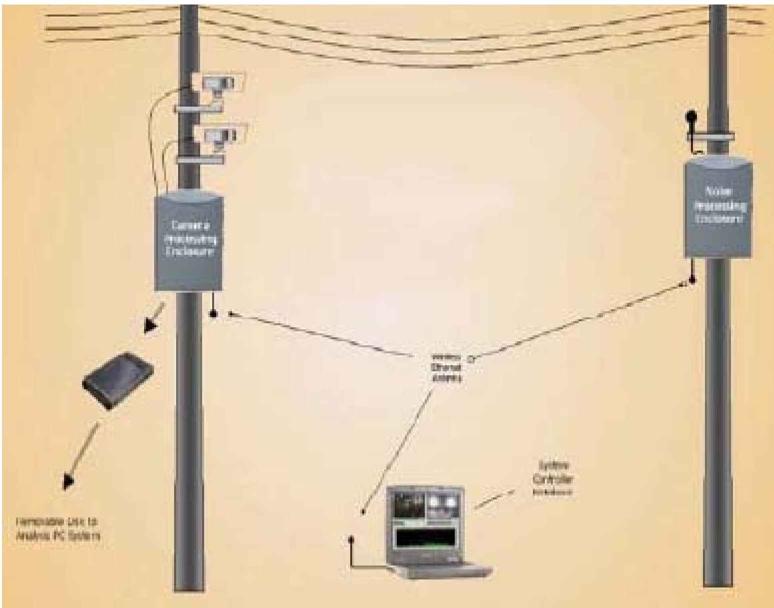


Figure 3.1: Roadside configuration of the Noise Camera (Acoustic Laboratories Australia)

Some of the features of the system that would be possibly useful in the design of an automated noisy vehicle detection system are:

- The components are mountable on existing infrastructure (see Figure 3.1). This reduces mounting costs and hinders vandalism
- The system may be accessed and the data downloaded without removing the device from its location
- The recording microphone is mounted above the traffic flow thus reducing the amount of possible sound attenuation from objects or vehicles on the roadway

3.1.2 Smartek™ Acoustic Sensor

The SmarTek Systems Acoustic Sensor (SAS-1) is a non-contact, passive acoustic (listen only) sensor for multilane traffic monitoring at all speeds from free flow to stop and go. It is completely non-intrusive to the roadway and operates from a "sidefire" position (Figure 3.2).

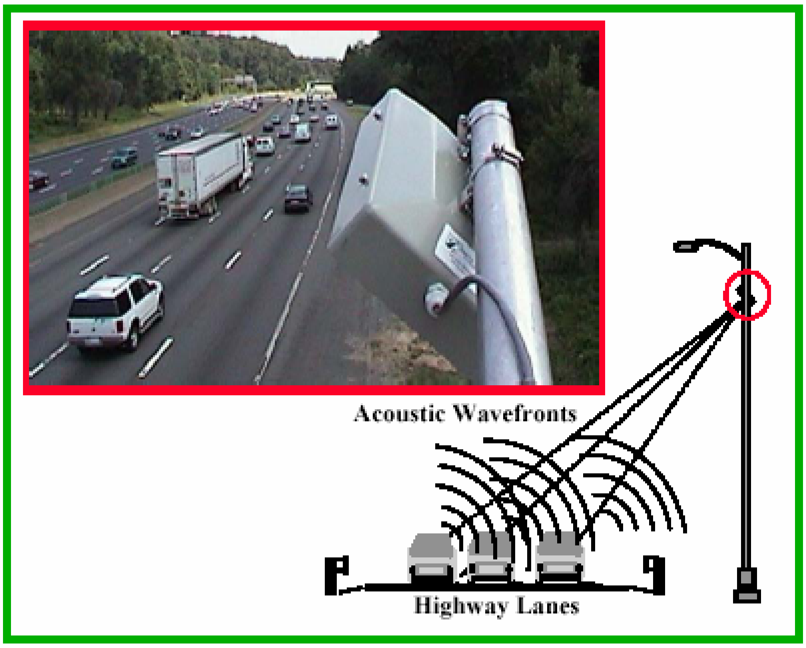


Figure 3.2: Sidefire operation of the Acoustic Sensor (SmarTek Systems)

The SAS-1 utilizes signal and spatial processing technology to provide high resolution "acoustic imaging" of all vehicle traffic passing below the sensor. This selective processing eliminates false vehicle detections caused by out of lane or off road noise. Up to five detection zones are formed by selecting the position and number of contiguous look directions (of which there are a total of 91) which are combined for actual vehicle detection.

Because the SAS-1 "acoustically images" the roadway traffic with a large number of high resolution cells (look directions), the end user is provided with flexibility to electronically position each detection zone and to set each detection zone's size and sensitivity. This capability eliminates the need for precise mechanical "pointing" of the device during installation.

The SAS-1 is comprised of an array of rugged microphones, analogue signal conditioning, and sampling circuitry for converting impinging acoustic signal wave fronts to digital signals. These digital signals are processed using a programmable Digital Signal Processor with associated memory and communication circuitry. The processing software then implements the required signal processing, spatial processing, and vehicle detection algorithms. The use of dedicated signal and spatial processing works to create 91 acoustic signal arrival direction channels (look directions) each of which are 15 degrees wide as shown in Figure 3.3.

Some notable features of the sensor with respect to the development of a noisy vehicle detection system are:

- Due to the "sidefire" mounting, no lane closures are needed for installation
- The device may be mounted on existing infrastructure
- The use of digital signals processing to create effective "look directions" and therefore limit the detection area

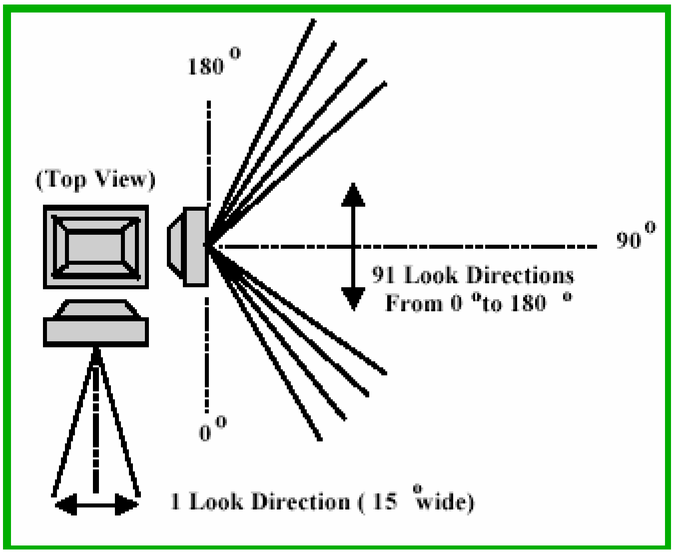


Figure 3.3: Effective look directions of the Acoustic Sensor (SmarTek Systems)

3.2 Roads in Australia

3.2.1 Organizations

State Departments

There are a number of organizations that control and contribute to the design and operation of roads in Queensland and Australia. Within each state and territory the majority of significant roads are owned and serviced by a separate road transport authority such as the Department of Main Roads in Queensland. The remaining roads are the responsibility of local councils.

Austroads

Austroads is the association of Australian and New Zealand road transport and traffic authorities. Austroads members are the six Australian state and two territory road transport and traffic authorities, the Australian Department of Transport and Regional Services (DOTARS), the Australian Local Government Association (ALGA), and Transit New Zealand.

The National Transport Commission

The National Transport Commission (NTC) is an independent body established to assist Australian Governments in achieving a jointly agreed objective of:

"…improving transport productivity, efficiency, safety and environmental performance and regulatory efficiency in a uniform or nationally consistent manner." (NTC website)

The NTC (formerly known as NTRC – National Road Transport Commission) acts to develop national approaches to road transport regulations by working in close partnership with transport sectors, governments, regulation associations and police.

3.2.2 Relevant Literature

There are a number of publications available from state and federal transport authorities which outline road specific design and analysis procedures. The Austroads research report: **AP-R277/05 - Modelling, Measuring and Mitigating Road Traffic Noise** examines the role of noise impact assessments in traffic management, noise modelling, traffic noise calculation methods, noise measurement and monitoring and traffic noise mitigation. The report discusses appropriate noise measurement equipment and procedures as well as investigating noise on the roadway.

Compression Braking and the NRTC

Australia's vehicle noise standards have not been reviewed for some years and at present, they permit significantly higher noise levels than standards elsewhere in the world. The NRTC is addressing the matter of engine brake noise as part of the Third Heavy Vehicle Reform Package. It is expected that the project will result in new noise standards and a new regulation for engine compression brake noise. The NTC report *ENGINE BRAKE NOISE: DEVELOPMENT OF A ROADSIDE TEST PROCEDURE*, (August 2003) confronts the issue of recording, identifying and prosecuting the noise produced by heavy vehicle compression braking.

The report concludes that the characteristics which are common amongst recordings of annoying engine break noise are a maximum instantaneous (Lmax) sound level greater than 95dB(A) and 5 or more modulations over 7dB(A) in 0.5 seconds with a maximum level of at least 80dB(A). The study involved taking roadside recordings of 600 passing heavy vehicles in various locations and analysing maximum sound levels, frequency content and modulation rate. It is important to note however that the proposed regulation acts only to identify compression braking as annoying or not. The actual automated identification of the application of compression braking has not been addressed.

Australian Standards and Design Rules

There are a number of Australian standards and design rules which specify the equipment and procedures required when conducting a compliant sound level survey. The maximum allowable noise limits for motor vehicles are defined in a series of Australian Design Rules. The major standards developed for the measurement of environmental and road traffic noise are summarized below.

AS 2702-1984 Acoustic methods of measurement of road traffic noise - Sets out the method for the measurement of road traffic noise indexes such as Leq(A).

AS1259.2-1990 Acoustics - Sound Level Meters - describes instruments for the measurement of frequency weighted and time averaged sound pressure levels, or sound exposure levels.

AS1633 Acoustics – Glossary of Terms and Related Symbols.

AS2659.1-1988 & AS2659.2-1983 - Guide to the use of sound measuring equipment, which provides guidance on the use of sound level meters.

AS2900 Quantities and Units of Acoustics.

AS/NZS4476 Acoustics – Octave-Band and Fractional Octave-Band Filters.

ISO9613 Acoustics – Attenuation of Sound during Propagation Outdoors - This standard is used as a basis for estimating noise propagation effects in several noise calculation
methods.

AS IEC 61672.1-2004: Electroacoustics -. Sound level meters - Specifications. Class 1

Australian Design Rules
The Australian Design Rule 28/01 – "External noise of motor vehicles" is the current Australian compliance standard for all Australian motor vehicle noise emissions. The standard was developed with the objective of limiting the overall contribution of motor traffic to community noise. The ADR is used as an industry limit for car manufacturers as it applies to all motor vehicles manufactured after 1993. The Design Rule states the maximum allowable noise level that can be produced by a number of separate vehicle classes either while in motion or while stationary. A table of noise limits is provided in Chapter 9. Similar design rules exist for motorcycles (ADR 39/00) and for mopeds (ADR 56/00).

3.3 Relevant Research

Classification
There are a number of published academic research reports pertaining to areas of interest for a noisy vehicle detection system. Acoustic source recognition is by no means a new area of research and has been extensively covered particularly in speech

recognition applications. The classification of vehicle types is a somewhat newer application of acoustic recognition and may be said to be relatively undeveloped. One paper of interest is *Neural Fuzzy Techniques In Vehicle Acoustic Signal Classification* by Somkiat Sampan.

The dissertation involves classification system design for vehicle acoustic signal classification. There are three parts considered in this design: acoustic sensor design, feature analysis and extraction, and classifier design. The main focus is on the classifier design. Two main networks, multilayer perceptron networks (MLPs) and adaptive fuzzy logic systems (AFLSs), are considered, analysed, and used as classifiers.

After a vehicle is detected significant features are extracted. There are 30 features extracted for each vehicle. All features are time domain features extracted directly from acoustic energy information. Since different classes of vehicles have different sizes, the ending point of feature extraction was adapted according to an initial prediction by a fuzzy logic system. The information of duration and loudness are used as the inputs to this prediction system.

The report demonstrates the possibility of an acoustic sensor system with 97.95% correct classification rate between small and large vehicles on 1327 vehicles, 92.24% correct classification rate in four-class problem on 1327 vehicles, and 78.67% correct classification rate in five-class problem on 1327 vehicles.

The sole reliance on the time domain energy information limits the process to accepting only uncorrupted (isolated) vehicle pass-by waveforms. The feature extraction technique also introduces significant processing latency, as a complete vehicle pass-by of (typically 5-6 seconds) is required.

Localization

The problem of source localization is also far from being a new research concept. The enormous field of radar has for a number of years helped in the development of effective source localization techniques. Acoustic source localization applications such as conference speaker localization and enemy firing localization in military operation have also recently appeared. The paper *Acoustic Source Localization Using Time-Delay Estimation* by Brent C. Kirkwood, August 4, 2003 provided insight into the application of far field localization for jet aircraft.

The angular location of a jet airplane was estimated by measuring an acoustic direction of incidence based solely on the noise produced by the aircraft. Cross-correlation based algorithms for passively locating jet aircraft were chosen as most suitable and implemented in a four-element multi-dimensional microphone array. The investigation was performed through simulations, laboratory measurements, and field measurements. Simulations in which the accuracy of the method was investigated provided verification of the theoretical analysis. It was shown that the cross-correlation time-delay estimation was not so dramatically affected by reflections or multiple sources in the field as were the other techniques considered.

CHAPTER 4 - LOCALIZATION OF A SOUND SOURCE

4.1 Sound Attributes

There are a number of attributes in a sound signal that may be exploited in determining the location and dynamic nature of a sound source. Some common examples are signal amplitude attenuation, frequency shift, phase delay and time delay. These signal attributes may be physically observed and provide information dependant on the speed and location of a sound source with respect to single or multiple observation points.

It is well known in many localization cases including acoustics that a larger number of observation points provide greater accuracy. One of the most common examples of multi-point acoustic observation is seen in binaural localization. In this natural process, the position of a sound source is interpreted by the brain using time delays and amplitude differences between two observation points (the ears). Although this is a somewhat simplified explanation of the process, it can be seen that these acoustic properties (particularly time delay) have been widely integrated into the signal processing field of source localization. One major use of signals processing in source localization is the field of Radar which has been responsible for the development and advancement of processing techniques in detection, tracking and imaging.

The requirement for a noisy vehicle detection system is a relatively simple and robust localization device. The signal attributes initially considered for providing source location information are

- Signal amplitude
- Time delay
- Frequency shift (Doppler Shift)

4.1.1 Signal Amplitude

Equation 2.1 in section 2.1 states that the amplitude of a signal at a distance r from a point source will be inversely proportional to r^2. An observation of the signal amplitude at multiple points otherwise known as an array could therefore be used to determine the bearing of the source from the array. A greater number of points would provide for greater accuracy. One problem that is foreseen with such an application is the effect of signal attenuation in a traffic environment. It is quite possible that the signals from noisy vehicles may be attenuated by other vehicles on the road. This could lead to an omission of a noisy vehicle measurement or a false location for the vehicle. One way of avoiding such attenuation would be to place the microphone array above the traffic flow so that vehicle noise is not intercepted.

The presence of multiple vehicles on a roadway will also cause disturbances in localization data.

4.1.2 Time Delay

Each acoustic signal produced by a vehicle may be considered as a time variant waveform $x(t)$. Due to the finite propagation velocity of sound, the waveform as it appears at a location a distance r from the vehicle may be considered as $x(t-r/c)$. Where c is the speed of sound. The waveform at the observation point may therefore be said to be equivalent to the source signal delayed by r/c seconds. In a single microphone application, a knowledge of the source waveform and its time behaviour will allow for the calculation of the distance r using the observed waveform. As the behaviour of vehicle waveforms is unknown to a passive sound system, the use of more than one microphone is required in locating the source. This technique is called "time difference of arrival" (TDOA) and is heavily utilized in radar applications.

The TDOA between two microphones is the difference in the times of the arrival of a particular wave front at each point. Figure 4.1 illustrates the TDOA as perceived between two microphones detecting a wavefront.

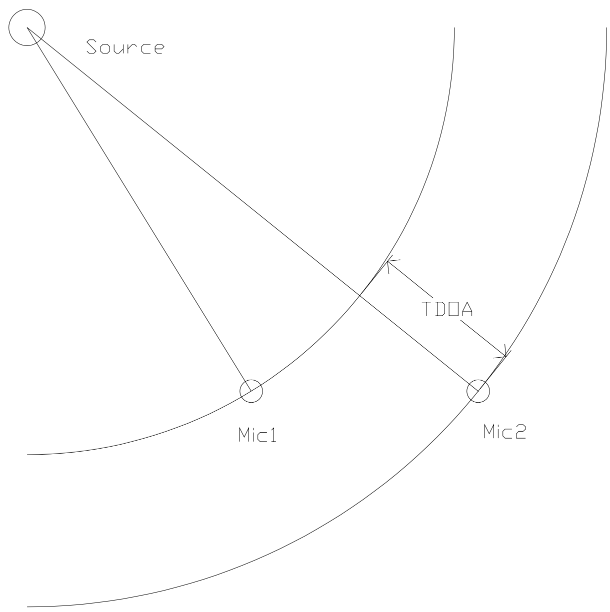


Figure 4.1: TDOA between two microphones

Unlike the signal amplitude, the time delay between a source and microphone will remain relatively un-affected by the presence of physical objects in the propagation path. The most common technique used to determine the TDOA is the cross-correlation of the signals from two microphones. As with most localization methods, a greater number of observation points (or in this case, microphone pairs) will produce a more accurate location estimation. The orientation of microphone pairs in three dimensions can also provide a location estimation in three dimensional space.

4.1.3 Frequency Shift

A shift in frequency of a sound signal (otherwise known as Doppler shift) occurs to different degrees when the source is moving at a velocity with respect to the observer. The amount by which the frequencies of the source signal are shifted is proportional to the relative velocity of the source. A single observation point will receive a signal that is higher in frequency than the source signal when the source is moving towards the point. The frequency will appear to be lower when the source is moving away from the point. Of course, without knowledge of the actual frequencies produced by the source, it is difficult to determine when the source actually passes the microphone. The use of several microphones however can overcome this by comparing frequency shifts over time.

In order to perform an analysis on the frequency waveform of a time varying sound signal, a fairly high resolution will be required. Subtle shifts in the frequency bands would be too small to be observed in a 1/3 octave response available from a sound level meter.

An initial analysis of single vehicle pass-by recordings found that the multiple frequency components making up the vehicles acoustic waveform are by no means constant over time neither do they shift linearly during the pass-by. This is due to the inconsistent operating nature of the vehicle and also the un-controlled effects of a moving vehicle. Although the frequency components of the signal may be similar over the pass-by and exhibit a linear shift trend, the field data is considered too erratic for any accurate localization interpretation.

4.2 Localization Using TDOA

From the signal attributes discussed, the time difference of arrival (TDOA) is considered as the most effective and accurate in determining the location of a sound source. As mentioned above, a value of TDOA may be obtained using the signals from two spatially separated microphones in a process called cross-correlation. This will be discussed further in Chapter 5. There are a number of ways by which the TDOA between microphone pairs may be used in estimating the position of the source. Even the simplest of these requires more than one TDOA value and hence more than microphone pair. As the identification of a vehicle on a roadway requires only location information in two dimensions, a three dimensional microphone array is not required. Therefore in keeping with the requirements for a simple array design, a three microphone array has been chosen for analysis. An array of three microphones arranged in a straight line provides two unique microphone pairs and will therefore provide adequate two dimensional location estimates. Figure 4.2 shows the linear array and the associated time delays for a passing vehicle.

Two alternate approaches to TDOA localization are the near field and far field methods. The use of TDOA in radar applications is usually accompanied by the far field method which utilizes the far field approximation. The far field approximation states that the noise source is far enough away from the observation points that the wavefront may be considered to be a straight line (rather than a circular arc). The less commonly used near field approach considers the propagating wavefront to be curved thus being more accurate. This additional accuracy becomes of little significance where the source is far from the microphones.

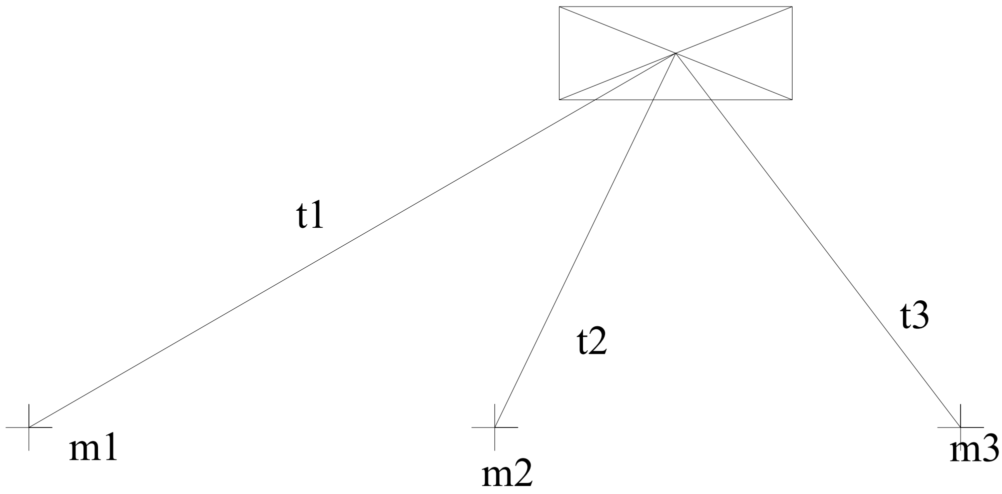


Figure 4.2: 3 microphone array and sound source

4.2.1 Far Field Approach

It is generally accepted that a source is in the far field and that the wavefront may be considered as a straight line where

$$r > \frac{2D^2}{\lambda} \tag{4.1}$$

r being the distance from the array to the source, *D* being the aperture of the array and λ being the wavelength of the signal.

The primary use of the TDOA in a far field application is the determination of the angle of incidence of a waveform. The approximate angle towards the source can be found using the delay and distance between two microphones as stated in the equation

$$\cos\theta_i = \frac{c\,\tau_d}{d_{mic}} \tag{4.2}$$

where θ_i is the angle to the source (or the waveform incidence angle), τ_d is the TDOA between the two microphones and d_{mic} is the distance between the microphones (or the microphone aperture).

The geometry of this technique is illustrated below

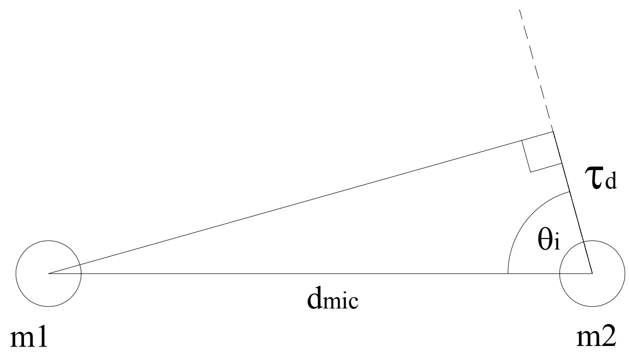


Figure 4.3: TDOA used to find angle to source

The angle to the source can often provide adequate location information however if a position is required, multiple microphones pairs may be used to provide a point at the intersection of signal incident lines. When the true nature of sound propagation is shown using a circular propagation pattern, the inaccuracy of this technique can be seen in Figure 4.4.

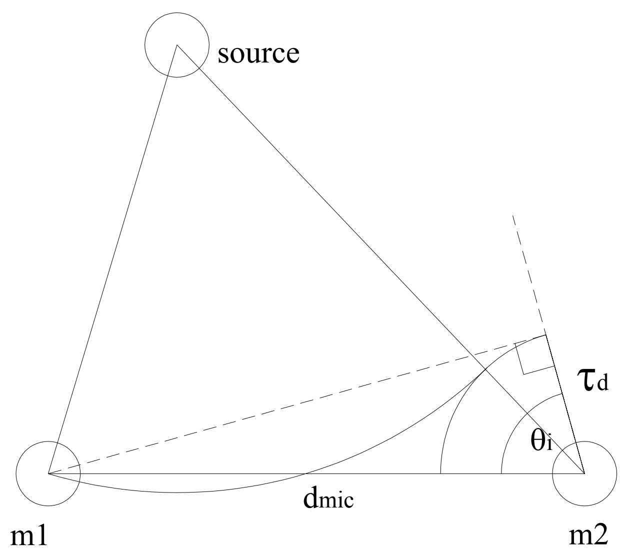


Figure 4.4: Far field vs. near field localization

In the figure above, the arc passing through microphone 1 shows the circular propagation of the acoustic wave front. The same time delay has been used by marking the distance τ_d around microphone 2. It is seen that the true angle of incidence is smaller than that estimated assuming a linear wave front (dashed). The above figure shows the inaccuracy of the plane wave approximation where the source is relatively close to the array. Such errors will become insignificant where the source is greater than the distance r in equation 4.1.

The equation states that in observing signals containing frequencies up to 3000Hz at the standard distance of 7.5m, the maximum array aperture will be approximately 0.65 meters (for the purposes of this investigation, a frequency range of 50Hz to 3000Hz has been considered as the desired analysis range).

4.2.2 Near Field Approach

The near field approach to source localization considers the signal wavefront to be curved. A practical solution to solving the location of an acoustic source using the time delays between microphone pairs is described below. The technique relies on three microphones of known co-ordinates (m1, m2 and m3) which observe a sound source at differing times due to the sound propagation delay. Figure 4.5 shows how the time difference of arrival may be used to construct a series of circles that define the location of the source relative to the microphones.

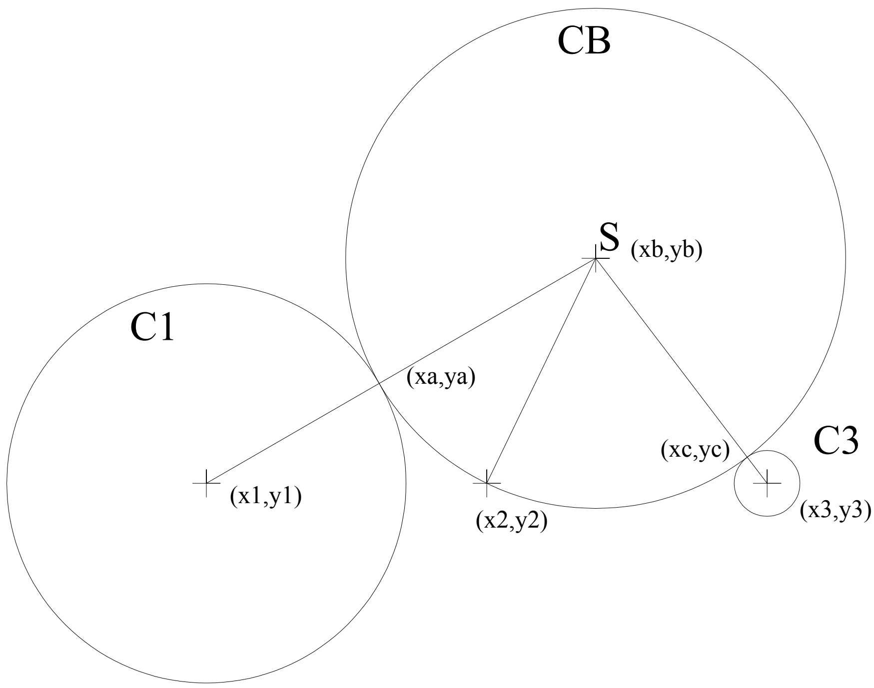


Figure 4.5: Near field localization of source

In Figure 4.5 above, the radius of C1, let us say rC1 is the time difference of arrival between microphone 2 and microphone 1. Likewise, the radius of C3, rC3 is the time difference of arrival between microphone 2 and microphone 3. Here m2 is the origin microphone, however, such a diagram may be created using m1, m2 or m3 as the origin microphone (the microphone for which all time delays are relative).

The objective now in locating the source is to determine equation of a circle CB which passes through points (xa,ya), (x2,y2) and (xc,yc) and is centred at (xb,yb). From the set of circles, a set of simultaneous circle equations may be derived.

Using the general expression for the equation of a circle:

$$(x_n - x_0)^2 + (y_n - y_0)^2 = r_n{}^2 \tag{4.3}$$

Where the point (xn, yn) lies on the circle centred at (x0,y0) with a radius of rn.

$$(x_a - x_1)^2 + (y_a - y_1)^2 = r_{C1}{}^2 \qquad (1)$$
$$(x_c - x_3)^2 + (y_c - y_3)^2 = r_{C3}{}^2 \qquad (2)$$
$$(x_a - x_b)^2 + (y_a - y_b)^2 = r_{CB}{}^2 \qquad (3)$$
$$(x_c - x_b)^2 + (y_c - y_b)^2 = r_{CB}{}^2 \qquad (4)$$
$$(x_2 - x_b)^2 + (y_2 - y_b)^2 = r_{CB}{}^2 \qquad (5)$$

These equations describe the time difference circles C1 (1) and C3 (2) as well as the source centred circle CB (1, 2 &3). The solution to the unknown points on the circles may be found through parameter optimization. By using a simple numerical code, the unknown coordinates may be varied through a select range. The most likely set of coordinates will provide a minimal optimization error in the set of rearranged equations below.

$$err_1 = (x_a - x_1)^2 + (y_a - y_1)^2 - r_{C1}{}^2 \qquad (1a)$$
$$err_2 = (x_c - x_3)^2 + (y_c - y_3)^2 - r_{C3}{}^2 \qquad (2a)$$
$$err_3 = (x_a - x_b)^2 + (y_a - y_b)^2 - r_{CB}{}^2 \qquad (3a)$$
$$err_4 = (x_c - x_b)^2 + (y_c - y_b)^2 - r_{CB}{}^2 \qquad (4a)$$
$$err_5 = (x_2 - x_b)^2 + (y_2 - y_b)^2 - r_{CB}{}^2 \qquad (5a)$$

Hence the most likely estimation of the unknown coordinates will give

$$err_T = err_1 + err_2 + err_3 + err_4 + err_5 = Minimum \qquad (4.4)$$

There appear to be three unknown points on the set of equations which must be substituted with various estimations for the optimization process. These points being (xa,ya), (xb,yb) and (xc,yc). However, with knowledge of the time delays between the microphone pairs and the speed of sound, the distances rC1 and rC3 are easily calculated. Thus points (xa,ya) and (xc,yc) may be derived from the coordinates (xb,yb) as shown below.

$$x_a = (x_b - x_1)\frac{r_{C1}}{r_{C1} + r_{CB}}$$

$$y_a = (y_b - y_1)\frac{r_{C1}}{r_{C1} + r_{CB}}$$

$$x_c = x_3 - (x_3 - x_b)\frac{r_{C3}}{r_{C3} + r_{CB}}$$

$$y_c = x_3 - (y_3 - y_b)\frac{r_{C3}}{r_{C3} + r_{CB}}$$

As these points may be calculated from the coordinates (xb,yb), being the estimated coordinates of the sound source itself, a numerical optimization process may be used that varies the values of xb and yb through a select range and hence acquires the minimum optimization error for the equations 1a to 5a.

4.2.3 Array Selection

As mentioned previously, a far field approximation used with a signal having a frequency range up to 3000Hz and a minimum propagation distance of 7.5m requires an array aperture of no more than 0.65 meters. Higher frequency components or a closer propagation distance will further reduce this value.

Apart from the accepted inaccuracy of the far field approach, localization errors are likely to occur due to correlation errors and the fact that vehicles behave quite differently to single point noise sources at a proximity pf around 7.5m. These inaccuracies will contribute to errors in the approximated angle of the far field technique thus returning a false source location. Figure 4.6 below shows how slight angular errors can affect the estimated location of a source in using the intersection of incident lines.

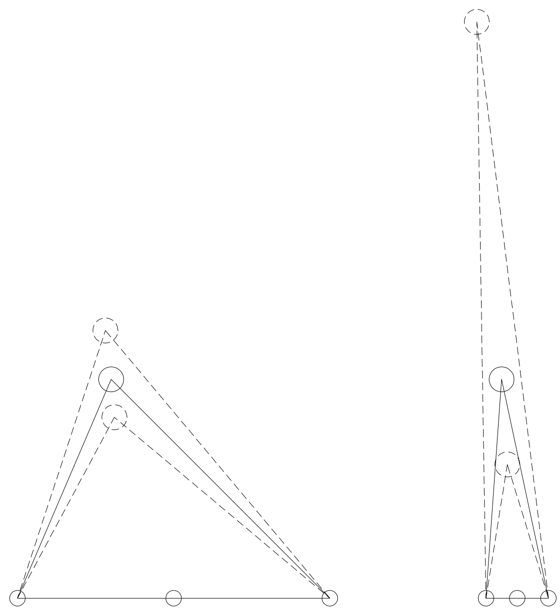


Figure 4.6: Localization errors due to far field inaccuracy

The example shows that a small array aperture will yield extremely inaccurate location values where incident angle errors are present. On a roadway, using the maximum aperture size of 0.65m, a total angular error of only 2.8 degrees would return the location of a vehicle as being in the next lane (3m further) possibly falsely identifying an un-offending vehicle. It is therefore desirable to use an array aperture as wide as possible to minimize such errors.

The approach of near field localization has been chosen for the roadside system. The method provides greater accuracy in the interpretation of the TDOA and allows for a wider array aperture which also minimizes the discussed consequential errors.

4.2.4 Accuracy of Localization

The identification of the coordinates of a sound source will improve with accuracy as the number of microphones increase. The sampling rate and the ability of the system to remove the effects of noise and interference will also improve localization accuracy. For this application, the sound source takes the form of a vehicle. It is well known that the sound created by a vehicle is not emitted from a single point on the vehicle. Therefore it is unlikely that a system could identify the single point noise sources on a car neither is it necessary. It is only required that the system can estimate the whereabouts of a vehicle in relation to the microphone array, the lane markings and other road features. This will allow for either manual evaluation of a captured image to identify an offending vehicle or a guided application of an ANPR (automatic number plate recognition) system.

Therefore it may be said that the localization algorithm is required to be accurate to a lane number and a chainage from a known point identifiable in the image for manual identification purposes. A similar tolerance would be required if another computer system were to analyse a particular area of an image for number plate recognition. Of course, in combining systems, additional road monitoring devices such as inductive loops and pneumatic traffic sensors could be used for extremely accurate localization of vehicles.

CHAPTER 5 – MICROPHONE ARRAY AND SIGNAL CORRELATION

5.1 Cross Correlation

In signals processing, correlation may be defined as the mathematically determined similarity between two signals. Cross correlation is the correlation of two different signals; autocorrelation is the correlation of a signal with itself. Correlation is calculated as a function of lag (Harris & Ledwidge). This lag can be both positive and negative and is applied to one of the signals in the correlated pair. A lag may be expressed as a time, number of wavelengths, a phase angle or in the case of discrete processing, as a number of samples. The equation for continuous cross correlation is (Leis)

$$R_{xy}(\lambda) = \int_T x(t)y(t-\lambda)dt \tag{5.1}$$

Where x and y are the two signals, t is time and λ is the lag time. The value is often calculated for all possible values of t and λ, the extents of which rely on the size of the correlated data sets.

The example figures below show how the cross correlation between discrete signals x and y is calculated as the product of a reference signal (x) and a lagged signal (y). The example shows that cross correlation is simply the multiplication of the contents of a sample set by another sample set in a time (or sample) shifted sequence

For the two discrete signals x and y comprised of four sampled values.

Lag (k) = 0

$x(0)$	$x(1)$	$x(2)$	$x(3)$
$y(0)$	$y(1)$	$y(2)$	$y(3)$

$$R_{xy}(0) = \frac{1}{4}\left[x(0)y(0) + x(1)y(1) + x(2)y(2) + x(3)y(3)\right]$$

Lag (k) = 1

$x(0)$	$x(1)$	$x(2)$	$x(3)$	-
-	$y(0)$	$y(1)$	$y(2)$	$y(3)$

$$R_{xy}(1) = \frac{1}{4}\left[x(1)y(0) + x(2)y(1) + x(3)y(2)\right]$$

Lag (k) = -1

-	$x(0)$	$x(1)$	$x(2)$	$x(3)$
$y(0)$	$y(1)$	$y(2)$	$y(3)$	-

$$R_{xy}(-1) = \frac{1}{4}\left[x(0)y(1) + x(1)y(2) + x(2)y(3)\right]$$

The example above shows how the cross correlation of two discrete signals is calculated as a function of sample lag (k). Here only a lag of +1, 0 and –1 are shown, however for the full correlation of the two sets of data one set's alignment is completely shifted past the other. The discrete cross correlation of two signals is therefore (Leis)

$$R_{xy}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)y(n-k) \qquad (5.2)$$

Where N is the number of samples and k is the sample lag. Note that correlated signals must be of equal length. In the discrete application, data sequences are usually padded with zeros where overlap occurs (for calculations at a lag other than zero). The significance of cross correlation is quite large in regards to signal processing. One very useful application is the ability to determine whether two signals are alike enough to be the same signal. The technique is used extensively in radar applications where a reflected signal is compared the original transmitted signal.

5.2 Cross-Spectrum

The discrete cross correlation of two signals has been defined as

$$R_{xy}(k)=\frac{1}{N}\sum_{n=0}^{N-1}x(n)y(n-k) \tag{5.3}$$

The frequency domain counterpart of the cross-correlation function relating two sequences $x(n)$ and $y(n)$ is called the *cross spectral density* or simply the *cross spectrum*. Rather than time delay, the cross spectrum holds phase information concerning a frequency present in two separate signals. Therefore, if two sequences have no shared frequencies, similar to their cross-correlation, their overall cross-spectrum will be approximately equal to zero. In the same manner that cross-correlation is derived from the product of the aligned components in a sequence, the cross-spectrum is found as the product of the frequency components of the Fourier transfers of the signals (Bendat & Piersol)

$$S_{xy}(\omega)=X(\omega)Y^*(\omega) \tag{5.4}$$

Where $X(\omega)$ and $Y(\omega)$ are the Fourier transforms of $x(t)$ and $y(t)$ and the * denotes the complex conjugate.

The cross-spectrum is related to cross-correlation by (Harris & Ledwidge)

$$S_{xy}(\omega)=\sum_{k=-\infty}^{\infty}R_{xy}(k)e^{(-j\omega k)} \tag{5.5}$$

And

$$R_{xy}(k)=\frac{1}{2\pi}\int_{2\pi}S_{xy}(\omega)e^{(j\omega k)}d\omega \tag{5.6}$$

In other words, the cross-spectrum is found as the Fourier transform of cross-correlation. The difficulty in obtaining the cross-spectrum lies in the need for an infinite time sequence. The use of a rectangular window in conveying correlation or time domain data can result in unwanted side lobes in the cross-spectrum (known as

Gibb's phenomenon). A reasonably reliable estimate of the cross-spectrum can be found for time limited sequences if the time domain function is tapered using a suitable window function before transformation.

From the equations 5.6 and 5.4 above, cross-correlation may expressed as

$$\begin{aligned} R_{xy}(k) &= \frac{1}{2\pi}\int_{2\pi} X(\omega)Y^*(\omega)e^{(j\omega k)}d\omega \\ &= F^{-1}\{X(\omega)Y^*(\omega)\} \end{aligned} \tag{5.7}$$

5.3 Application of cross-correlation theory

In order to determine the location of the source in the localization method described in Chapter 4, values for the time difference of arrival between the microphone pairs are required. One of the most traditional techniques for calculating the time delay between two signals employs the use of cross-correlation. The technique assumes that both signals are similar apart from scaling factors and a time shift. In the case of a simple sound field consider the signals observed at two microphones x and y described by

$$\begin{aligned} x(t) &= \alpha_x s(t) + n_x(t) \\ y(t) &= \alpha_y s(t-\tau) + n_y(t) \end{aligned}$$

Where α_x and α_y are propagation attenuation coefficients due to the different distances from the source to the microphones, $s(t)$ is the original source acoustic signal, τ is the TDOA between the two microphones and n_x and n_y are independent noise components present at each microphone. These noise components are assumed to be uncorrelated with themselves and the source signal.

The value for TDOA may be found geometrically as

$$\tau = \frac{d_x - d_y}{c} \tag{5.8}$$

where d_x and d_y are the distances from the microphones x and y to the signal source. The speed of sound c is found using the equation 2.7 in Chapter 2. The TDOA may be estimated using the classical time domain cross-correlation approach

$$R_{xy}(\lambda) = \int_T x(t)y(t+\lambda)dt \tag{5.9}$$

Then at $\lambda = \tau$

$$\begin{aligned} y(t+\lambda) &= \alpha_y s(t+\lambda-\tau) + n_y(t) \\ &= \alpha_y s(t) + n_y(t) \end{aligned}$$

and the cross-correlation function will be a maximum. Therefore τ is found as

$$\tau = \max_\lambda R_{xy}(\lambda) = \max_\lambda \int_T x(t)y(t+\lambda)dt \tag{5.10}$$

The average cross-correlation of any un-correlated signal is equal to zero. This is shown for random Gaussian noise in Figure 6.3. As the additive noise signals n_x and n_y are considered to be random un-correlated noise, they will contribute a value of zero to the cross-correlation function (both when correlated with each other, themselves and the source signal). The cross-correlation of the two simple sound field signals $x(t)$ and $y(t)$ is therefore

$$\begin{aligned} R_{xy}(\lambda) &= \int_T (\alpha_x s(t) + n_x(t))(\alpha_y s(t-\tau) + n_y(t))dt \\ &= \int_T (\alpha_x \alpha_y s(t)s(t-\tau) + \alpha_x s(t)n_y(t) + \alpha_y s(t-\tau)n_x(t) + n_x(t)n_y(t))dt \\ &= R_{ss}(t-\tau) + R_{sn} + R_{sn} + R_{nn} \\ &= R_{ss}(t-\tau) \end{aligned}$$

where R_{ss} is the autocorrelation of the source signal transposed to have its peak value at $\lambda = \tau$. As explained above, the components containing the random Gaussian distributed noise will integrate to zero. This shows how un-correlated noise will have little effect on the time delay estimation process.

5.3.1 Sample window

The comparison of the acoustic signal received by two separate microphones will clearly be limited in signal duration by the capability of the analysis system. The fact that near instantaneous correlation data is required also prevents the use of long signal sequences for time delay estimation. Theoretically, the sequence length should be long enough to provide sufficient signal alignment for a reasonably weighted correlation value. This concept is illustrated in the example in section 5.1. As the alignment of the sample windows is offset by an increasing lag, the number of sample values multiplied by zero will increase thus providing a smaller correlation set and peak value (shown simply in Figure 5.1). The sequence length must also be longer than the expected delay factor between similar data in the sequences or else no exact correlation is possible.

The range of expected time delay values may be derived from the pre-defined range of detection for the system. The limits of this range define the area in which noisy vehicles will be detected and hence define a maximum TDOA between microphone pairs. The maximum expected TDOA and indeed the maximum theoretically possible delay (being the propagation time over the aperture size) will most probably require a much smaller lag range than what is available with full length cross correlation. Performing cross correlation using this limited lag range will restrict the detection limits and save in computation.

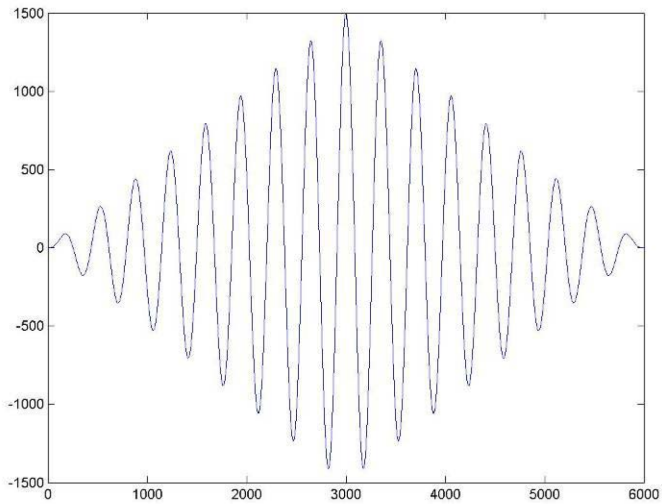


Figure 5.1: Decreasing correlation peaks for a periodic signal due to overlap

The choice of 125ms sampling windows and a microphone aperture array of less than 1m is discussed in the proceeding chapters. This chosen aperture size introduces a maximum signal delay between microphones of approximately $1\text{m}/340\text{ms}^{-1} = 2.94\text{ms}$. The available detection range in the 125ms recordings is adequate for the derivation of an un-diminished correlation peak representing the TDOA if not more than adequate. Knowledge of the maximum possible peak correlation delay time can assist in reducing the number of computations immensely. In fact, a maximum delay of +/- 2.94ms applied to the correlated signals as opposed to the full length +/-125ms applied lag reduces the number of computations by a factor of 42.

5.4 Application of cross-spectrum theory

The Fourier transform of the signal received at a microphone $x(t)$ may be expressed as (Bendat & Piersol)

$$x(t) \Leftrightarrow X(\omega) = \alpha_x A(\omega) e^{-j\omega\Delta x} \qquad (5.11)$$

Where α_x is an attenuation coefficient, $A(\omega)$ is the Fourier transform of the source signal and Δx is the propagation time between the source and the microphone. The signal received a second microphone $y(t)$ may be stated similarly as

$$y(t) \Leftrightarrow Y(\omega) = \alpha_y A(\omega) e^{-j\omega\Delta y} \qquad (5.12)$$

From equation 5.4 the cross-spectrum is then (Bendat & Piersol)

$$\begin{aligned} S_{xy}(\omega) &= X(\omega) Y^*(\omega) \\ &= \alpha_x A(\omega) e^{-j\omega\Delta x} \alpha_y A^*(\omega) e^{j\omega\Delta y} \\ &= \alpha_x \alpha_y \left|A(\omega)\right|^2 e^{-j\omega(\Delta x - \Delta y)} \\ &= \alpha \left|A(\omega)\right|^2 e^{-j\omega\tau} \end{aligned} \qquad (5.13)$$

Where $\tau = (\Delta x - \Delta y)$ and is the value of time difference of arrival (TDOA) between the microphone pair. The component of the source Fourier transform squared is also known as the power spectrum or the auto-spectrum of the source. The above equation may therefore be further simplified as

$$S_{xy}(\omega) = \alpha S_{AA} e^{-j\omega\tau} \tag{5.14}$$

Applied to the context of this project, if we consider again a simple sound field with the signals received at two microphones

$$x(t) = \alpha_x s(t) + n_x(t)$$
$$y(t) = \alpha_y s(t-\tau) + n_y(t)$$

The value for cross-correlation given by

$$R_{xy}(\lambda) = F^{-1}\{S_{xy}(\omega)\} \tag{5.15}$$

will therefore be a maximum when $\lambda = \tau$.

The reason for the use of the frequency domain information of signals in cross-correlation has come about due to computational efficiency in some applications and its robustness in the presence of certain interference (Harris & Ledwidge) Both the time domain and cross spectrum correlation methods are simulated and assessed in Chapter 7.

5.5 Cross correlation in a noisy environment

To some extent, the simple sound field model described will be spoiled by external noise and reverberations in a reflective environment. The presence of additional correlated information in the two acoustic signals (having a different peak delay) can reduce the accuracy of the system.

In the presence of a noisy vehicle as well as a number of other vehicles, it would be expected that among the several time delays provided by the cross-correlation process, the greatest value would belong to the noisy vehicle. There will be a point however where the addition of other point source signals to the dominant source signal will corrupt the signal beyond accurate time delay estimation. This is shown in Chapter 7 along with appropriate techniques for ensuring system accuracy.

CHAPTER 6 - POSSIBLE FEATURE EXTRACTION TECHNIQUES

There are many applications and studies that focus on being able to interpret sound signals to classify the source. Such techniques usually use a number of primary defining feature classes and fuzzy logic systems to group signals into likely source classes. Here more detailed analysis can be performed which is specific to the sound class. Often thousands of known time signals or frequency spectra are stored to be used for comparison. The aim of this project is to investigate the possible use of signals processing for source recognition rather than classification. While a sound based vehicle classifier would be of great use to road authorities, the research procedures and resources lie beyond the capabilities of this project. More importantly, the system being discussed has a greater need to identify vehicles that belong to a certain sound class rather than classify them even if they are not offending. For this reason, the research in this section has been oriented towards vehicle operation recognition rather than vehicle classification. More specifically, the recognition of heavy vehicles using compression braking has been chosen, both as a research example and due to the practical priority in the field.

With the aim of classifying or identifying noise sources by their acoustic emissions, appropriate features that describe the spectral and temporal structure of the sounds should be used. Obviously these features will differ according to the type of noise being classified and a priori knowledge of the analysed noise type can be used in their selection. In this case, the type of noise being discriminated is vehicle noise and the resolution of classification aims in particular to isolate heavy vehicles using compression braking. There exist a large number of attribute comparison techniques by which noise signals may be classified. These techniques can draw information from the short time behaviour or the long lime behaviour of a signal. Short time analysis may examine a signal window as small as 20ms where long lime behaviour can extend up to one second. The signal features selected for investigation in this report exhibit both short and long time analysis properties as well as analysis of the time and frequency domain components of a signal.

The possible features for source classification that have been selected for investigation are:

- Zero-crossing rate (ZCR)
- High zero crossing rate ratio (HZCRR)
- Noise level modulation rate (NLMR)
- Low and high short time energy ratio (LSTER & HSTER)
- Spectrum flux
- Noise frame ratio
- Third octave spectra comparison
- Time domain energy envelope

6.1 Zero-Crossing Rate (ZCR)

Zero-crossing rate is a useful feature in signal classification used in a variety of speech and music algorithms. A zero-crossing is the occurrence of a sign change in the amplitude of the audio signal. The zero-crossing rate is therefore the number of zero crossings within a certain time period. In most applications, the ZCR is given as signal crossings per second. The ZCR may be defined by (Lu, Jiang & Zhang)

$$ZCR = \frac{1}{2T}\sum_{n=0}^{N-1}\left|\operatorname{sgn}(x(n+1)) - \operatorname{sgn}(x(n))\right| \qquad (6.1)$$

Where T is the duration of the signal

$$T = \frac{N}{f_S} \qquad (6.2)$$

N is the number of samples in the signal and f_S is the sampling frequency. And sgn(x) is the sign of the signal amplitude (giving -1, +1 and 0).

The Zero-crossing rate of the acoustic signal emitted by a passing vehicle could be derived from either the long time or short time noise signal. The inconsistent nature of long time noise signals will reduce the discriminative ability of the classifier. If short time periods are used, a more select value for vehicle operation identification can be used. In practice, this infers that a vehicle may be classified by short bursts of particular noises rather than the accumulated noise events that occur in a pass by. Short time intervals may range from 20ms to 125ms whereas long time signals could be up to a few seconds long.

6.2 High Zero-Crossing Rate Ratio (HZCRR)

In many applications, the variation of an extracted feature can provide more discrimination than the feature by itself. Such is the case with zero-crossing rate. A *high zero-crossing rate ratio (HZCRR)* may be defined as the ratio of the number of frames whose ZCR are above a certain product of the average ZCR in a longer time window. The lengths of the frames and the longer window are dependant upon which type sound information is being targeted. The *HZCRR* for a threshold level of 1.5 times the average *ZCR* is found by (Lu, Jiang & Zhang)

$$HZCRR = \frac{1}{2N}\sum_{n=0}^{N-1}\left[\operatorname{sgn}\left(ZCR(n) - 1.5avZCR\right) + 1\right] \tag{6.3}$$

Where the average ZCR is taken from the frames as

$$avZCR = \frac{1}{N}\sum_{n=0}^{N-1} ZCR(n) \tag{6.4}$$

Where *n* is the frame index, *N* is the number of frames in the time window and *ZCR(n)* is the zero-crossing rate of the n^{th} frame. It is likely that the *HZCRR* for relatively constant sound signals will be low

The *HZCRR* of environmental noise will vary dramatically due to the characteristics of the large number of sound components picked up. On a roadway, the environmental contributors will cause variance in the absence of a dominant sound source however the *HZCRR* may exhibit some discriminative ability where the resonance of a vehicle dominates the sound signal.

6.3 Noise Level Modulation Rate

The noise level provided by precision sound level equipment describes the equivalent continuous noise level over an integrated time period. This level will modulate according to variations of the signal energy. Depending upon the integrating time period, the noise level can give a representation of quick modulations in acoustic energy or smoother trends in a signal. The shorter the averaging time, the faster and more detailed will be the response of the Leq.

In the study of compression braking noise characteristics, the NRTC suggest that a Leq averaging time of 5ms be used in order to avoid the omission of large short time modulations. The relative magnitude and rate of these modulations in Leq or Lmax could be used in classifying vehicle type and behaviour. Figure 6.1 shows how different RMS averaging times affect the recorded noise levels for a signal.

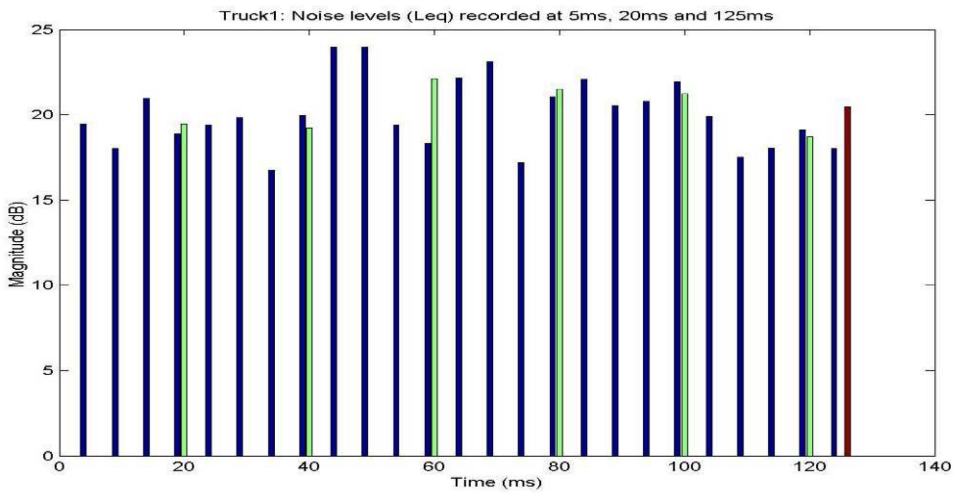


Figure 6.1: Comparison of 125ms, 20ms and 5ms Lmax Noise Levels

The NTRC report suggests a standard of feature extraction which considers the number of modulations exceeding a pre-defined value in a period of time

6.4 Low and High Short Time Energy Ratio (LSTER & HSTER)

The energy contained in a sound signal varies along with the signal magnitude. A comparison of the energy held in small frames of the signal with the energy of the entire signal can give the classifier information concerning the dynamic nature of the sound. Both low and high short time energy ratios represent the variation of short time energy. The low short time energy ratio (*LSTER)* may be defined as the ratio of the number of short time frames whose *STE* is less than a certain factor of the average energy in a longer time window. The *HSTER* is the same but for the frames whose *STE* is greater than a product of the average waveform energy.

The *LSTER* of a signal using a threshold of 0.5 times the averaged energy is given by (Lu, Jiang & Zhang)

$$LSTER = \frac{1}{2N}\sum_{n=0}^{N-1}\left[\operatorname{sgn}\left(0.5avSTE - STE(n)\right)+1\right] \tag{6.5}$$

Where the average *STE* over the signal is

$$avSTE = \frac{1}{N}\sum_{n=0}^{N-1} STE(n) \tag{6.6}$$

Where *n* is the frame index, *N* is the number of frames in the time window and *STE(n)* is the short time energy of the n^{th} frame.

Short time energy by itself does not provide any classification guidance. The energy in a window of sound will vary depending on the loudness of the source and the distance to the observation point.

6.5 Spectrum Flux (SF)

Another feature used in general audio classification applications is Spectrum Flux (SF). This is given as the average variation of spectrum between adjacent frames in a waveform. Clearly the procedure will yield different results with the use of various frame sizes and signal lengths. The *SF* may be obtained using different resolutions of spectrum, hence giving a range of detail. The general formula for obtaining the spectrum flux of a signal is (Lu, Jiang & Zhang)

$$SF = \frac{1}{(N-1)(K-1)}\sum_{n=1}^{N-1}\sum_{k=1}^{K-1}\left[F(n,k) - F(n-1,k)\right]^2 \tag{6.7}$$

Where *F(n,k)* is a discrete frequency spectrum of the signal frame, *N* is the number of frames and *K* is the order of the spectrum (the number of discrete frequency bins). As stated above, the frequency spectrum may be of any desired resolution and range. It would seem appropriate to investigate the use of *SF* with the standard 1/3 octave spectrum provided by the majority of sound level equipment. Such feature extraction would not be labour intensive concerning computation and uses readily available captured data.

6.6 Noise Frame Ratio (NFR)

As explained in Chapter 5, the auto correlation of a signal can enhance and assist the detection of signal periodicity. A completely periodic signal will have relatively high local correlation values (Figure 6.2) whereas a non-periodic signal such as white noise will have no significant correlation peaks besides the one that indicates zero sample shift (Figure 6.3). As a signal is subject to more and more noise, the magnitude of the accessory periodic correlation peaks will decrease (Figure 6.4).

In feature extraction, the *NFR* is defined as the ratio of noise frames in a signal. A noise frame is a sample frame whose maximum local autocorrelation peak is below a pre-set value. The *NFR* depends heavily on how noisy a signal is. It may be said that signals with a low *NFR* are likely going to be more easily classified than those with a high *NFR*. Using the same technique for extracting the *NFR* from a signal, the general signal periodicity may be found. This could also prove to be a useful feature in source classification.

Autocorrelation of a sampled signal is defined as the product

$$R_{xx}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)x(n-k) \qquad (6.8)$$

where k is lag or delay.

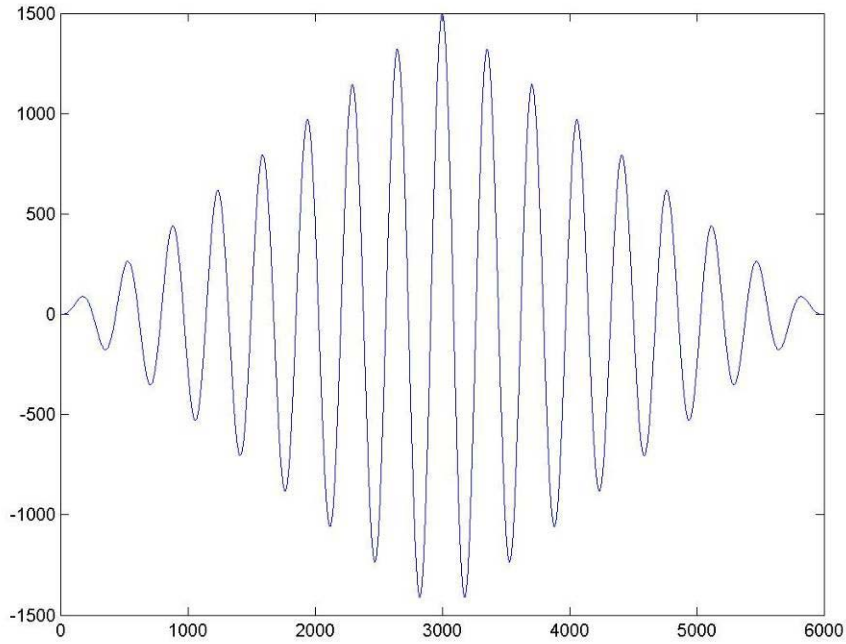


Figure 6.2: Autocorrelation of periodic waveform (sine wave)

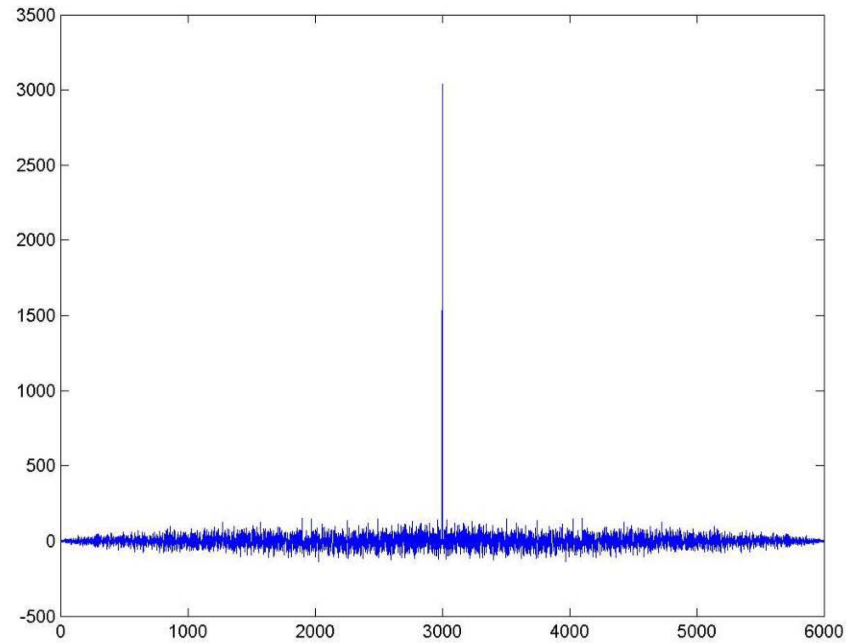


Figure 6.3: Autocorrelation of white noise

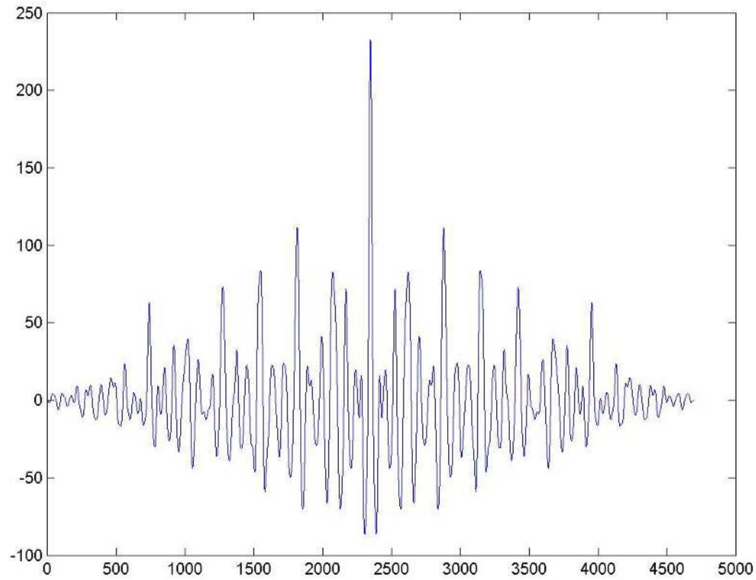


Figure 6.4: Autocorrelation of semi periodic waveform (Truck01.wav)

6.7 Third Octave Spectra Comparison

The conversion of a signal into frequency components allows for a more extensive range of identification and comparison features. The most common form of spectrum representation in the field of sound level measurement is the 1/3 octave spectrum. The 1/3 octave spectrum gives the RMS averaged sound pressure levels over the range of frequencies that lies within one third of an octave of the frequency spectrum. There are approximately ten octaves between the limits of audible frequencies 20Hz and 20,000Hz. There are therefore a possible 30 1/3 octave bands that lie within this range. The most common reference points for 1/3 octave bands can be seen in the figure below.

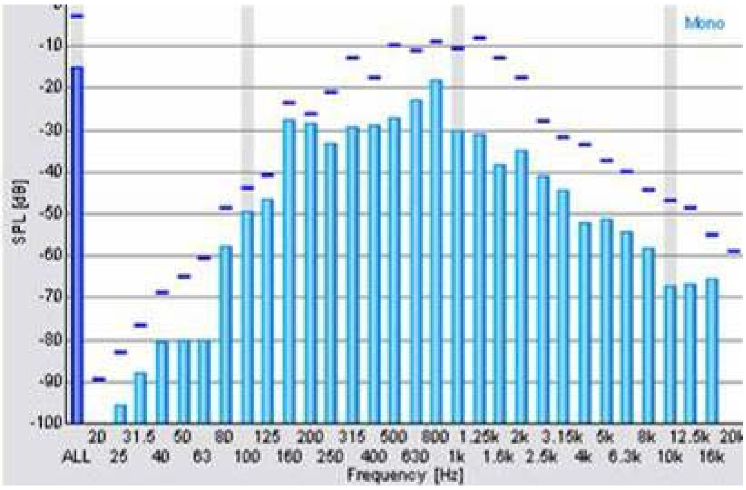


Figure 6.5: 1/3 octave spectrum from 20Hz to 20 kHz

If we were to use a 1/3 octave spectrum to identify a source, the source would need to produce a fairly stable spectrum over time or at least produce a similar spectrum periodically. The identification of a source using the spectrum would entail direct comparison between the features of a reference spectrum and an observed spectrum.

6.7.1 Direct Comparison

One obvious method of comparison would be to compare the levels of each of the frequency bands of the normalized spectra from the road and a reference spectrum. If the two spectra are alike, then the probability of classification is increased. One particular method used in sound classification applications is where a signal's 1/3 octave spectra is compared to a library of known spectra. From this, the closest match can be assigned for classification. In the case of roadside classification, we desire to identify if a characteristic noise event occurs rather than to classify every noise event. The spectrum library approach may here be useful where only reference spectra of desired noise types are included (i.e. spectra that represent compression braking).

A simple comparative method here would be to use the sum of the difference magnitudes between each band in two 1/3 octave spectra. A small value suggests a high similarity whereas a large difference value rules out correlation. If the 1/3 octave spectrum is continually calculated using small time periods, short busts of characteristic noises can be isolated. This means that if a passing truck using compression braking has its acoustic signal contaminated by other environmental noise, a "clean" segment may still be detected when the noise backs off.

It may be the case that one short time 1/3 octave spectrum is not enough to represent a noise type. Here several reference images could be utilized.

6.7.2 Feature Extraction

Rather than comparing the exact images of source spectra, a classifier could compare other characteristics of a spectrum to those of the reference spectrum. Such characteristics could include

- level ratios between different bands
- ratio between the total noise level and band levels
- what bands have the highest levels
- the amount of bands that lie within a certain level percentage

These characteristics would be best derived from normalized spectra. In this way, signals of varying noise levels (due to distance, attenuation or low noise emission) could be compared without bias.

6.8 Time Domain Energy Envelope

This feature extraction technique works primarily with long time signal information. This is because the signal from an entire vehicle pass by is required for analysis. The time domain energy envelope is the energy produced by a source over a certain time period. In this application, the energy waveform produced by a vehicle as it passes by an observation point may be useful in source classification or recognition. The vehicle energy E is calculated as (Sampan)

$$E = \frac{\sum_{k=1}^{N} s^2(k)}{N} \tag{6.9}$$

Where s(k) is the k^{th} sampled signal and N is the number of samples in the window. A typical energy waveform for a passing vehicle is shown in Figure 6.6. The figure also shows the boundaries of feature extraction regions. The features that could be used for signal recognition are listed below.

1. Maximum value of average energy
2. Approximated location of centroid of energy from starting point
3. Location of the peak value of energy from the starting point
4. Difference in centroid location and peak location
5. Max value of energy in each window (1 & 2)
6. Number of windows from SP to EP
7. Approximated location of maximum value of E/Eav in 1st region
8. Approximated location of centroid of energy in 1st region
9. Mean of energy in 1st region
10. Approximated location of maximum value of E/Eav in 2nd region
11. Approximated location of centroid of energy in 2nd region
12. Mean of energy in 2nd region
13. Number of windows having E/Eav > 1
14. Number of windows having Eav > 50% of Emax
15. Number of windows having Eav > 25% of Emax

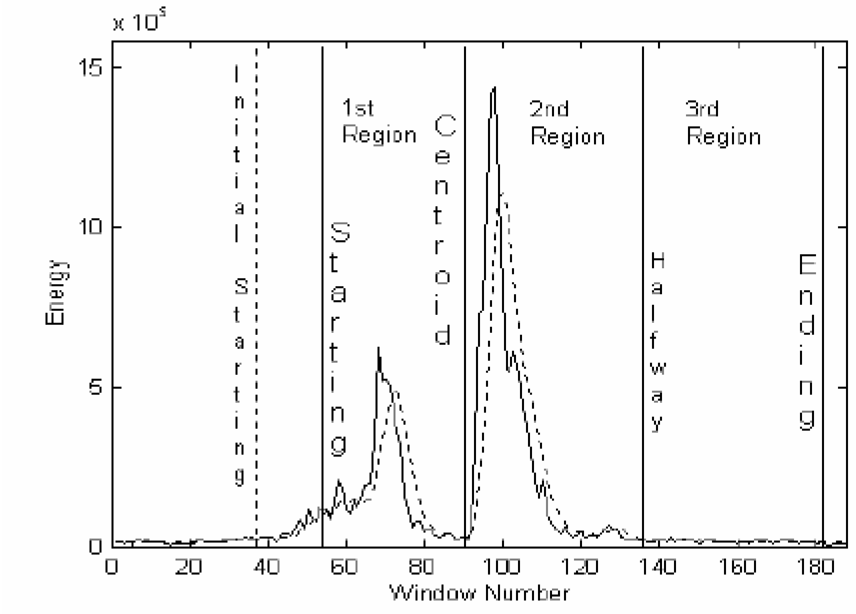


Figure 6.6: Energy waveform with feature extraction regions (Sampan)

6.9 Field Recordings

As explained at the start of the chapter, the vehicle characterization research in this project is specifically oriented towards the recognition of heavy vehicles using compression braking. The selected feature extraction techniques introduced in Chapter 5 address a wide range of signal attributes both in the time and frequency domain. In Order to test for the class separability and hence the usefulness of these features, a number of sample recordings were collected in the field and analysed according to each of the described extraction techniques.

6.9.1 Equipment

The device used to record the sample vehicle pass-bys was a Sony Portable Minidisk Recorder (MZ-NH700) configured to record un-compressed PCM (pulse code modulated) audio at 44100Hz. A standard *SHURE* dynamic cardioid voice microphone was used connected by a low noise audio cable. Figure 6.7 below shows the typical symmetrical polar sensitivity pattern for the semi directional diaphragm.

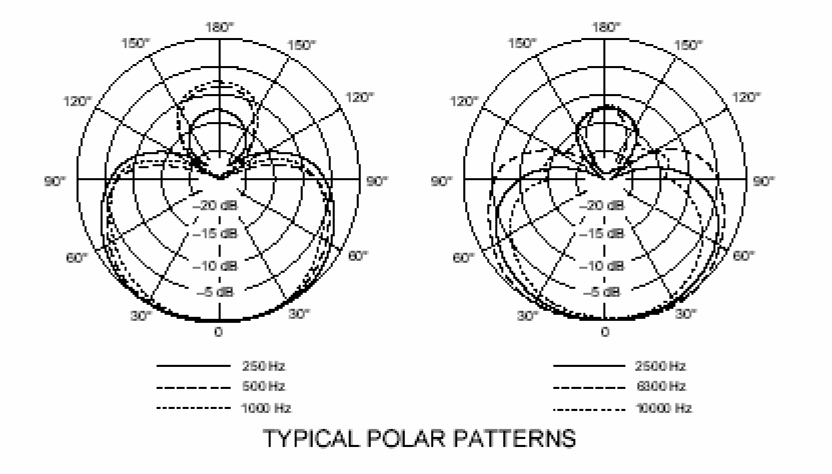


Figure 6.7: Polar sensitivity pattern for dynamic cardioid microphone (Shure Inc.)

The polar pattern shows the sensitivity roll-off exceeding 5dB for angles greater than +/- 60 degrees. This sensitivity roll-off acts as a shield from environmental noise behind and beside the microphone. It is seen however that there is a significant sensitivity lobe directly behind the microphone. This could be offset in future recordings by using acoustic shielding to shield from unwanted directions. The recorded sound files were then transferred to a PC and converted into wave files.

6.9.2 The Test Group

An analysis of suitable signal features to identify heavy vehicle compression braking must not only be successful in identifying the sound type but also in ruling out similar sound types. What is referred to here as class separability aims to recognize features common to compression braking as well as identifying features common to vehicles that clearly aren't using compression braking.

In the collection of the field recordings it was noted that the vehicle types that are most likely to sound like a vehicle using compression braking (and hence cause false identification) are passing heavy vehicles and accelerating trucks. Therefore adequate consideration has been given to the definition between these operation types in the study of feature extraction techniques.

With the issue of class separability in mind, a wide range of vehicle types and operations were captured in the form of roadside pass-bys. The selected recordings used as experimental references are described below in table 6.1 along with their observed acoustic properties.

Table 6.1: Field recording properties

Wave file	Vehicle Type and Operation	Acoustic Properties
Car1.wav	Sedan accelerating uphill	Gradual increase in frequency
Car2.wav	Diesel 4WD accelerating uphill	Clear resonant engine frequency with high freq. chatter
Car3.wav	Sedan passing downhill	Mostly low frequency exhaust sound
Car4.wav	Sedan passing downhill	Mostly low frequency exhaust sound
Car5.wav	Car Pass	Particularly noisy muffler
Car6.wav	Car Pass	Noticeable high frequency tyre on road noise
Truck1.wav	Truck using compression braking	Mostly low frequency content
Truck2.wav	Truck using compression	Noticeable shift down in frequency.

	braking	High pitch braking squeal
Truck3.wav	Truck using compression braking	Large amount of high pitch truck noise.
Truck4.wav	Truck using compression braking	High resonant frequency. Large amount of high pitch noise
Truck5.wav	Truck using compression braking	Low resonant frequency. High pitch squeal
Truck6.wav	Truck using compression braking	Very faint compression braking observable. Mostly low freq.
Truck7.wav	Truck accelerating uphill	Jittery gear operation causing noticeable modulations.
Truck8.wav	Truck accelerating uphill	Noticeable upwards freq shift. Mostly low freq.
Truck9.wav	Truck pass uphill	Loud. Constant freq.
Truck10.wav	Truck pass downhill	Mostly low freq.
Mot1.wav	Passing motorbike	Low frequency exhaust noise (chopper)

Each of the waveforms is one second in duration, selected from a longer recording to best represent the vehicle operation and to be free of irregularities. The recordings of heavy vehicle compression braking represent a range vehicles applying their compression brakes but having uniquely differing acoustic signals. This selection was made to hopefully broaden the recognition criterion and prevent only the recognition of a single type of compression braking sound. The additional truck recordings were also chosen to challenge the separability of each classification feature.

6.9.3 Recording Procedure

The recordings were taken on the northern stretch of Anzac Avenue just before the intersection with Alderly St. in Toowoomba QLD. Anzac Avenue is subject to a high amount of heavy vehicle traffic as well as passenger vehicles. The particular section used for measurements is a downhill stretch before a set of traffic lights being the cause of constant application of compression braking for southern travelling trucks.

Heavy vehicles travelling north can also be observed when accelerating uphill from the traffic lights.

The recording point chosen was situated on the side of the road closest to the north bound traffic lanes. This was not optimal as the recordings of heavy vehicles using compression braking (south bound) were further than 7.5m away. It was considered necessary however to record from this side of the road due to the large number of Camphorloral trees placed along the roadside on the opposite side. It is believed that such large obstructions would produce periodic lulls in the recordings of passing vehicles. The recording position was free from surrounding reflection sources and recording was decisively undertaken on relatively calm days.

6.9.4 Recording Precision

The importance of compliance with Australian Standards for sound level measurement is discussed in Chapter 9. One of the fundamental requirements in sound level measurement is the use of an appropriate calibrated precision sound level meter. The equipment used in recording and processing the field traffic data is not in compliance with these requirements and therefore the sound levels obtained from the recordings cannot be considered to be accurate. It is apparent in many of the sound level waveforms throughout the analysis chapters that the sound levels for the test group are well below the expected levels. Although this excludes the recordings from noise level detection applications, feature extraction and correlation data may still be obtained from the recordings. The recordings are effectively level shifted from those that would be provided by precision instrumentation.

CHAPTER 7 - CORRELATION RESULTS

7.1 Waveform Simulation

To assist in the analysis of the array configuration and time delay estimation techniques, simulated waveforms were produced to resemble those signals reaching the microphones in the suggested array. Each of the signals created are identical in frequency components but vary due to the position of the source vehicle over the signal time.

The frequency components of the signals were chosen and weighted to resemble those of a heavy vehicle using compression braking. A number of sine waves having fundamental components at 125Hz, 160Hz, 250Hz, 500Hz, 800Hz and 1250Hz were synthesized and added together to create the source signal. The apparent signal at each of the microphones was simulated by applying time, frequency and amplitude functions to the original source signal. The 1/3 octave spectrum of the simulated source signal is shown in Figure 7.1.

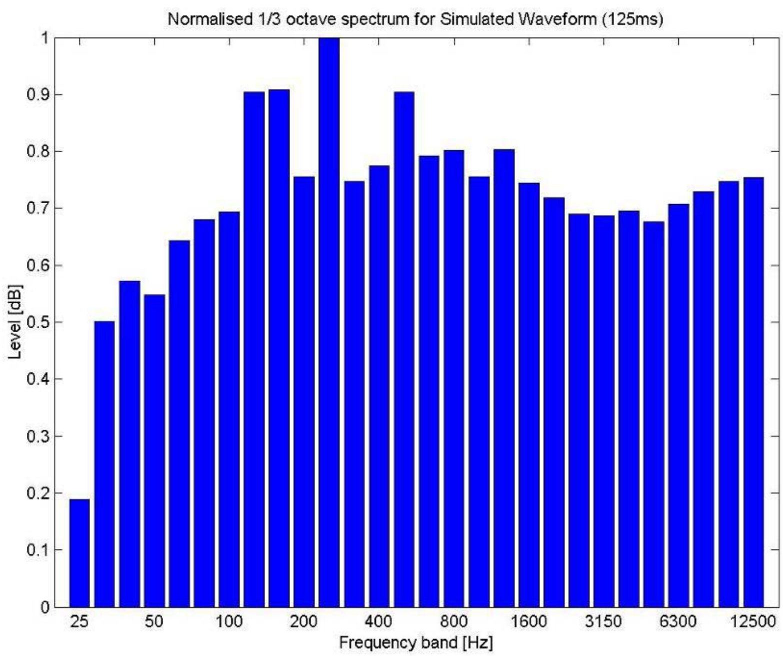


Figure 7.1: 1/3 octave spectra for simulated source waveform

The appearance of frequency components in bands other than those specified above is the result of the addition of a Gaussian distributed random noise waveform.

7.1.1 Amplitude Simulation

In the field, the sound intensity received at an observation point will vary depending on the proximity of the source. The relationship from section 2.1 can be used here to approximate the variance in signal amplitude as a vehicle passes an observation point.

$$W = 4\pi r^2 I \tag{7.1}$$

This shows that the acoustic intensity at the observation point I is proportional to the source acoustic power times $1/r^2$. This factor was applied to the generated waveform to simulate a vehicle pass by at 60km/h. the line of travel for the vehicle was offset from the observation point by 7.5 m to simulate a roadside measurement scenario. Figure 7.2 shows the attenuation factor applied over six seconds. Figure 7.3 shows the attenuation waveform applied to the simulated waveform.

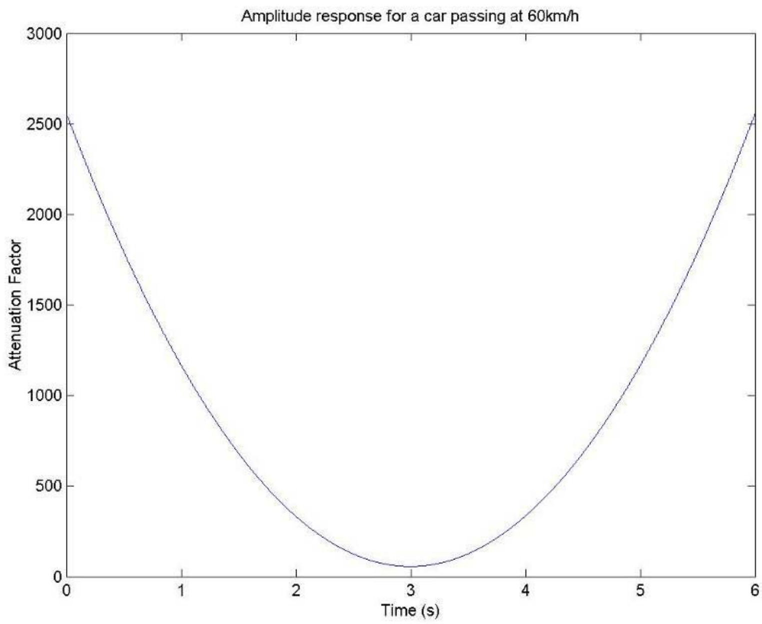


Figure 7.2: Attenuation factor for passing sound source at 60km/h

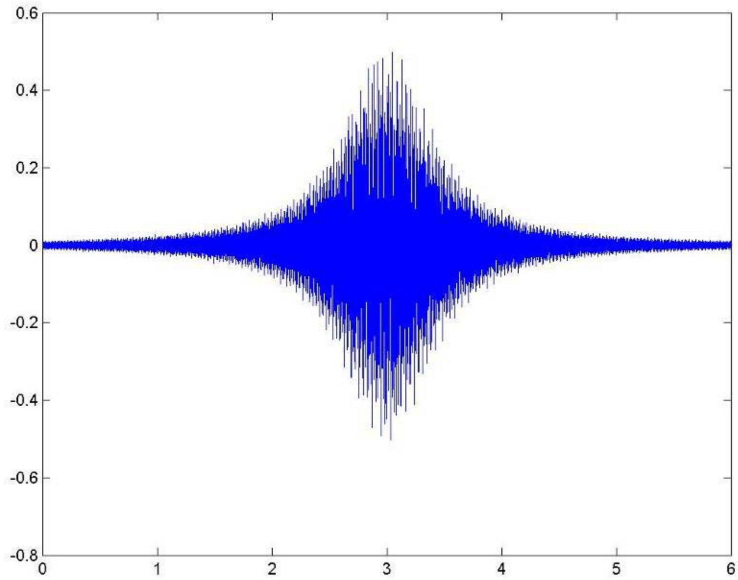


Figure 7.3: Attenuated simulated vehicle waveform

7.1.2 Time Delay

The original generated waveform was treated as the source signal. Therefore the waveforms received at the multiple observation points were obtained by applying a varying time delay dependant upon the distance between source and microphone. Figure 7.4 shows the distance between the source and microphone over the pass by.

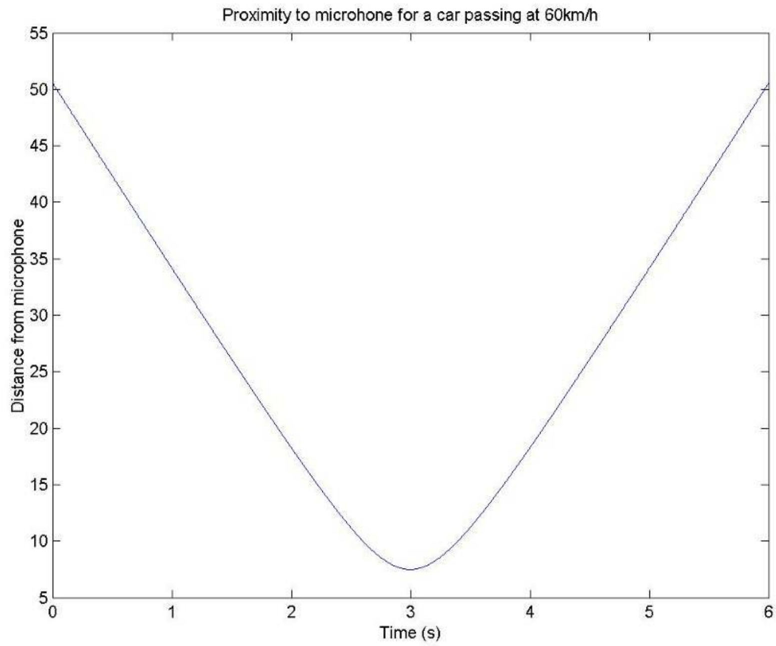


Figure 7.4: Distance between microphone and source

The time delay of the signal is then given by

$$\tau = \frac{d}{c} \tag{7.2}$$

where c has been approximated as 340m/s

7.1.3 Frequency Shift

It was decided to include the effects of Doppler shifting in the simulation to determine its effects on the correlation process. It is not so much the shift in frequency of the signal in a sample window that is of concern but the difference in frequencies of a signal reaching two microphone locations. According to equation 2.8 in section 2.1.2, the shift in frequency of a passing noise source was modelled and is shown in Figure 7.5.

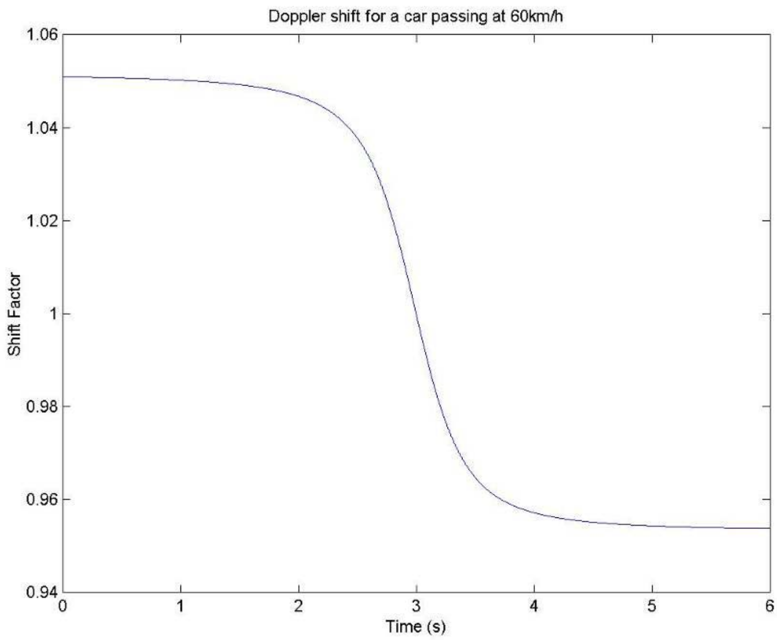


Figure 7.5: Frequency shift factor for moving source

It is seen in the figure above that the greatest shift in frequency occurs where the source passes closest to the microphone (at the 3 second mark). The 1/3 octave spectrum below shows the frequency components of the signal received where the source is stationary in relation to the receiver (i.e. sound propagation path is at right angles to the direction of travel) and hence the frequency shift is zero.

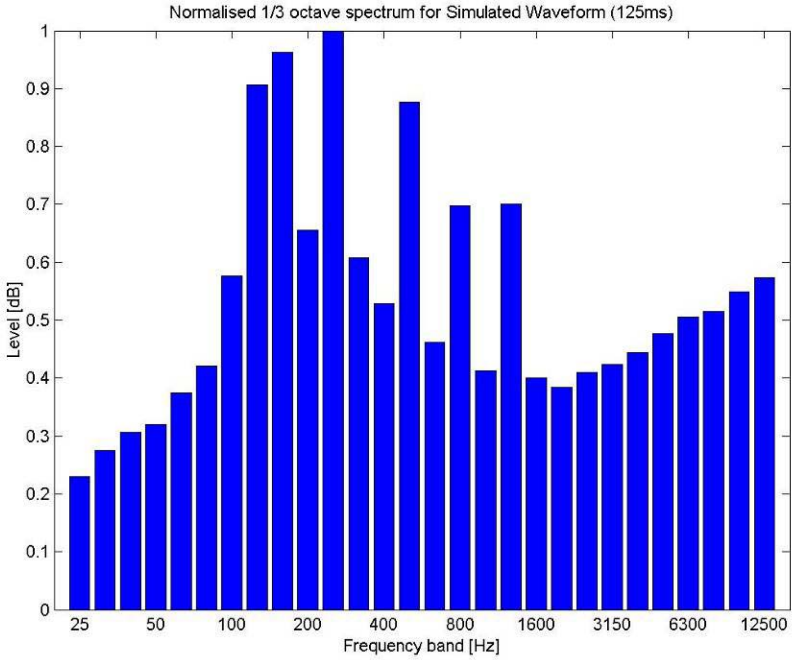


Figure 7.6: Normalized 1/3 octave spectrum of signal passing through zero freq. shift

The following figure represents the 1/3 octave spectrum of the signal received where the source is 5 meters past the observation point.

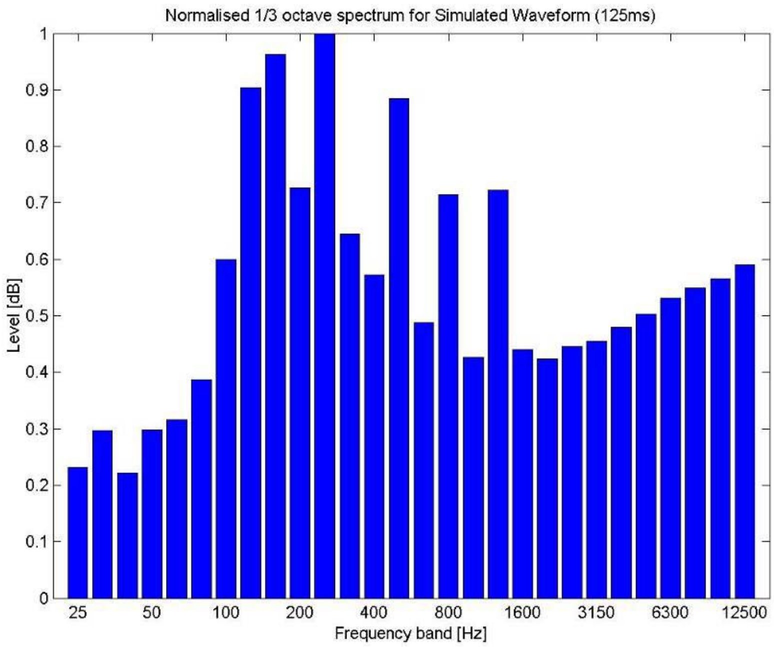


Figure 7.8: Normalized 1/3 octave spectrum of signal passing 5 meters ahead of the observer

The effective shifting is seen particularly in the lower frequency bands. It must be noted here that these 1/3 octave spectra were taken from sample windows having a length of 125ms. The source may be travelling normal to the propagation path at the start of this window, but after 125ms will have moved through approximately 2.08 meters. Frequency shift does occur therefore.

Figure 7.9 below shows a 125ms and 20ms window of the simulated waveform.

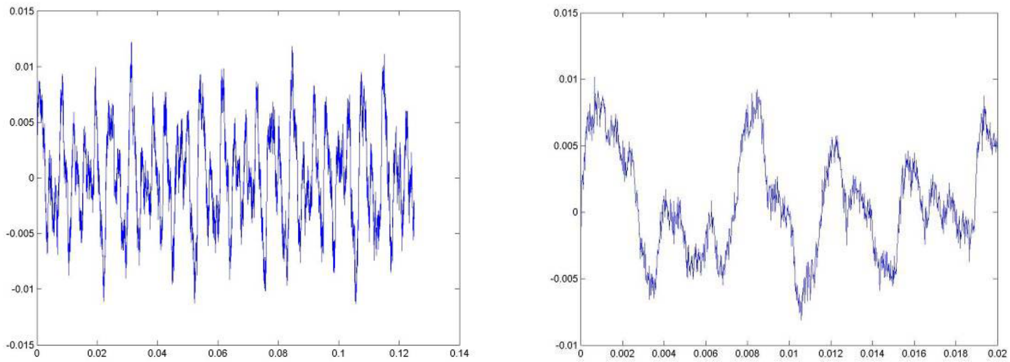


Figure 7.9: 125ms and 20ms windows of the simulated waveform

7.2 Application of the time domain cross-correlation technique

7.2.1 Experimental Technique

The technique of time domain cross correlation was applied to a number of the simulated waveforms produced using Matlab. For this group of experiments, correlation was calculated between the simulated signals reaching two microphones a certain distance *d* apart. This was done by simply correlating a signal simulated as *x* meters from a microphone and the signal simulated as being *x*-*d* meters from a microphone as outlined in the previous section. The correlated signal frame size used was 125ms; this is well beyond the requirements for finding any time delay to be interpreted by the system.

Figure7.10 below shows the cross correlation of the two simulated microphone signals where the source is passing the first microphone and is hence seen at a position *x*=0 for the first microphone and *x*=-5 for the second microphone.

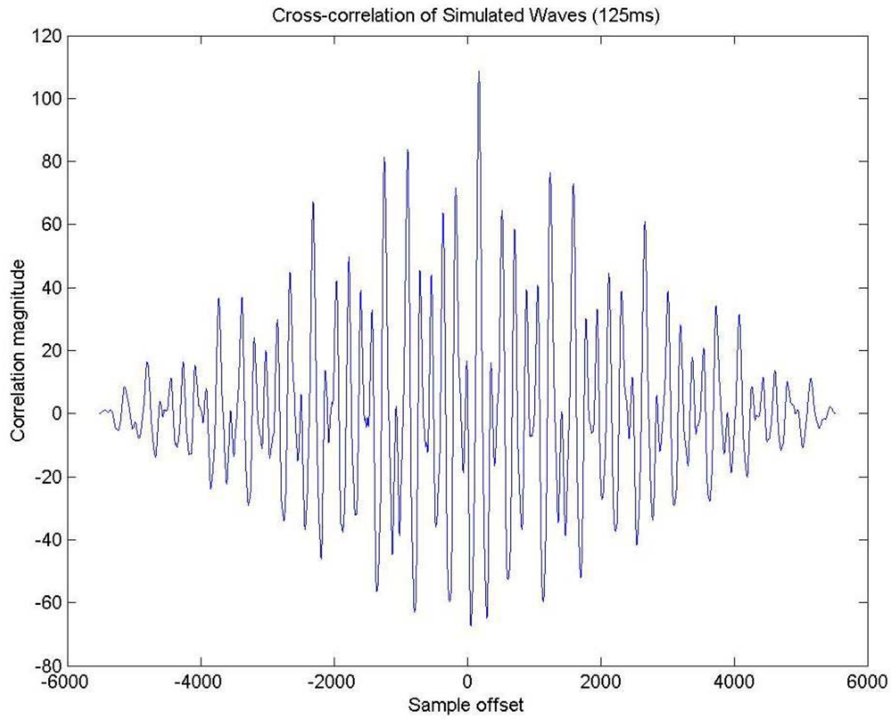


Figure 7.10: Cross-correlation of signals received for window start position x = 0

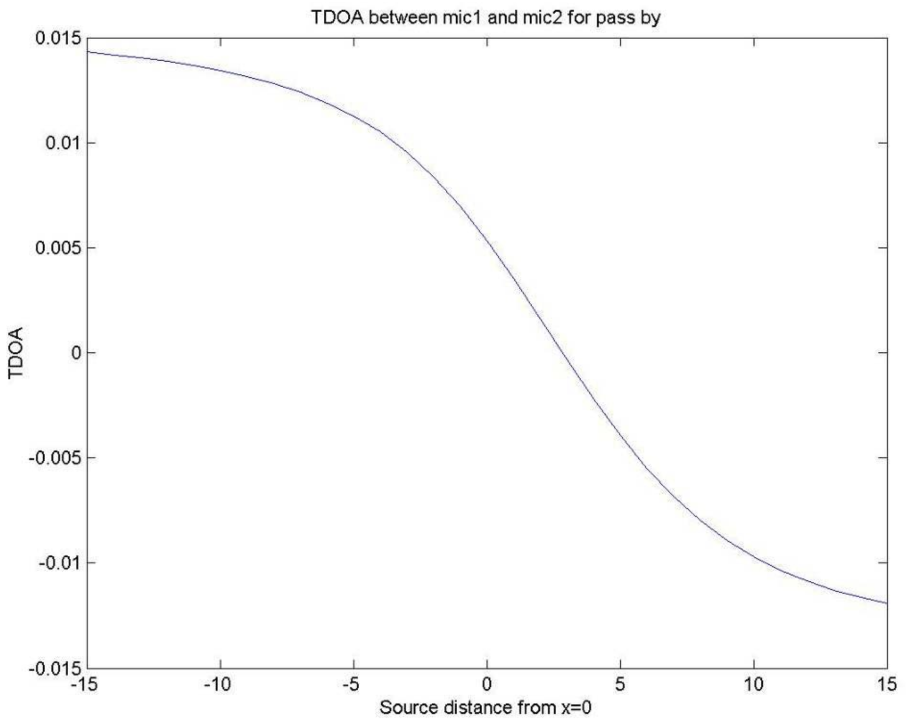


Figure 7.11: TDOA from maximum cross-correlation value

This cross correlation was conducted for signals produced as the source passed by the microphone pair where x=0 represents the horizontal position of the first microphone.

The figure above shows the theoretical TDOA between two microphones. The 'source distance from $x = 0$' on the x axis states the source position at the start of each 125ms processing frame.

As mentioned above, the source does move approximately 2.08 meters during the 125ms signal recording (60km/h x 0.125sec). The time delay between the source and observation point therefore varies also. This has been illustrated in the TDOA estimation figures where the theoretical time delays derived at the start and finish of the 125ms waveform are represented by dashed lines.

7.2.2 Doppler Shift

The frequency shifting due to Doppler shifting will lead to a slight un-correlation between the signals at different microphones representing a signal sequence from the source. This is simply due to the different relative velocity of the source with respect to the two observation points at any given time. The difference between consecutive shift factors will be greatest where the source is between two of the microphones. The frequency shift factor is calculated at this point (for an aperture of 5m)

Microphone 2.5m behind = 1.0157
Microphone 2.5m in front = 0.9847
Total shift between both signals = 0.031
This amounts to a frequency shift of 15.5Hz for a 500Hz signal.

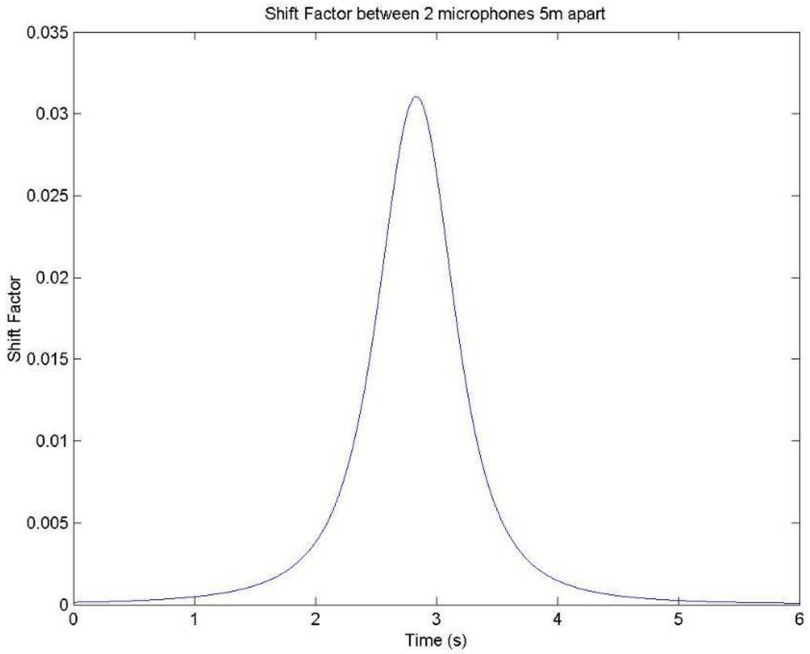


Figure 7.12: Difference in frequencies for two microphones 5m apart (vehicle passing at 60km/h)

The figure above shows the shift between the frequency components received at two microphones 5m apart as the source passes by at 60km/h. The difference is greatest where the source is directly between the two microphones (that is on the microphone axis) and has a peak value of 0.31. The further apart the microphones in the array, the greater the difference in frequency shifts and hence the signal difference will be.

As highlighted in section 8.9, the application of vehicle braking or acceleration also has an effect on the frequency components produced by a vehicle. This change in frequency will however be perceived at both microphones and will hence assist rather than impede the correlation of the two signals.

7.2.3 Signal Uniqueness

The simulated waveforms received at the microphone pair offset by the propagation time delay are identical but for amplitude difference, slight frequency shifting and the addition of uncorrelated Gaussian noise. It could possibly occur therefore that the correlation of time delayed signals could produce a correlation waveform similar to that in Figure 6.2, of the auto-correlation of a periodic signal. Here the peak correlation value is given where the two correlated data sets are completely in phase with maximum possible overlap. It can be safely assumed however that this will not be the case with the cross correlation of sound data recorded in the filed. This is due to the existence of non-periodic signal uniqueness. This signal uniqueness exists as a signal attribute from the source that is not periodic (at least within the processing window) and will be clearly correlated between the two recorded signals. Higher correlative definition due to signal uniqueness is seen in section 7.9 where actual field recordings are correlated. In order to simulate signal uniqueness, a random waveform was generated and included in each of the simulated waveforms (see code, Appendix I).

7.3 Cross-Correlation Simulation Results

The cross-correlation technique for TDOA extraction was tested using the simulated pass-by waveforms discussed above. It was immediately apparent that the correlation process yielded faulty TDOA estimations for an aperture size of 5m. In Figure 7.13 below, the theoretical time delays at the start and end of the 125ms correlated waveforms are represented by green and red dashed lines respectively. It is seen that correlation peaks occur at times other than the theoretical TDOA values.

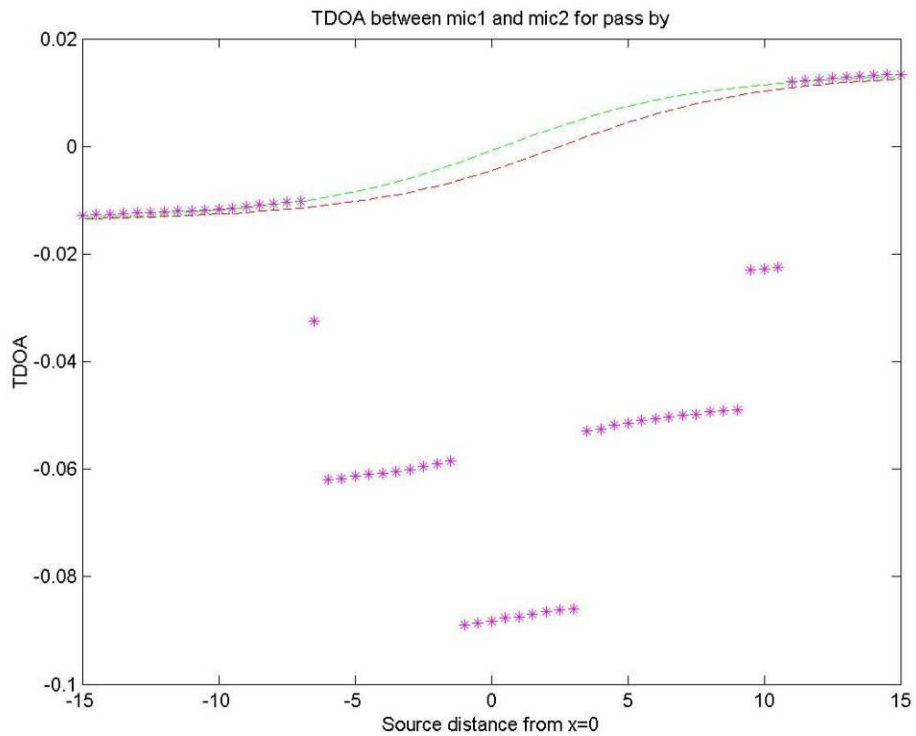


Figure 7.13: TDOA waveform of vehicle pass-by using an array aperture of 5m

The cause for these faulty values was investigated by individually removing applied attributes from the simulated waveforms. It was found that the Doppler shift had no effect on the correlation process but the applied time delay between the two signals was responsible for the false values. For a stationary source, the time delay between the signals received at different locations would remain constant and thus be easily estimated using cross-correlation. In the case of a moving source however, the time delay applied to the individual signal frequencies varies proportionally to the changing distance between the source and observer. Due to the different location of each microphone, the varying time delay over the pass-by will be different for each microphone signal.

The difference between these time delay sequences will increase with the distance between microphones. The simulation was therefore repeated using smaller array aperture values to reduce the adverse effect of differential time delay. Figures 7.14 to 7.16 below show the increase in TDOA estimation accuracy by using smaller microphone separation distances.

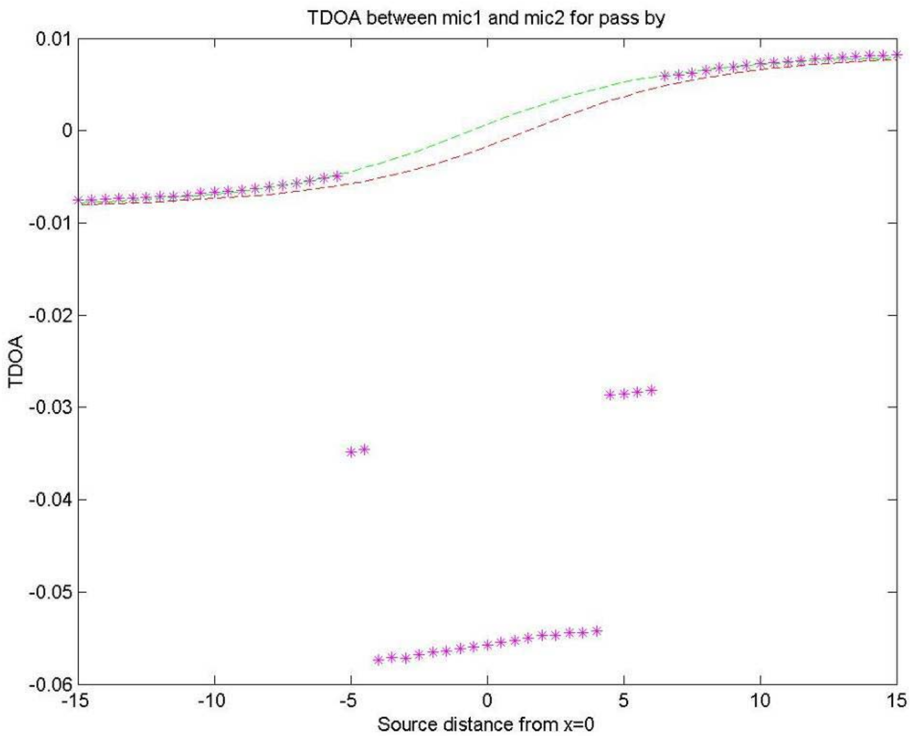


Figure 7.14: TDOA waveform of vehicle pass-by using an array aperture of 3m

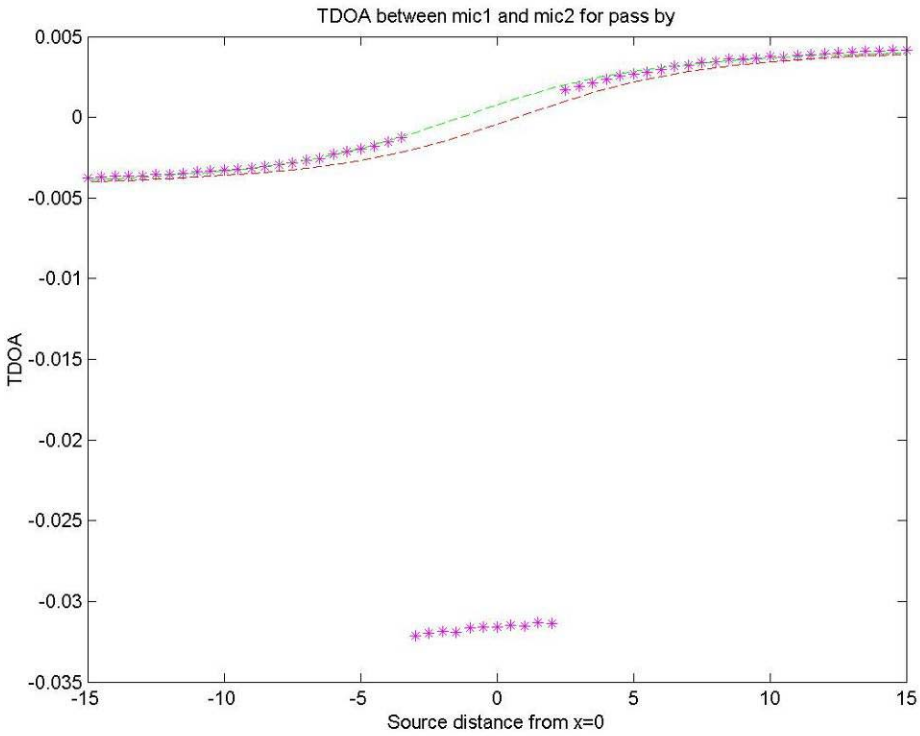


Figure 7.15: TDOA waveform of vehicle pass-by using an array aperture of 1.5m

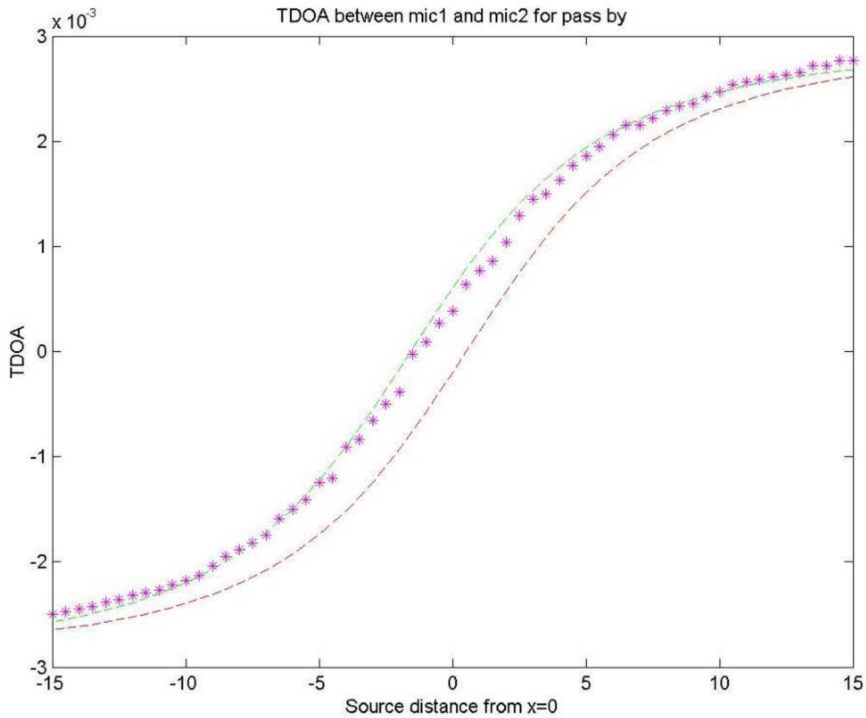


Figure 7.16: TDOA waveform of vehicle pass-by using an array aperture of 1m

The series of figures above show an improvement in TDOA estimation with the use of a small array aperture. The following simulations were therefore processed using an aperture size of 1m. Appendix G gives a comprehensive series of figures obtained using various aperture sizes and lane offsets.

7.4 Peak Correlation Magnitude

The TDOA figures above are a series of time delay values corresponding to the peak vale of the cross-correlation between two waveforms. It was seen that the peak magnitudes of the correlation waveforms were somewhat diminished where false TDOA estimations were present. Figure 7.17 below shows the low peak correlation values corresponding to the false TDOA values given for an array aperture of 5m compared to the higher values present for the accurate 1m aperture simulation.

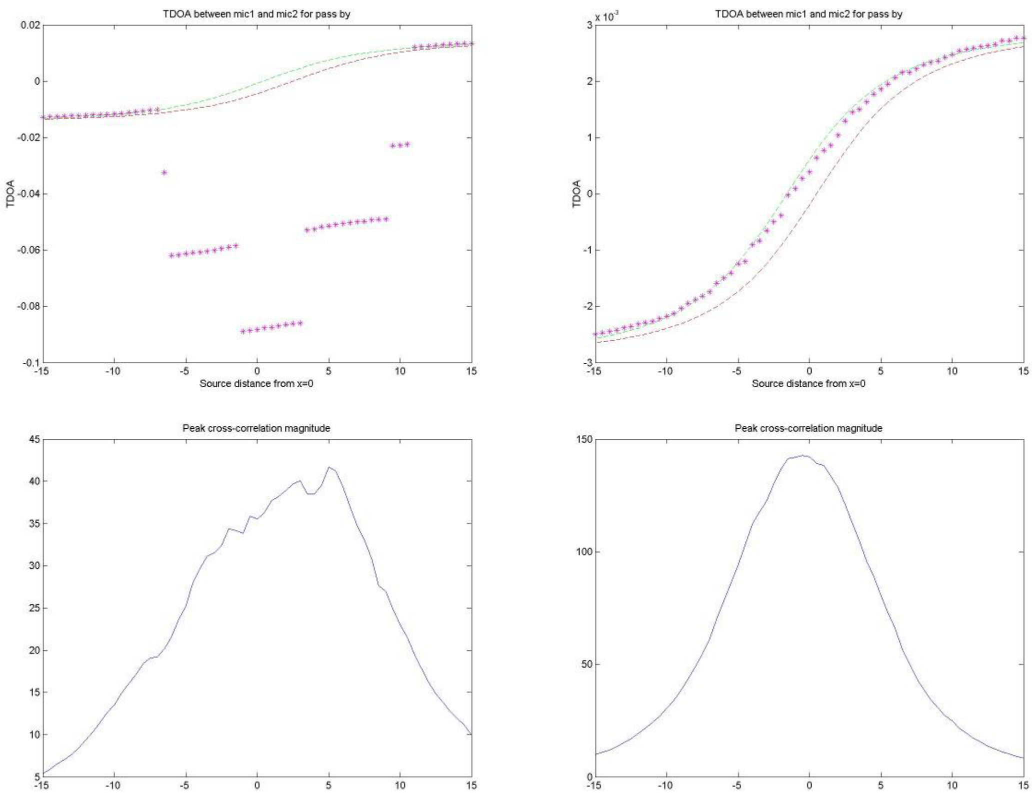


Figure 7.17: Peak correlation magnitudes corresponding to TDOA estimates. Aperture of 5m (left). Aperture of 1m (right).

The 1m aperture waveform exhibits correlation peaks above 140 where the 5m array waveform lies below a maximum value of 43. The peak correlation magnitude therefore provides a substantial indication of the accuracy of the estimated TDOA series. This property is discussed further as a location accuracy factor in Chapter 9.

The peak correlation magnitude graphs are given alongside the TDOA waveforms in Appendix G.

7.5 Multiple Lane Detection

The simulations discussed so far utilized an offset distance from the microphones to the vehicle lane centreline of 7.5m. The effect of increased lane offsets was investigated using the standard lane centreline separation width of 3.5m. The simulations show that the additional propagation distance leads to a reduction in correlation magnitudes. Figures 7.18 to 7.20 below show the peak correlation magnitudes and TDOA waveforms for vehicle pass-bys at lane offsets of 11m, 14.5m and 18m.

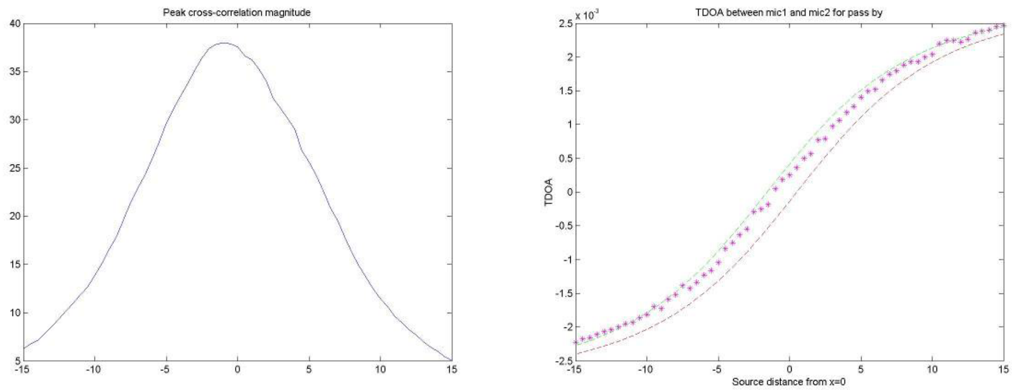


Figure 7.18: Peak correlation magnitude and TDOA waveform for a pass-by offset of 11m

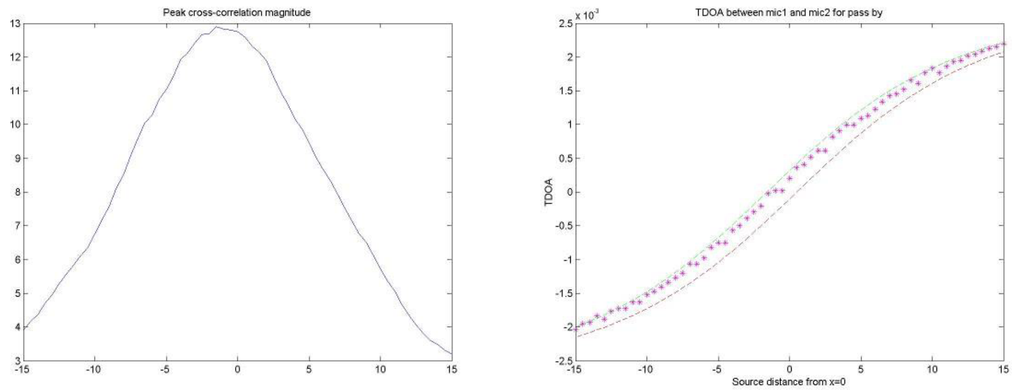


Figure 7.19: Peak correlation magnitude and TDOA waveform for a pass-by offset of 14.5m

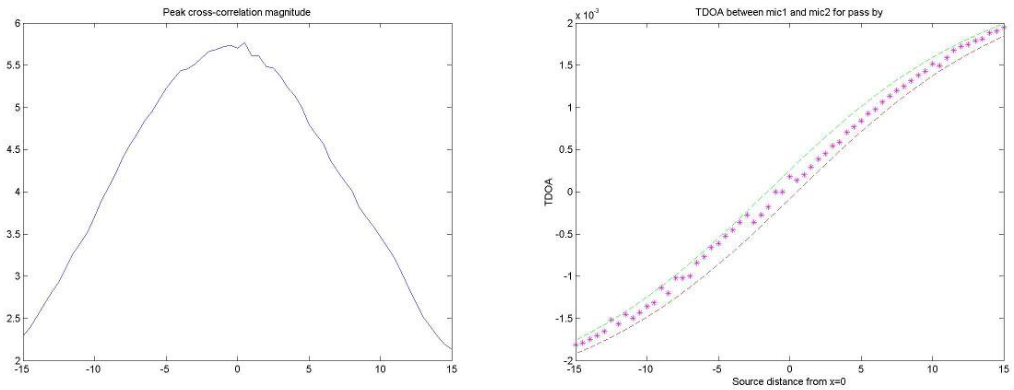


Figure 7.20: Peak correlation magnitude and TDOA waveform for a pass-by offset of 18m

The effect of signal attenuation is apparent through the reduced correlation magnitudes. This effect indicates that in the case of multiple vehicle pass-bys, the system will be most sensitive to the closest lanes. Section 7.8.1 investigates the signal amplitude requirements where noisy vehicles in the far lanes are to be detected above vehicles in the near lanes. The vast difference in peak correlation magnitudes for both lanes also suggests that a location accuracy factor will not be universal but will be specific to each lane.

7.6 Restricted Cross-Correlation Lag

It was mentioned in section 5.3.1 that the maximum possible time delay between a signal reaching two microphones can be derived from the distance between them. Using an array aperture of less than 1m limits the maximum TDOA to $1m/340ms^{-1}$ = 2.94ms. This delay will occur where the source is travelling along the axis passing between the two microphones. This can only occur in the lane offset scenario where the source is an infinite distance from the array. The true maximum delay will therefore be less where a restricted detection range is used. Figure 7.21 below shows a series of correlation waveforms using full signal length lag to determine the TDOA value. The reduction in processing is illustrated in the restricted lag correlation waveforms in Figure 7.22.

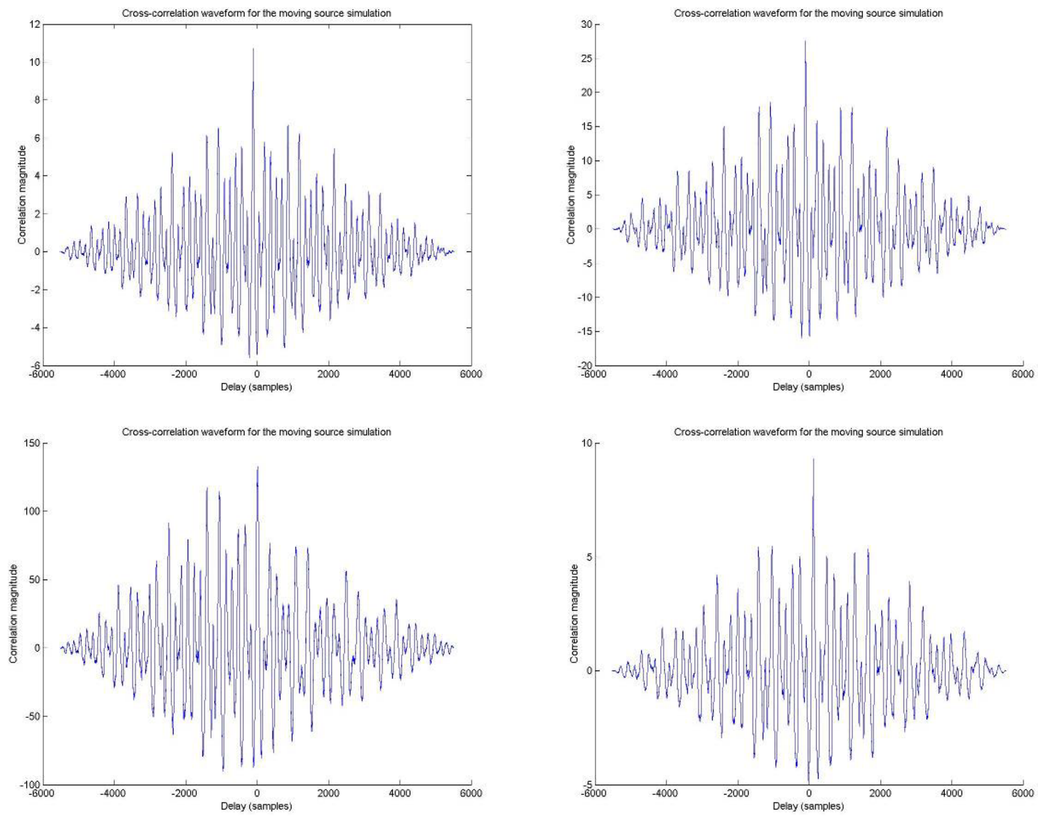


Figure 7.21: Cross-correlation series for a vehicle pass-by using a full length lag

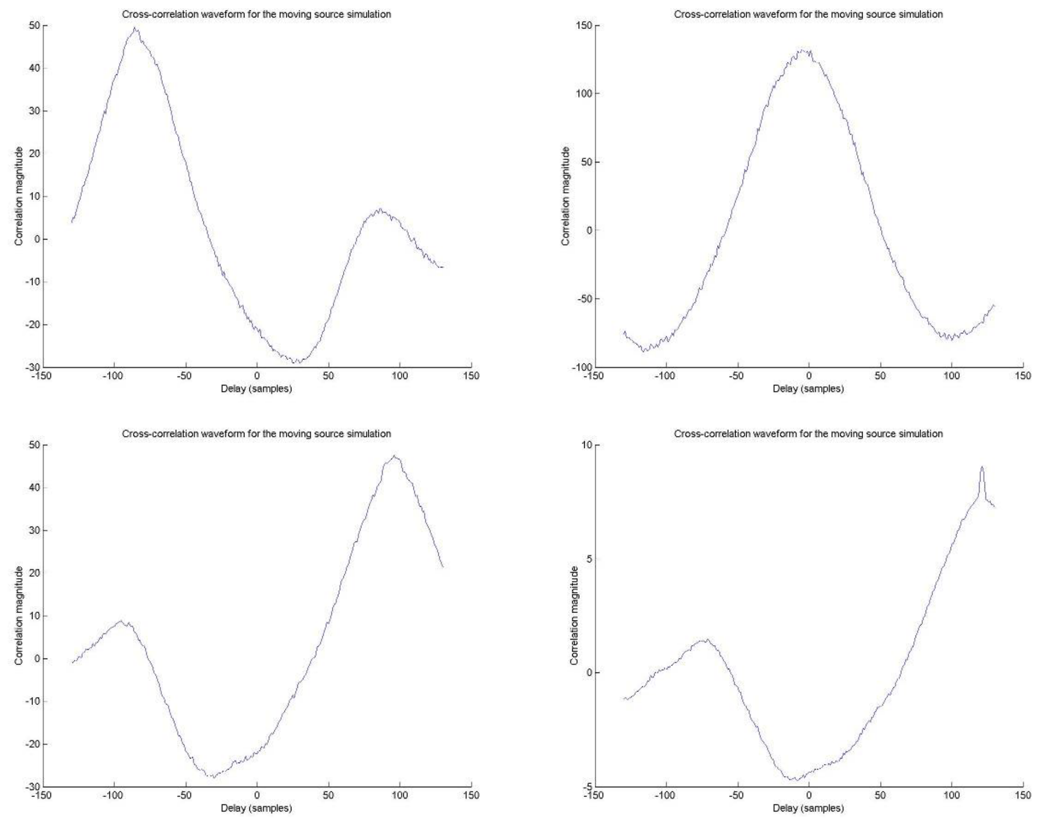


Figure 7.22: Cross-correlation series for a vehicle pass-by using restricted length lag

Both of the above figures portray the pass-by occurrence. Figure 7.21 uses 0.125 x 44100 = +/- 5513 lag calculations whereas Figure 7.22 uses 0.00294 x 44100 = +/-130 lag positions. This is an effective reduction in computations by a factor of 42.4. The Restricted correlation process also acts to inhibit the detection of correlation peaks outside the possible detection range. This is confirmed by the restricted lag waveforms in Appendix G.

7.7 Application of the cross-spectrum correlation technique

As with the experimental procedures described in the previous section, the technique of cross-spectrum correlation was applied to several simulated waveforms to test its viability. The cross-spectrum technique appeared to provide inferior handling of the variably delayed signals. Figure 7.21 gives the TDOA waveform of a simulated vehicle pass-by using a microphone separation of 1m. The false TDOA values indicate the techniques inferiority to the time domain cross-correlation technique seen in Figure 7.16.

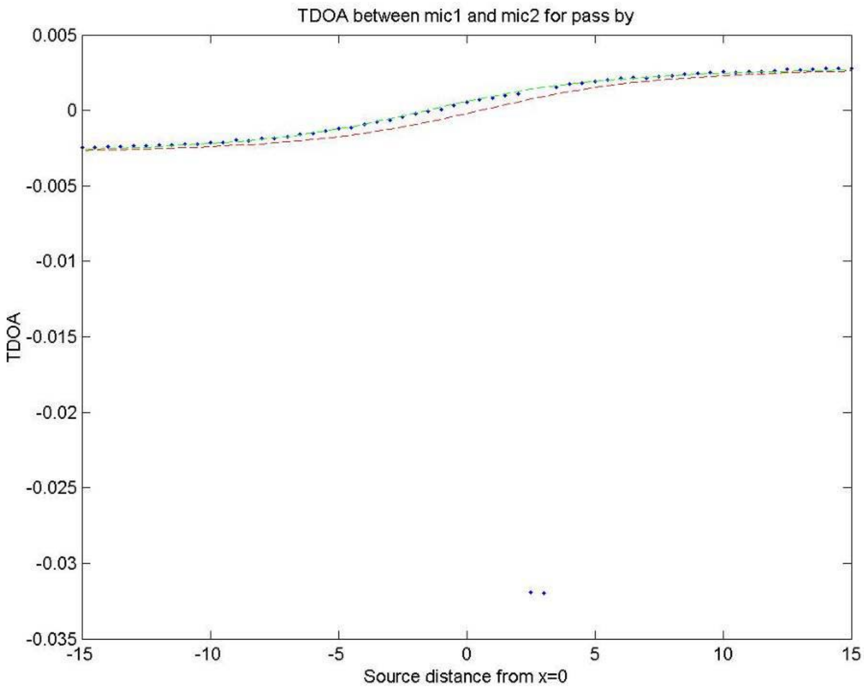


Figure 7.23: Cross-spectrum TDOA waveform of vehicle pass-by using an array aperture of 1m

Figure 7.24 shows the correlation waveform where the source was at a position of x=0 with respect to the first microphone in a 5m array. The peak magnitude definition is much less than that seen in Figure 7.10 for time domain cross-correlation. The lower definition will see the appearance of false peak values in the presence of interference much earlier than with the time domain approach. For these reasons and also the fact that the cross-spectrum technique cannot be simplified computationally by using restricted lag, the time domain option for TDOA estimation has been selected as the most robust and applicable. An extensive set of cross-spectrum experimental waveforms is provided in Appendix G.

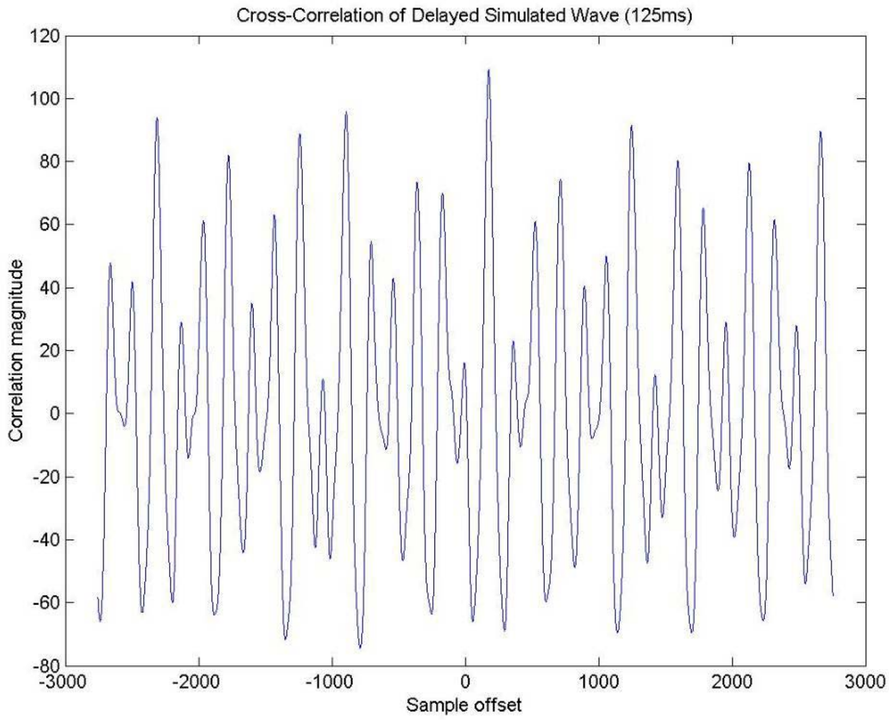


Figure 7.24: Cross-correlation derived from cross spectrum for window start position of x = 0

7.8 Performance in a multi source or reverberative environment

The introduction of additional noise sources on the roadway will cause further correlation peaks at different delays. The magnitude of these correlation peaks will depend upon the sound intensity of the source as well as its proximity to the array. The assumption that the loudest vehicle will be perceived dominantly by the system is challenged by the attenuation evident in section. False localization peaks may also occur where correlation values add. The local side peaks in Figure 7.10 for example may add to the central peak for another noise source thus giving a false time delay. In practice, additional correlation peaks may be reduced by ensuring the array is located in a non reverberative environment.

7.8.1 Dual Vehicle Simulations

A number of experiments were conducted where the simulated pass-bys of two vehicles were combined at each microphone. The dual vehicle configurations consisted of two vehicles travelling in adjacent lanes in the same direction. Vehicles travelling in different lanes travelling in opposite directions were also tested. The waveforms attained from the experiments are given in Appendix H. Figure 7.25 shows a series of TDOA waveforms for a dual pass-by where the sources travel in adjacent lanes.

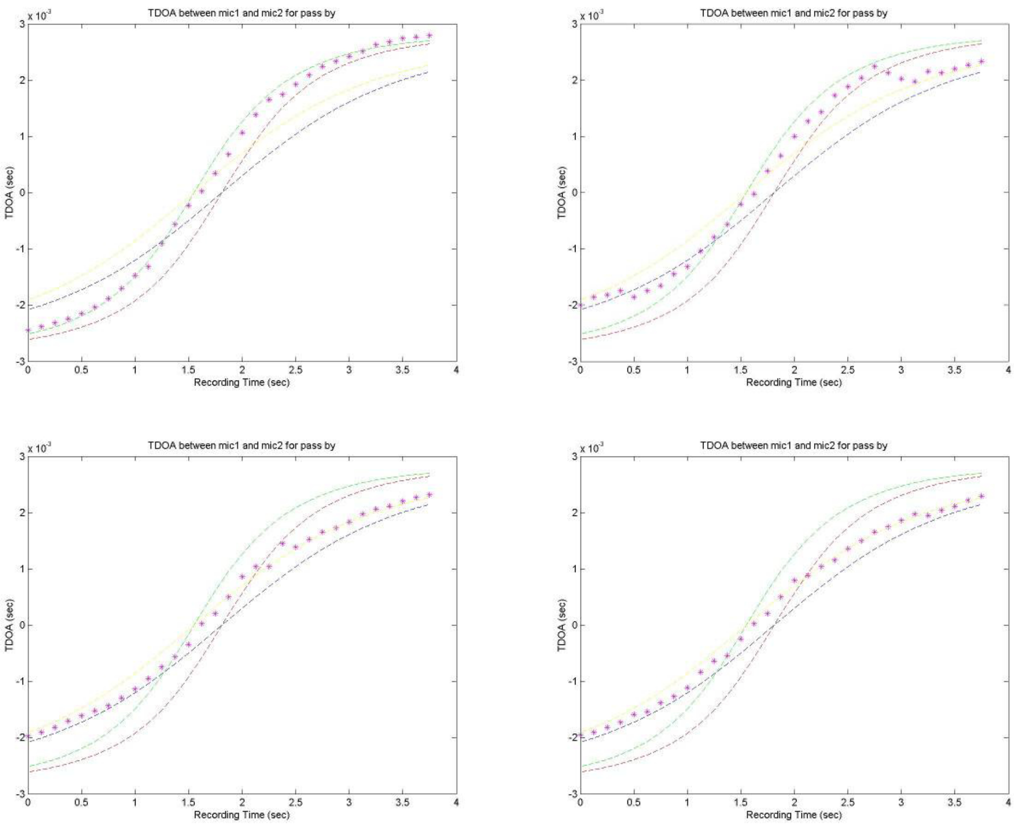


Figure 7.25: TDOA of dual vehicle environment, various signal magnitudes

The four waveforms represent the pass-by for different vehicle signal magnitudes. For the first plot (top left), the vehicles have equal magnitude and so the closest vehicle is detected by the system. As the far vehicle's signal magnitude is incremented to 2, 3 and 4 times the magnitude of the near vehicle, it is seen that the far vehicle path is detected more dominantly. These dual vehicle comparisons were conducted for several configurations and are illustrated in Appendix H. The approximated minimum amplitude differences between simulated near and far vehicles for path recognition are presented in table 7.1.

Table 7.1: Signal differences required for dual vehicle far lane recognition

Source1 Path	Source2 Path	Source1 offset	Source2 Offset	Required multiplier for far source recognition	Required signal difference
-20m to 20m	-20m to 20m	7.5m	11m	3	2dB
-20m to 20m	-20m to 20m	7.5m	14.5m	4	2.5dB
-20m to 20m	20m to -20m	7.5m	11m	4	2.5dB
-20m to 20m	20m to -20m	7.5m	14.5m	4	2.5dB
-20m to 20m	20m to -20m	7.5m	18m	7	5.0dB
-20m to 20m	-10m to 30m	7.5m	7.5m	4 for complete detection	2.5dB

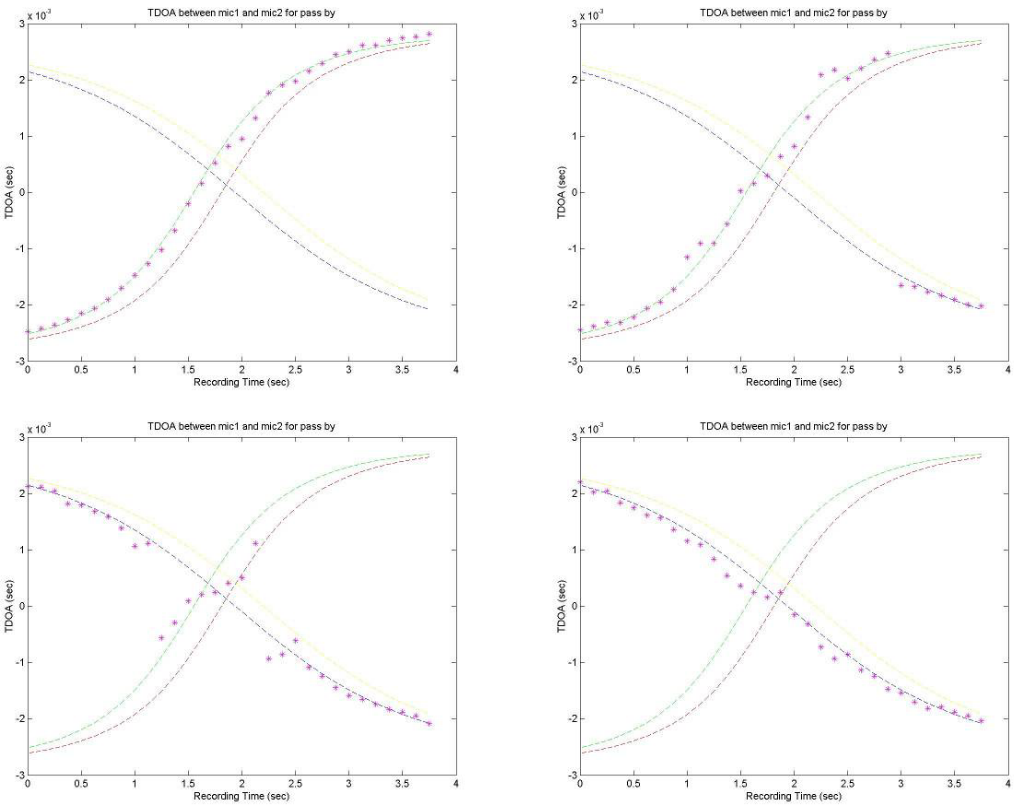


Figure 7.26: TDOA of dual vehicle environment (opposite travel directions), various signal magnitudes

The required signal difference levels in table 7.1 indicate that even a vehicle in the far lane of a four lane roadway will be detected above a vehicle in the near lane if it exceeds the near vehicle sound level by 5dB. Vehicles in the second and third lanes required levels 2.5dB grater than the near vehicle for detection. In practice, the system could be installed on the side of the roadway which experiences most noisy vehicle behaviour (i.e. the downward travelling lanes for targeting compression braking) to increase sensitivity for those lanes.

7.9 Field Recording Localization

A number of field recordings were obtained using a two microphone array. The recordings were taken 7.5m back from the four lane road using the measurement procedures described in section 6.9.

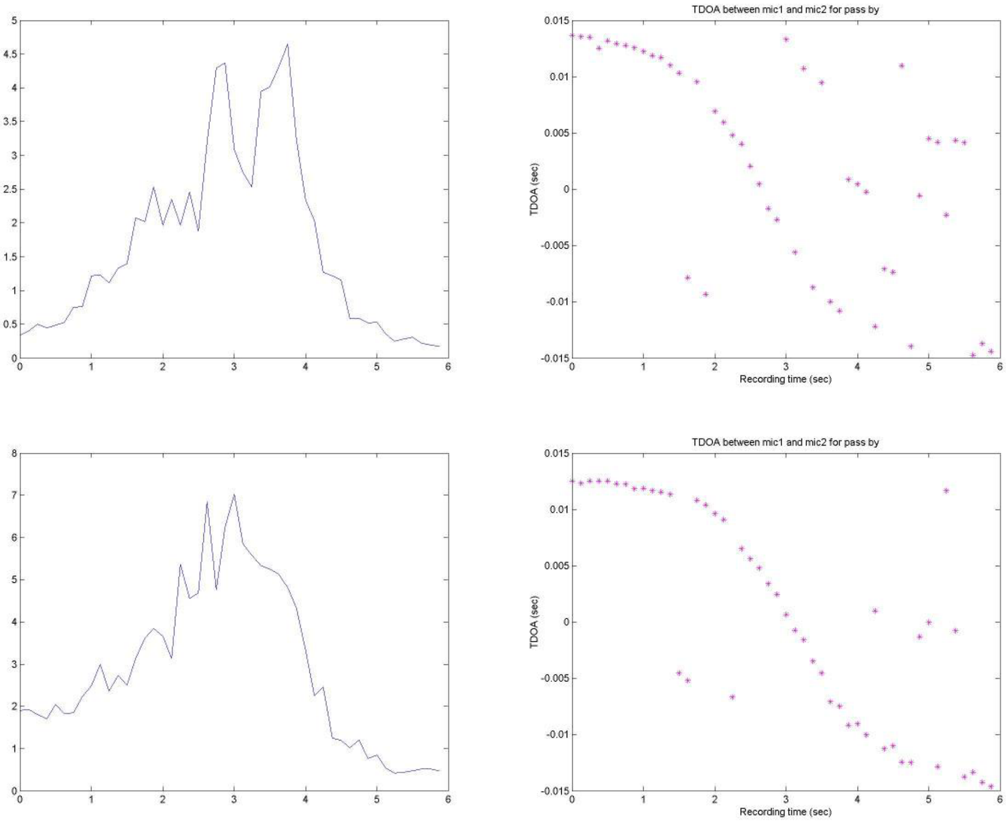


Figure 7.27: TDOA of estimation of field array signals using an aperture of 5m

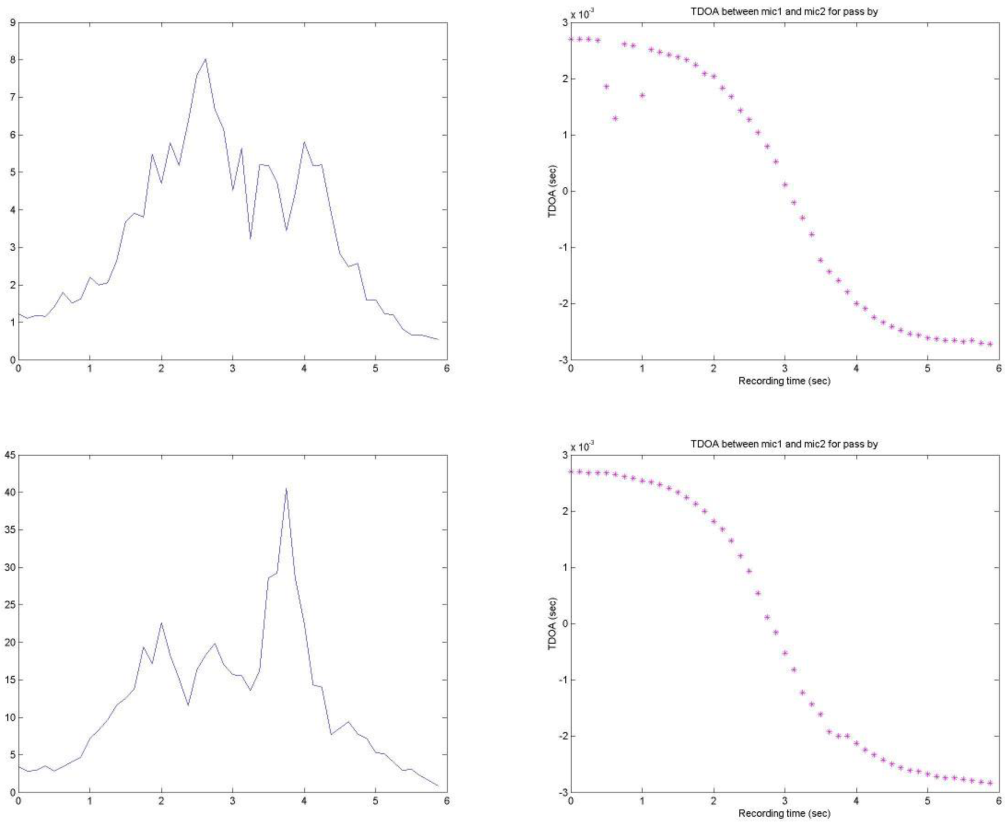


Figure 7.28: TDOA of estimation of field array signals using an aperture of 1m

Figure 7.27 above shows the results obtained from recordings taken using a 5m separation distance. The recordings in Figure 7.28 were taken using a 1m separation distance. The TDOA waveforms derived through the cross correlation of the field samples confirm the increase in accuracy for a smaller array aperture (the 1m waveforms). The functionality of the peak correlation magnitude as an indication of location accuracy is apparent in Figure 7.28. The first TDOA waveform exhibits some false values and is accompanied by a low scaled peak correlation plot. The lower TDOA waveform however does not exhibit false values and is accompanied by a large scale peak magnitude plot.

The research discussed in this chapter has confirmed the greater passive localization accuracy provided by small aperture array size. The data sets provided also suggest that the accuracy of a TDOA estimate may be confirmed by the magnitude of its respective cross-correlation waveform peak. The time domain cross-correlation technique has been selected over the cross-spectrum technique due to its greater robustness, peak definition and compatibility with restricted lag processing.

CHAPTER 8 - VEHICLE RECOGNITION RESULTS

8.1 Observations of the recordings

8.1.1 Proximity of test data

In choosing the sound frame for analysis, the layout of the microphone array was considered. For a three microphone array with a total length of 3m (suggested in section 7.3), the distance between microphones being 1m. Considering the range and capturing ability of a site camera, the field of detection has been decided as no larger than 30m from the middle microphone. On a roadway with a speed limit of 60km/h (assumed as the average travel speed), the time either side of the waveform peak for which data should be analysed is 1.8 seconds.

It was illustrated in Figure 4.6 that at a closer proximity to the microphones, the frequencies of the vehicle waveform will appear to shift more due to the change in relative velocity of the vehicle in the direction of the microphone. The test frames were all taken from the approach side of the vehicles waveform to provide some continuity with frequency shifting and the received sound. This also provides sound type continuity as the sound perceived is not only dependant on the relative velocity of the vehicle but also the position of the observation point due to different sounds produced by the vehicle (stepping away from the point source assumption).

From the analysis of the recorded pass-bys, it was noted that the vehicle tyre on road noise is only apparent when the vehicle is in close proximity to the microphone. This applies also for noises made by loose components mounted on the vehicle. This was noticed particularly for trucks, which would rattle and clang as the passed. However when beyond the immediate vicinity of the microphone, only the vehicle engine noise was clearly apparent. This is undoubtedly due to a faster decline in high frequency signal magnitude over distance. The appropriate use of filters or the analysis of select frequency ranges has helped to eliminate complications caused by such noise events.

8.1.2 Window size

The information that is held in a sound recording can give a large amount of detail to an observer of the type of vehicle passing, the size of vehicle, its mode of operation and even its speed. These conclusions may be made from the sound information using various lengths of recorded data. By observing the recording of a complete vehicle pass by for example, it is would be easier to hypothesize the size and speed of the passing vehicle than if the observer listened to a short burst of the recording. Feasible estimations concerning the mode of operation and the type of vehicle may still be made however with only relatively short lengths of sound information.

It has been noted through analysis of the recorded compression braking waveforms that short analysis periods (i.e. less than 1 second) are sufficient for the recognition of compression braking. The observation of longer signals can bring with it several additional interfering noise sources. These noise sources include rattling parts on the truck, pitch changes due to shifting gears, high pressure break blow offs and of course the presence of other passing vehicles (particularly ones passing in the opposite direction or closer to the microphone array). Taking short analysis windows of the acoustic signal of a passing vehicle will increase the chance of isolating the uncorrupted source signal.

In the following sound experiments, an analysis window of 20ms, 125ms and 1000ms was applied for feature extraction. In all cases, the window was aligned with the section of recorded waveform that best represented the vehicle operation and type. Where frame ratios are used, a long window of 1 second is used with the smaller windows being either 20ms or 125ms.

8.2 Zero-Crossing Rate

A Matlab program was developed to measure the zero crossing rates in the analysis windows of the test signals. The results are displayed in table 8.1 below.

Table 8.1: Zero Crossing Rate data for the test group

	ZCR		
	20ms	125ms	1000ms
Car1	1670.45	1425.15	1622.20
Car2	5609.21	5921.02	6186.46
Car3	1056.86	1269.81	1591.66
Car4	816.67	1117.67	1115.71
Car5	431.88	566.41	527.98
Car6	3803.3	3875.26	3657.50
Truck1	**1958.94**	**1985.77**	**2390.47**
Truck2	**2343.71**	**3107.77**	**4632.79**
Truck3	**5787.53**	**6407.28**	**5605.71**
Truck4	**7285.24**	**6663.68**	**6102.73**
Truck5	**5892.88**	**4297.58**	**4852.78**
Truck6	**2047.84**	**1938.81**	**2170.01**
Truck7	4910.70	4834.75	5134.95
Truck8	4719.89	3603.91	4177.27
Truck9	8000.86	9112.89	8494.37
Truck10	1952.59	1376.38	1719.7
Mot1	860.95	1646.06	1187.92

The above data set seems to indicate a generally lower ZCR for the car and truck signals which have predominantly low frequency content. As explained further on in this section, the ZCR value acts as an indicator of the ratio between the magnitudes of low and high frequency components in the signals. From this data set, a ZCR threshold of 1700 in the 125ms window could be used to eliminate several of the additional cars and trucks from identification whilst retaining the entire compression braking group.

8.3 Zero-Crossing Rate Ratio

Table 8.2: High and Low Zero Crossing Rate Ratios for the test group

	HZCRR		**LZCRR**	
Window size	20ms	125ms	20ms	125ms
Car1	0.12	0.13	0.18	0.13
Car2	0.00	0.00	0.00	0.00
Car3	0.10	0.00	0.18	0.00
Car4	0.08	0.00	0.14	0.00
Car5	0.10	0.00	0.12	0.00
Car6	0.00	0.00	0.06	0.00
Truck1	**0.06**	**0.00**	**0.08**	**0.00**
Truck2	**0.04**	**0.00**	**0.20**	**0.13**
Truck3	**0.06**	**0.00**	**0.16**	**0.13**
Truck4	**0.00**	**0.00**	**0.06**	**0.00**
Truck5	**0.02**	**0.00**	**0.02**	**0.00**
Truck6	**0.02**	**0.00**	**0.04**	**0.00**
Truck7	0.06	0.00	0.08	0.00
Truck8	0.00	0.00	0.02	0.00
Truck9	0.02	0.00	0.08	0.13
Truck10	0.10	0.00	0.14	0.00
Mot1	0.12	0.13	0.20	0.13

The ZCRR appears to provide little class separability for this test group. The low values for the HZCRR and LZCRR using a 125ms window indicate that the waveforms have a fairly controlled ZCR over time. The higher values given using a shorter comparison window (20ms) then indicate that variations in the ZCR do occur in short bursts but seem to average out to zero in longer samples.

The sound contribution from tyre on road noise can often dominate the recorded signal. This noise type appears to contribute mainly high frequency components which will have an affect on the zero crossing rate. A non characteristic ZCR will be perceived when the magnitude of the lower frequencies are at a magnitude that the waveform oscillates about a magnitude of zero but the magnitude of high frequency noise is sufficient to cause a high rate of zero crossings.

8.4 ZCR high frequency control

8.4.1 High frequency contributions

The relation of the magnitude of the vehicle engine acoustic emissions to the magnitude of high frequency emissions caused by tyre on road noise, high pitch braking and environmental noise can disrupt the expected zero crossing rate. If the magnitude of the low frequency components are not great enough to offset the high frequency components from the zero line, an uncharacteristically high zero crossing rate will be experienced. This will resemble more the frequency and noise produced by factors other than the vehicle engine itself. In order to be able to recognize a vehicle resonance accurately, the magnitude of vehicle noise must be at least greater than two times the magnitude of high frequency additional noise. Figures 8.1 and 8.2 below show how the relation of the low frequency component magnitude and high frequency component magnitude can affect the zero crossing rate for a dual frequency waveform. The frequencies used in the simulation were 400Hz and 1800Hz.

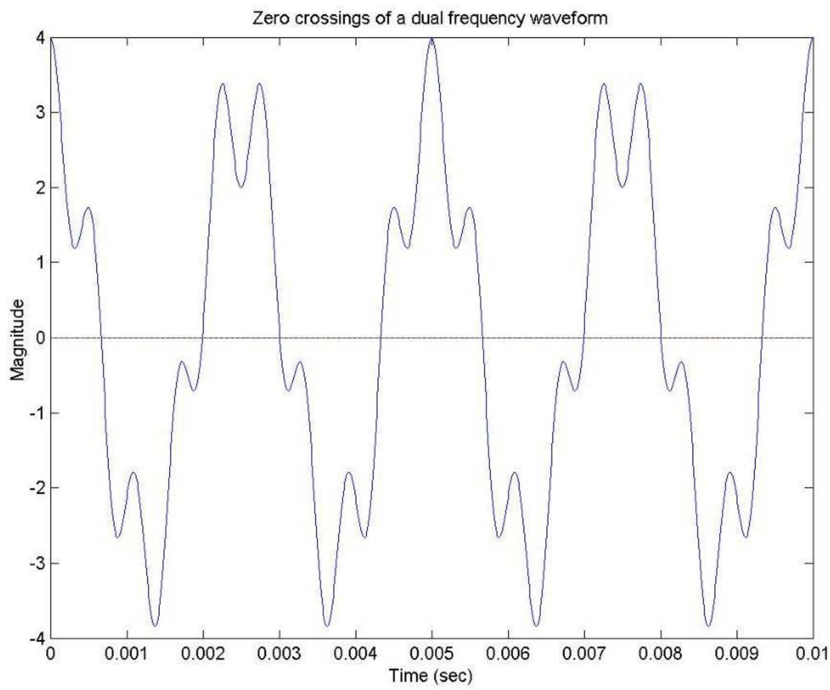


Figure 8.1: High ratio of low frequency magnitude to high frequency magnitude.

The figure above exhibits a large low frequency component where the resonant ZCR is unaffected by the high frequency component. The number of zero crossing in the 10ms period is eight. The ratio of low frequency wave magnitude to the high frequency wave magnitude is 3.

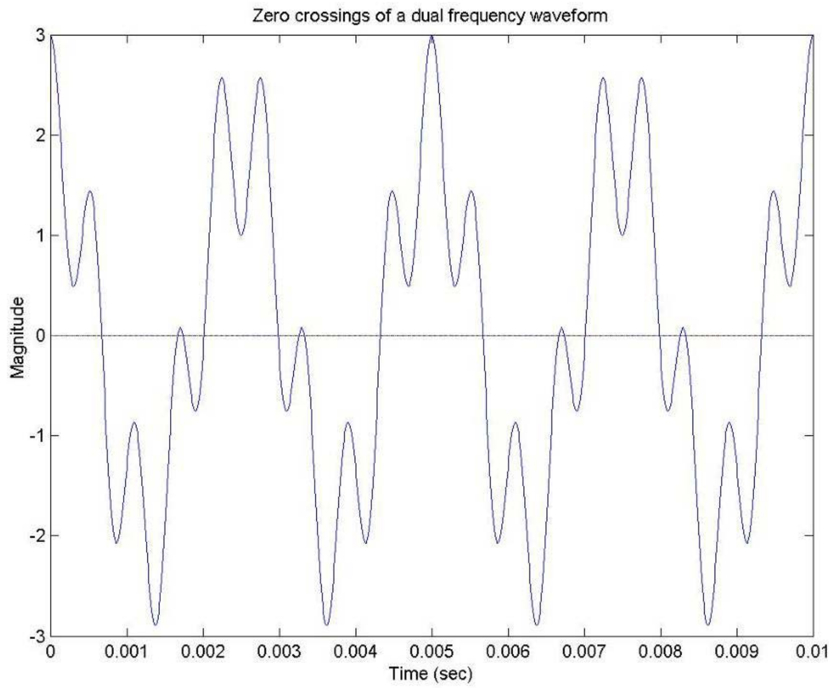


Figure 8.2: Higher high frequency magnitudes causing higher ZCR

In figure8.2 above, the larger high frequency magnitude relative to that of the low frequency component causes additional zero crossing of the signal. The number of zero crossings in the 10ms window is 16, twice that of the waveform in Figure 8.1. The ratio of low frequency wave magnitude to the high frequency wave magnitude is 2.

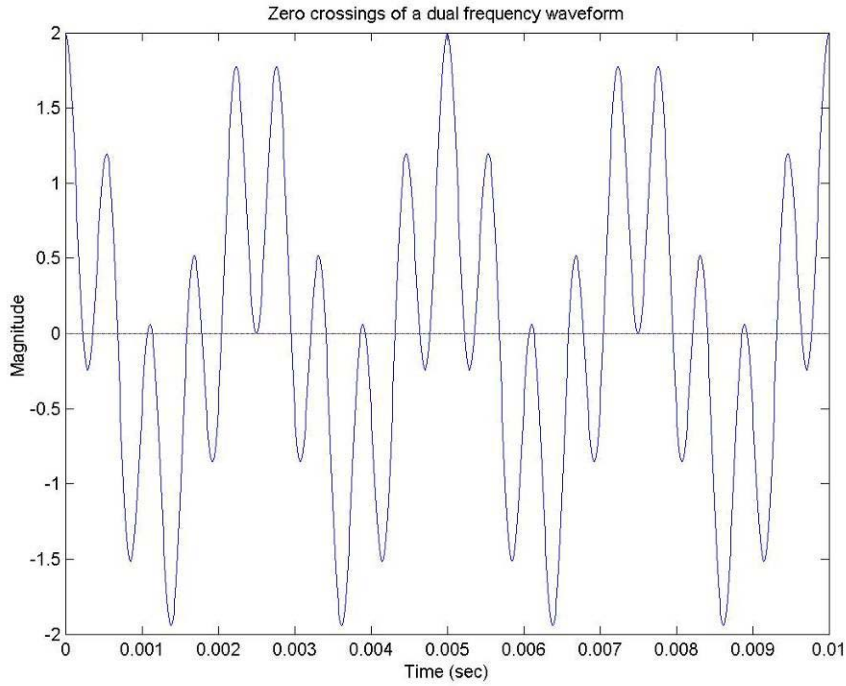


Figure 8.3: Low freq to high freq magnitude ratio of 1

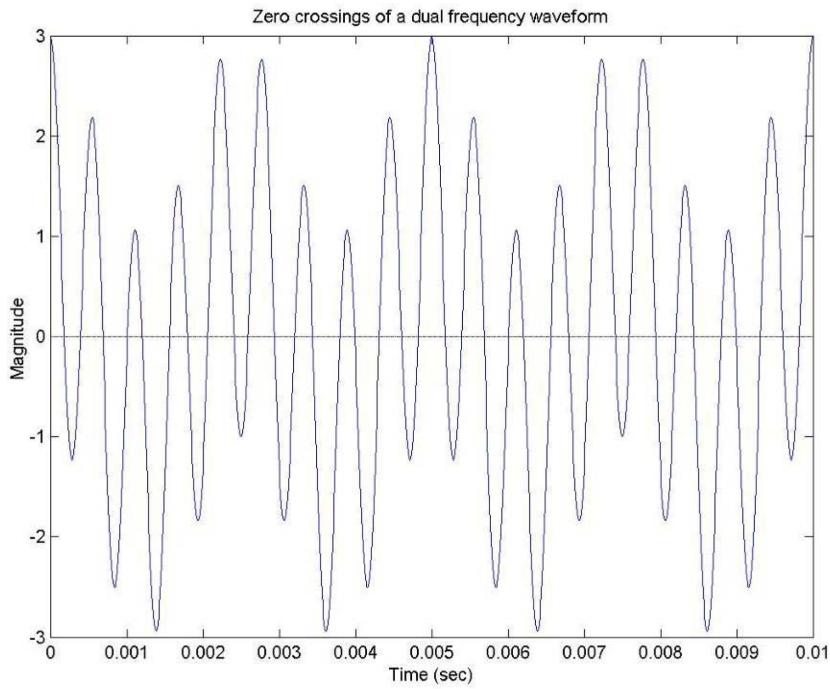


Figure 8.4: Low freq to high freq magnitude ratio of 0.5

Figures 8.3 and 8.4 above demonstrate even higher zero-crossing rates where the magnitude of the high frequency component is large in relation to the magnitude of the low frequency component. The waveform in Figure 8.3 has a low frequency to high frequency magnitude ratio of 1 resulting in 31 zero crossings in 10ms. Figure 8.4 has a magnitude ratio of 0.5 resulting in 36 zero crossings in the 10ms (equivalent to the un-modulated zero crossing rate of the high frequency component).

Examination of the data in table 8.1 showed that the trucks having the higher ZCR values had perceivably louder high frequency components. These were in the form of tyre on road noise from either the truck itself or other passing vehicles. A high pitch squeal is also produced by many trucks as they brake down a hill. The waveforms of some of the field recordings are shown below. The 20ms waveform for Truck6 having the lowest ZCR is shown below.

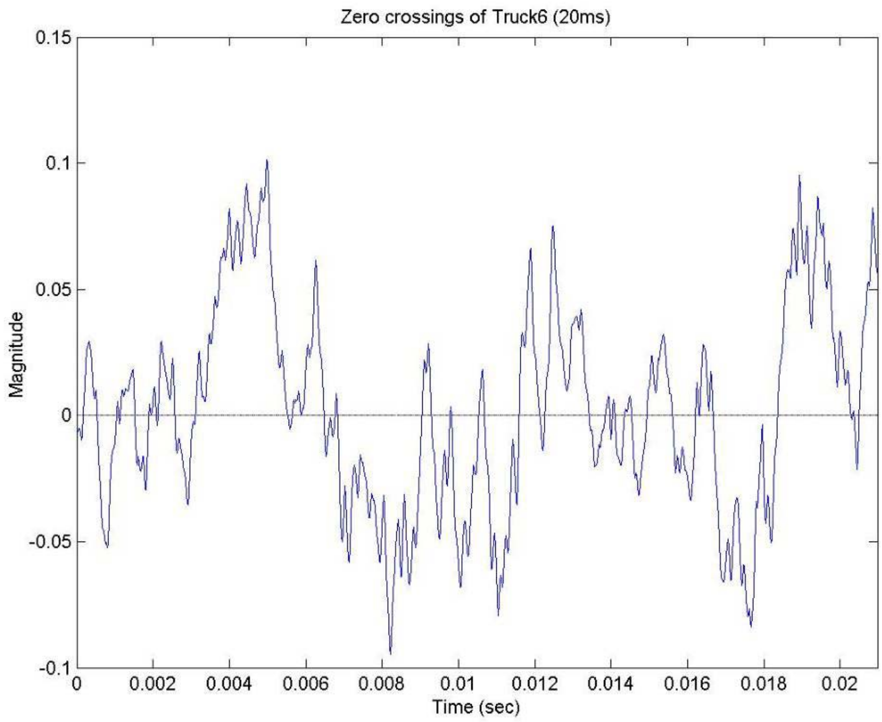


Figure 8.5: Truck6 20ms waveform

The fundamental frequency can be seen to have a period of around 7ms (143Hz). Additional low frequency oscillations of around 1000Hz (period of 1ms) can be seen in the waveform. High frequency components are also visible but they do not contribute to the zero crossing rate. This is not the case for the waveform for Truck4 which recorded the highest ZCR in the sample set. The waveform in Figure 8.5 exhibits an underlying low frequency component (perceivably around 200Hz).

However the magnitude of the high frequency additions is such that the ZCR is far greater than that for the waveform in Figure 8.6.

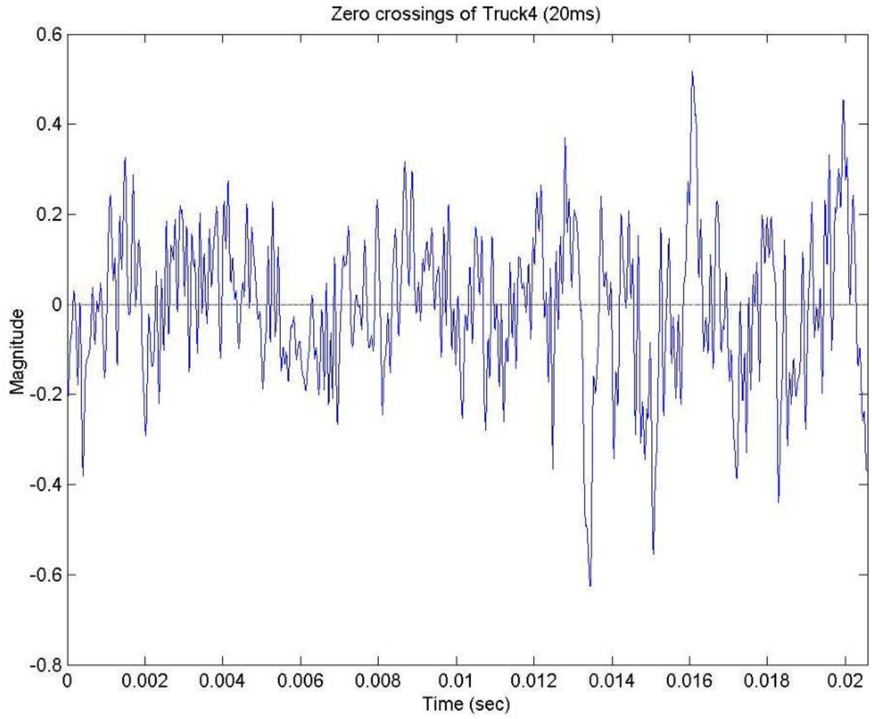


Figure 8.7: Truck4 20ms waveform

8.4.2 Signal Filtering

It is clear from the observations stated above that the information gathered from the ZCR of passing trucks will not be uniquely descriptive of the vehicle's engine where the sound signal consists of high magnitude high frequency components. The obvious way for removing these unwanted components (providing little information concerning the operation of the vehicle engine or exhaust) would be to filter them from the signal using a low pass filter.

The following set of figures illustrates the application of two types of lowpass filters, the Butterworth and the Chebyshev to a simulated dual frequency waveform. The simulations were conducted using Matlab functions and a low pass frequency of 441Hz.

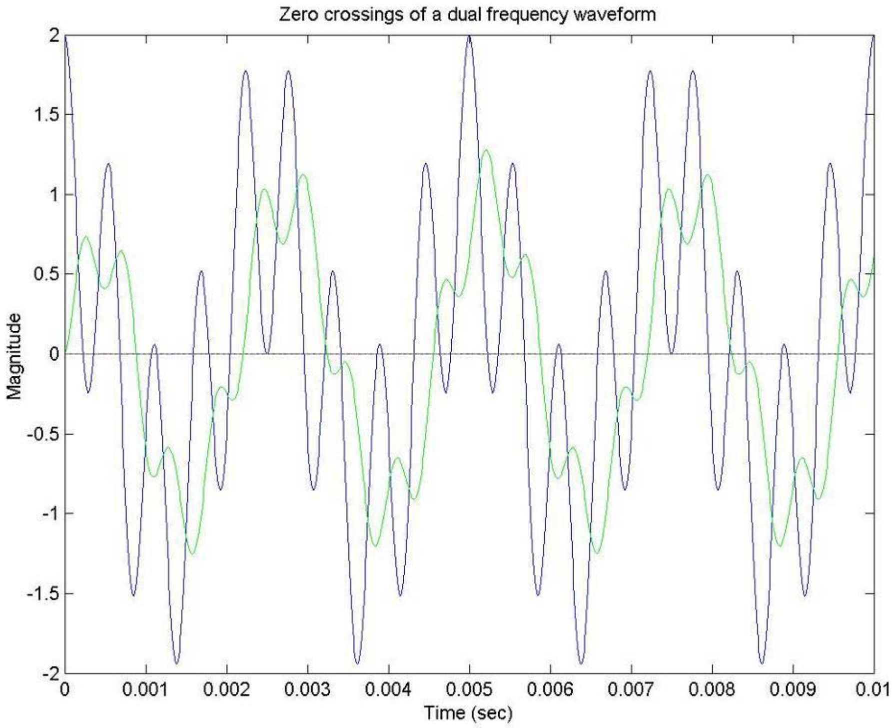


Figure 8.8: Signal filtered using 2nd order Butterworth filter with cut-off freq of 441Hz

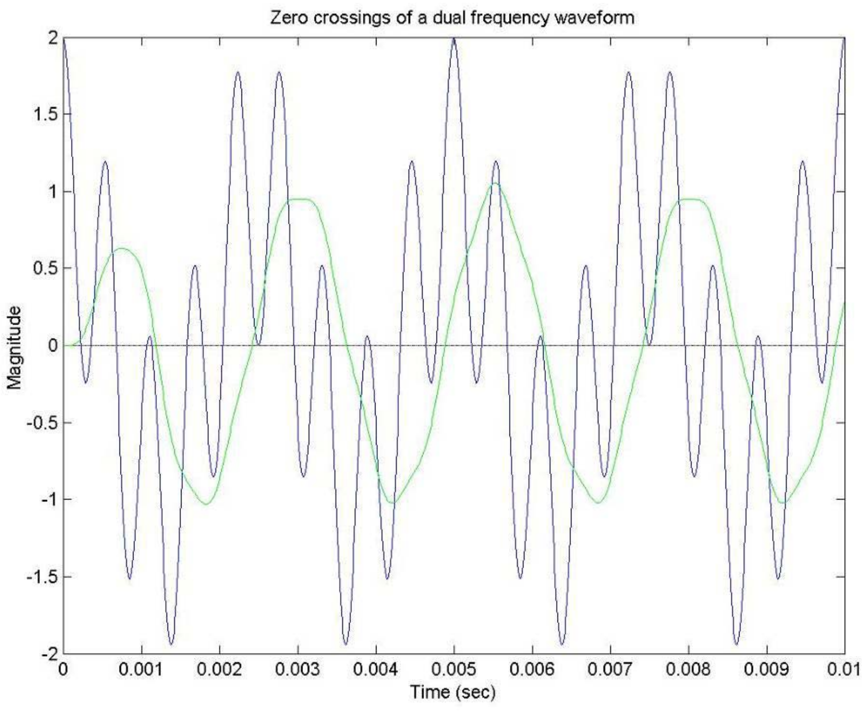


Figure 8.9: Signal filtered using 5th order Butterworth filter with cut-off freq of 441Hz

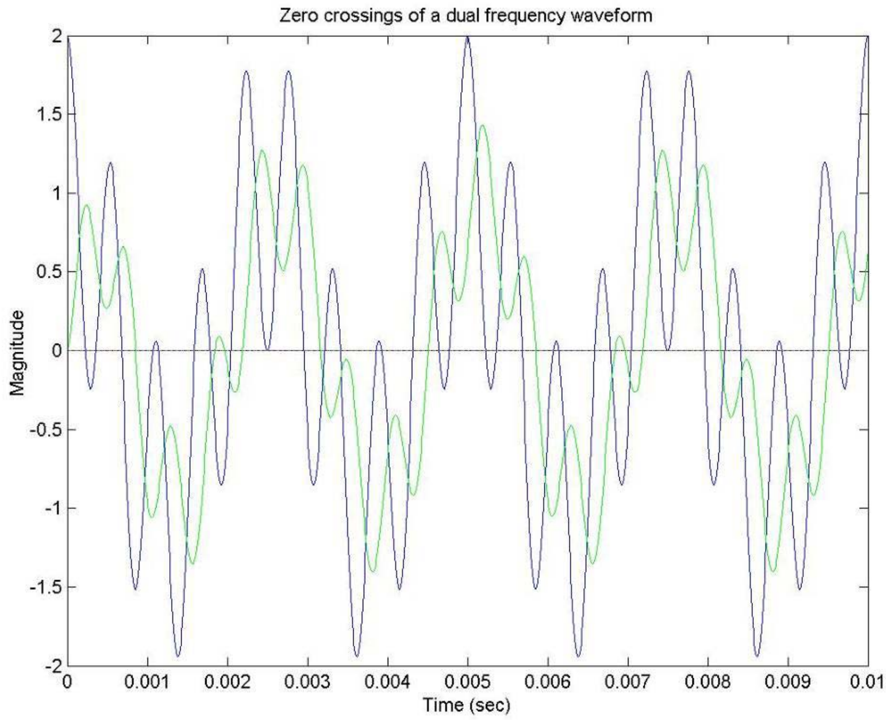


Figure 8.10: Signal filtered using 2nd order Chebyshev filter with cut-off freq of 441Hz and pass band ripple of 0.5dB

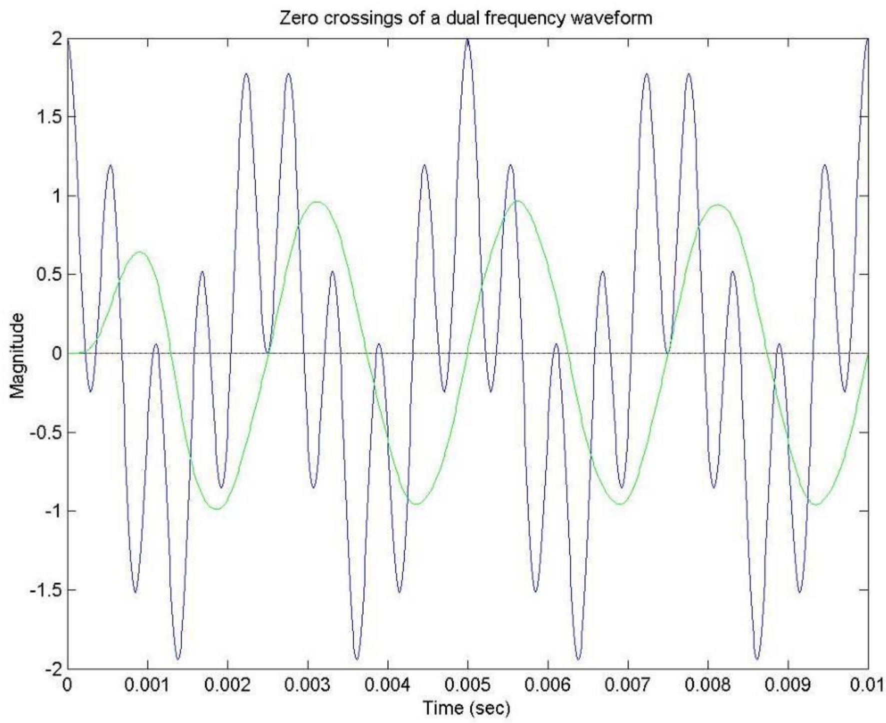


Figure 8.11: Signal filtered using 5th order Chebyshev filter with cut-off freq of 441Hz and pass band ripple of 0.5dB

It is apparent that by the application of a digital low pass filter to the recorded signal, the un-characteristically high ZCR may be removed along with unwanted high frequency information.

8.4.3 Analysis of Field Data

A 4th order Chebyshev filter with pass band ripple of 0.5 was selected for application to the field recordings. This was decided due to the adequately smooth output waveform and sharper frequency response to the Butterworth filter.

Using the same field waveforms analysed above, the Chebyshev filter was applied using Matlab and the zero-crossing rates were re-examined.

Figure 8.12 below shows the waveform for Truck6 (20ms) after the application of the 4th order Chebyshev filter with a cut-off frequency of 1200Hz. Note that the cut-off frequency is that which is equal to the negative bandpass ripple after filtering.

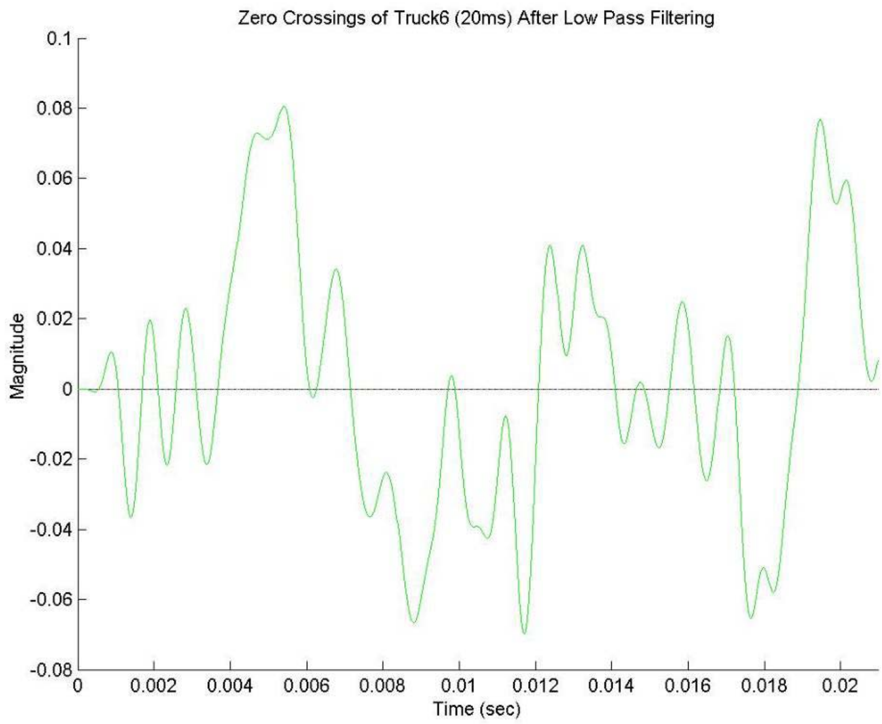


Figure 8.12: Truck6 (20ms) waveform after lowpass filter application with cut-off frequency of 1200Hz

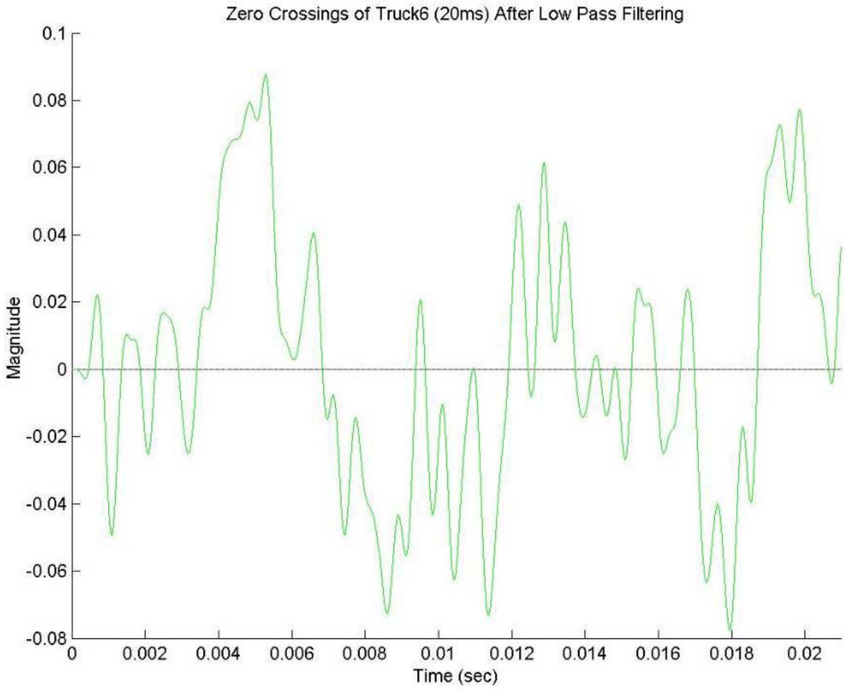


Figure 8.13: Truck6 (20ms) waveform after lowpass filter application with cut-off frequency of 2000Hz

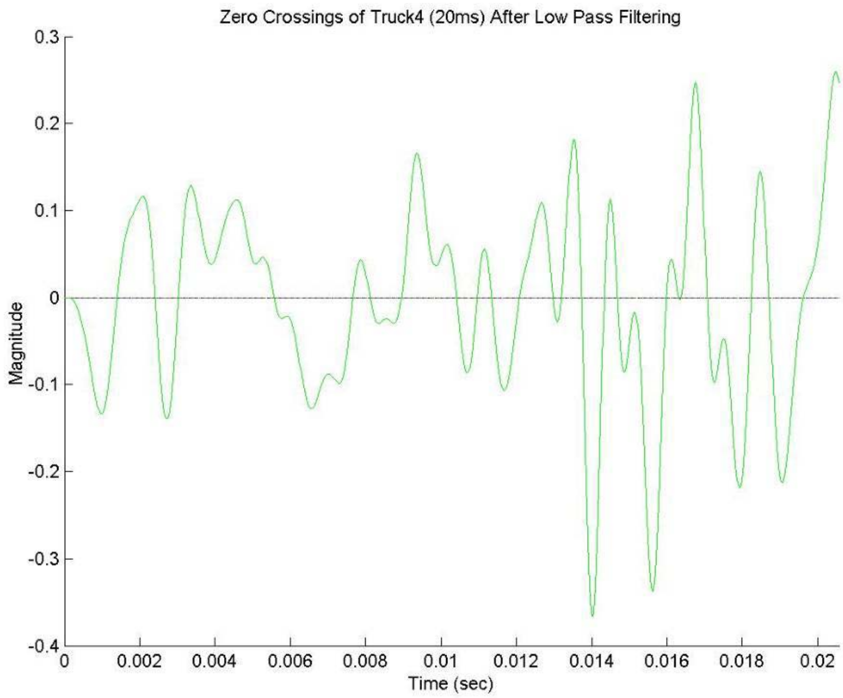


Figure 8.14: Truck4 (20ms) waveform after lowpass filter application with cut-off frequency of 1200Hz

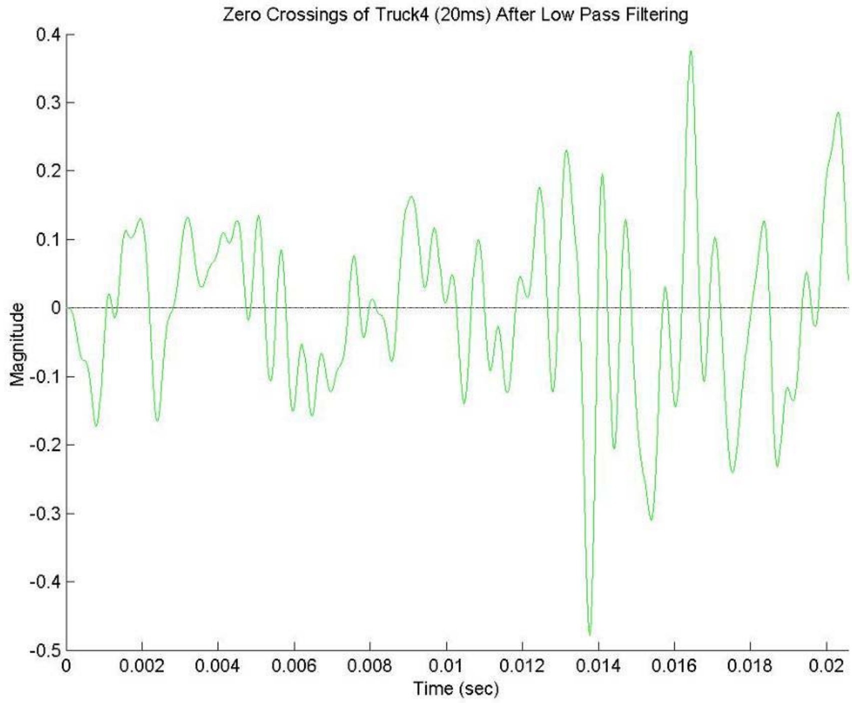


Figure 8.15: Truck4 (20ms) waveform after lowpass filter application with cut-off frequency of 2000Hz

It is seen in the figures above that the application of the low pass filter dramatically reduces the influence of high frequency noise on the zero-crossing rate of the signals. A cut-off frequency of 1200Hz and 2000Hz were applied in order to examine the most suitable approach for ZCR frequency limiting. It was decided that the 2000Hz cut-off frequency best selected the low frequency components of the signal for analysis while adequately attenuating any disruptive high frequencies.

After applying a lowpass filter with a cut-off frequency of 2000Hz, the zero-crossing rate was calculated for the test waveforms. The results are given in Appendix E. The ZCR data set after lowpass filtering appears to provide less class separability than the original set. It may be conceded that the low frequency engine noise attributes may be determined using alternate techniques that require less processing than the applied filters (such as 1/3 octave analysis).

It is seen from the analysis of the waveforms that the lowest dominant frequency of the vehicles (usually in the range of 100-200Hz) is not particularly helpful in providing information about the vehicle type. This frequency will be easily determined using the autocorrelation or 1/3 octave filtering procedure discussed later in this chapter. It appears that the dominant frequencies from the range of 400 – 3500 Hz provide more information concerning the type and operation of the vehicle. This is seen in the comparison of the sound files Car1 and Car2. Car1 is a sound recording of a reasonably quiet sedan accelerating up a hill. Car2 is a recording of a passing diesel 4WD wagon. Although these vehicles have quite a similar dominant low frequency of around 125Hz, the diesel vehicle is defined by quite a large component at around 2800Hz where the car exhibits a small frequency of around 3300Hz.

Most vehicles use similar revs in their operations and therefore the frequencies produced will be similar. It could be generalized that a car moving with 4000rpm firing a 6 cylinder engine in three stages will create a fundamental sound pressure wave at approximately 4000/60 x 3 = 200Hz. The equivalent frequency for a 4 piston engine would be
4000/60 x 2 = 133Hz.

The ZCRR values obtained after filtering the signals still provide little class separability for the test group. This was apparent for ratio thresholds of both 1.5 and 1.2. The data obtained for these test sets can be found in Appendix E.

Table 8.2a below shows the ZCR values obtained from the waveforms using a range of cut-off frequencies. Once again, the class separability appears to degrade and the technique is considered too computationally expensive to justify further application.

Table 8.2a: ZCR's of the 1 second waveforms using a lowpass filter

	ZCR				
Cut-off Freq	400Hz	800Hz	1200Hz	2000Hz	No Filter
Car1	318.45	334.42	401.31	538.07	1622.20
Car2	344.91	727.82	991.75	1994.50	6186.46
Car3	327.73	400.66	551.54	921.23	1591.66
Car4	352.59	388.55	526.39	712.18	1115.71
Car5	317.39	311.40	347.33	418.19	527.98

Car6	255.85	770.52	1804.86	2425.06	3657.50
Truck1	**429.39**	**625.20**	**889.59**	**1199.71**	**2390.47**
Truck2	**290.86**	**408.81**	**564.73**	**1001.52**	**4632.79**
Truck3	**159.93**	**369.85**	**688.72**	**1376.44**	**5605.71**
Truck4	**426.35**	**857.70**	**1160.24**	**1792.28**	**6102.73**
Truck5	**352.19**	**594.62**	**946.81**	**1510.51**	**4852.78**
Truck6	**377.26**	**608.20**	**864.02**	**1128.80**	**2170.01**
Truck7	276.67	381.55	605.29	1092.71	5134.95
Truck8	281.81	483.32	916.26	1395.08	4177.27
Truck9	163.31	358.50	770.77	1871.15	8494.37
Truck10	153.35	187.21	348.53	539.72	1719.7
Mot1	323.98	353.98	415.97	530.96	1187.92

8.5 Noise Level Modulation Rate

In order to analyse the noise level modulation rate, the A weighted equivalent continuous sound level (Leq(A)) was processed for the field data using Matlab. The equivalent level was produced using integration times of 5ms, 20ms and 125ms as illustrated in Figure 8.16 below. Figure 8.17 shows the Leq levels over a 125ms period of the Truck1 waveform. The 5ms Leq values occur every 5ms and the 20ms recorded valued occur every 20ms and so on.

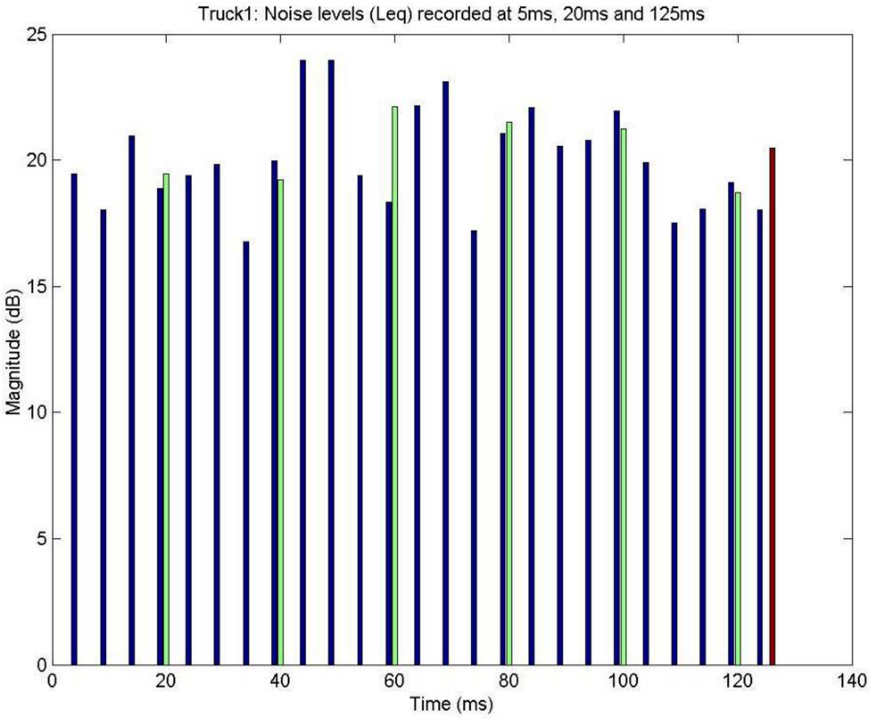


Figure 8.16: Leq levels for 'Truck1' recorded at intervals of 5ms, 20ms and 125ms

8.5.1 The magnitude level

It is important to note that in analysing the modulation properties of the recorded Leq levels, the magnitude of the waveform plays little importance. That is to say, the level offset from zero is not important; it is the magnitude between Leq levels that provides modulation data.

In the field this is easily visualized, where an acoustic signal is attenuated over distance, the noise level is lost but the nature of the sound is still discernable. This is because an attenuation of say 10dB will be applied to the high and low Leq levels alike, hence the modulation levels (the difference between the high and low noise levels) will remain the same. Therefore in this particular form of feature extraction, the level differences have been considered and not the level magnitudes themselves for any characteristic information. Taking this further, the Leq waveform in Figure 8.17 may be simplified by shifting the zero reference level to the minimum value. This is applied in Figure 8.18.

As explained in section 6.9.4, the actual Leq levels are not accurate (having not been recorded with a precision SLM) but the levels differences are adequately described.

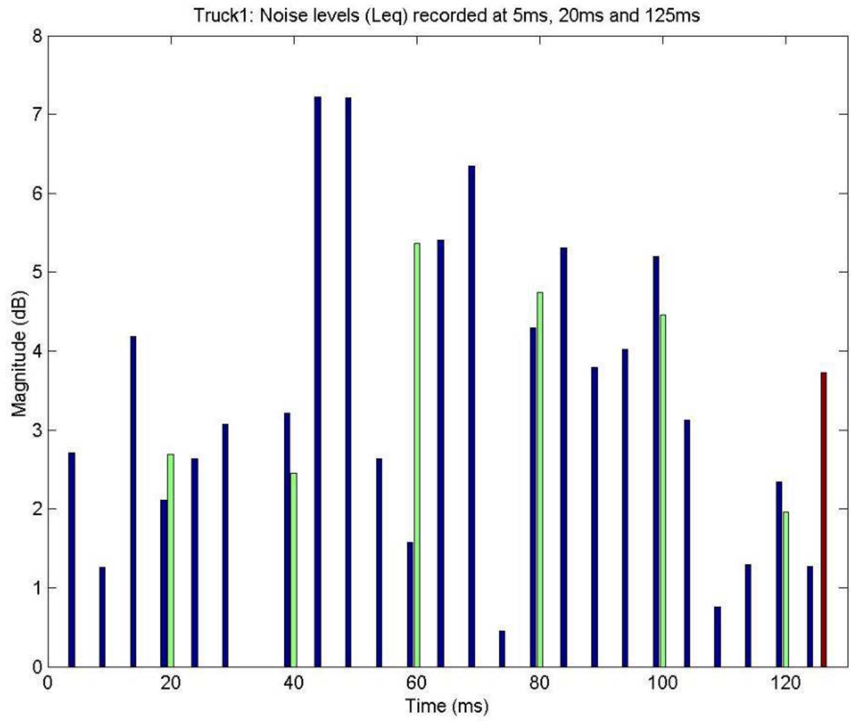


Figure 8.17: Reference level shifted Leq waveform

8.5.2 Results

Modulation number is defined as the number of times the Leq level series goes from a peak to a trough to a peak with a magnitude no less than that of a pre-set value. In the experiments displayed, values of 3dB, 5dB and 7dB were used. The value of a 3dB modulation was included to comply with the proposed selection criteria put forward by the NRTC report discussed in section 3.2.2. However as it was also stated, this criteria aims only to gauge the severity of known compression braking recordings, not to identify or recognize the sound type.

Figure 8.17 shows the modulations occurring in the 125ms waveform for Truck1 where the sound pressure level has been recorded at 5ms, 20ms and 125ms intervals. The 5ms levels have been applied to prevent the effect of aliasing that can be clearly seen in the figure where longer averaging times are used.

The results for the 5ms modulation rates for the 125ms and 1000ms waveforms are presented in Appendix E. Modulation magnitudes of 3, 5 and 7dB were used in the analysis. There appears to be little class separability available from the feature where the 125ms waveforms are assessed. In the 1000ms recordings, many of the non-compression braking vehicles could be ruled out by their high modulation rates above 3dB or 5dB. Six out of the eleven non-compression braking vehicles were well outside the range experienced by the compression braking group.

Figure 8.18 below shows the 5ms Leq levels for the 1000ms recording of Truck1. An analysis of the ratio of the average Leq value to the peak value (the average normalized value) was conducted. The results are given in tables 8.3 and 8.4 below.

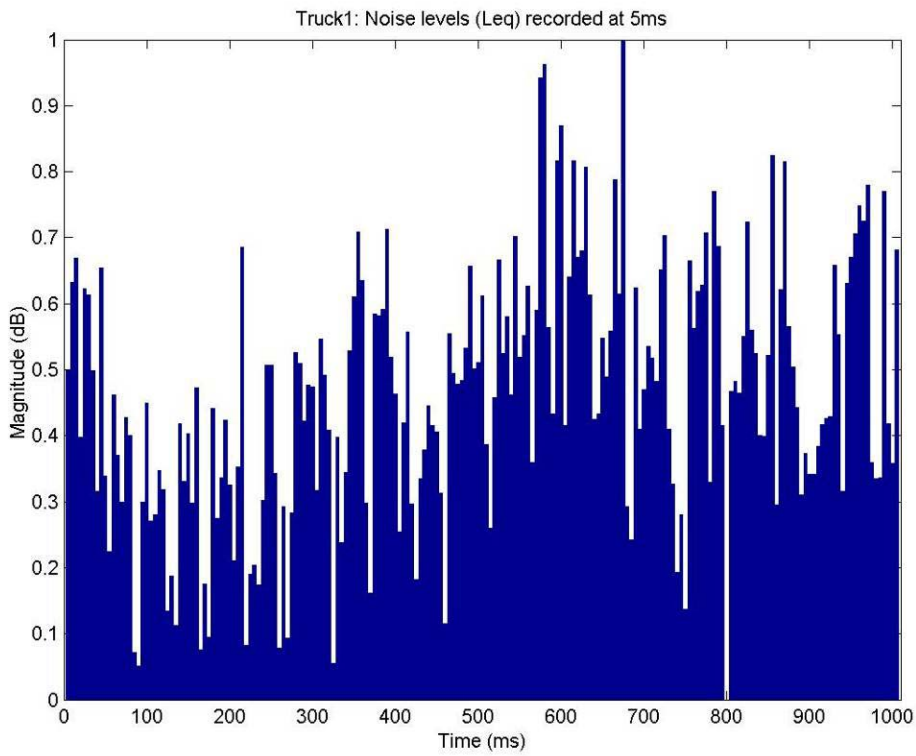


Figure 8.18: Normalized and reference shifted 5ms Leq waveform of Truck1

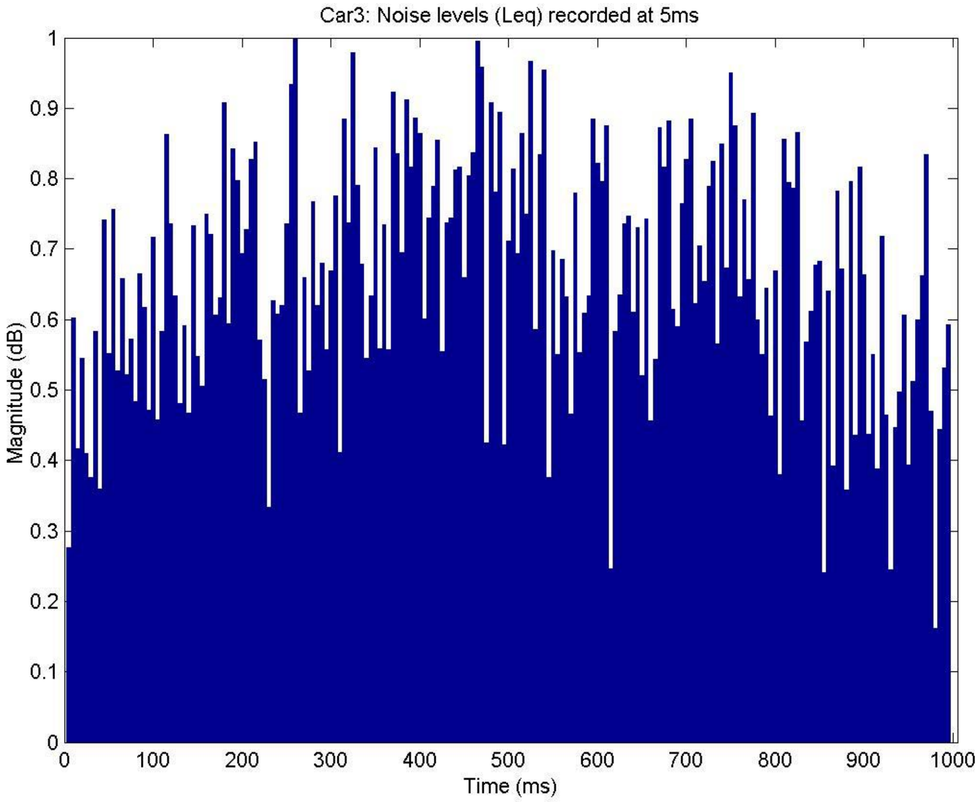


Figure 8.19: Normalized and reference shifted 5ms Leq waveform of Car3

Table 8.3: Averaged level of normalized Leq(A) waveform over 125ms

Wave File	Average Leq
Car1	0.55
Car2	0.44
Car3	0.47
Car4	0.56
Car5	0.47
Car6	0.53
Truck1	**0.45**
Truck2	**0.54**
Truck3	**0.62**
Truck4	**0.49**
Truck5	**0.37**
Truck6	**0.57**

Truck7	0.47
Truck8	0.63
Truck9	0.52
Truck10	0.66
Mot1	0.53

Table 8.4: Averaged level of normalized Leq(A) waveform over 1000ms

Wave File	Average Leq
Car1	0.58
Car2	0.58
Car3	0.66
Car4	0.58
Car5	0.59
Car6	0.54
Truck1	**0.46**
Truck2	**0.52**
Truck3	**0.45**
Truck4	**0.53**
Truck5	**0.45**
Truck6	**0.53**
Truck7	0.50
Truck8	0.56
Truck9	0.33
Truck10	0.60
Mot1	0.56

A low value for LeqAv indicates that there is a greater separation between the peak values of the Leq waveform and the average Leq level. Car4 has a higher average than Truck1. Is seen in waveforms 8.18 and 8.19 above where the peak levels for Car3 are much closer to the average level for the entire waveform. The group of cars seem to exhibit a generally higher average where the trucks using compression braking appear to lie in the fairly limited range of 0.45 – 0.53. Some of the other trucks also lie in this region but the truck accelerating uphill (Truck 10) is well clear of the range.

8.6 Low and High Short Time Energy Ratio (LSTER & HSTER)

Using short time windows of 5ms and 20ms and a long time window of 1000ms, the following results for low and high short time energy ratios were obtained. A tolerance factor of 1.5 was used for both the LSTER and HSTER selection.

The 20ms analysis provided little information concerning vehicle type. There appears to be some amount of grouping with the HSTER values in 5ms test, however there is still a great deal of multi-class inclusion. These data sets are given in Appendix E.

It was found that a short window of 5ms and a long window of 125ms yielded an adequate class separation (table 8.5 below). The use of the 125ms long window is also more desirable in capturing characteristic periods of engine operation. Using a low pass cut-off of 0.24 for the LSTER and of 0.16 for the HSTER will identify the compression braking group while excluding five of the additional vehicles (Car4, Trucks7, 8, 10 and Mot1).

Table 8.5: STER’s using small window of 5ms and long window of 125ms

Wave File	LSTER	HSTER
Car1	0.16	0.16
Car2	0.12	0.12
Car3	0.200	0.08
Car4	0.32	0.24
Car5	0.04	0.08
Car6	0.08	0.04
Truck1	**0.200**	**0.12**
Truck2	**0.240**	**0.120**
Truck3	**0.08**	**0.00**
Truck4	**0.200**	**0.16**
Truck5	**0.200**	**0.08**

Truck6	**0.12**	**0.04**
Truck7	0.1667	0.1667
Truck8	0.3600	0.200
Truck9	0.240	0.160
Truck10	0.320	0.200
Mot1	0.440	0.280

8.7 Spectrum Flux

The Spectrum Flux of the 1000ms and 125ms waveforms were extracted and are displayed in Appendix E. The information provided by the spectrum flux does seem to give some class separability information, particularly from the group of cars. The value given by the extraction process is an indication of the variance of the frequency signal over the entire sound recording. Being an average, the value may be offset easily by spurious events in the recording. It is typical that all spectrum waveforms should change over time and that the amount of change will be affected by additional noise and random events. It is seen by the comparison of the SF values for Truck5 in the 1000ms and 125ms analysis that the feature seems erratic in consistency. Figure 8.20 shows how the data sets from the 1000ms and 125ms analysis vary. 125ms (red dashed), 1000ms (blue solid)

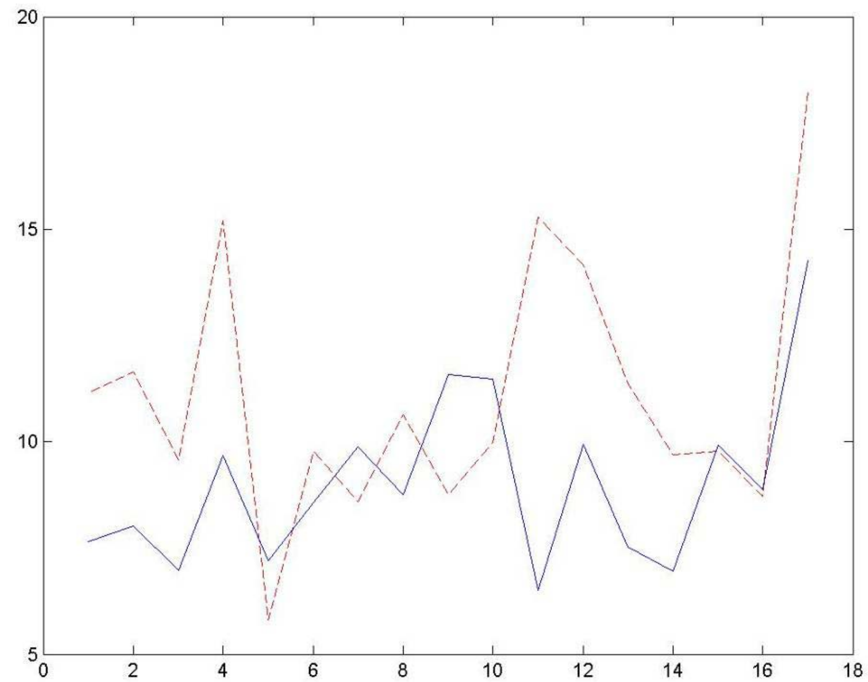


Figure 8.20: Comparison between 1000ms and 125ms SF analysis

Due to the influence from external and random noise events on the spectrum flux over time, the 125ms extraction technique was expected to be the most reliable for any type of characterization. The data presented above shows however that the trucks using compression braking are best classified using a range of SF values when a processing window of 1000ms was used. The suspicion arises as to whether the spectrum flux is being affected or biased due to high frequency sound components. The following section looks at using a smaller frequency range for spectrum flux extraction.

8.8 Narrowband Spectrum Flux

The information provided by the spectrum flux extraction procedure discussed above and in section 8.7 considers the entire frequency range of the sound signal. By reducing the number of frequency bands analysed in the 1/3 octave spectrum and hence discarding all un-important frequency bands, the results obtained will likely reflect more closely the operation of the investigated sound source. In this case the investigated sound source is the vehicle engine and exhaust system which have been shown in section 8.4.3 to be described by lower frequency bands in the 1/3 octave spectrum. The spectrum was therefore reduced to contain seven bands, ranging from band centre frequencies of 100Hz to 400Hz. All higher frequency information is considered un-important in this case for the recognition of compression braking.

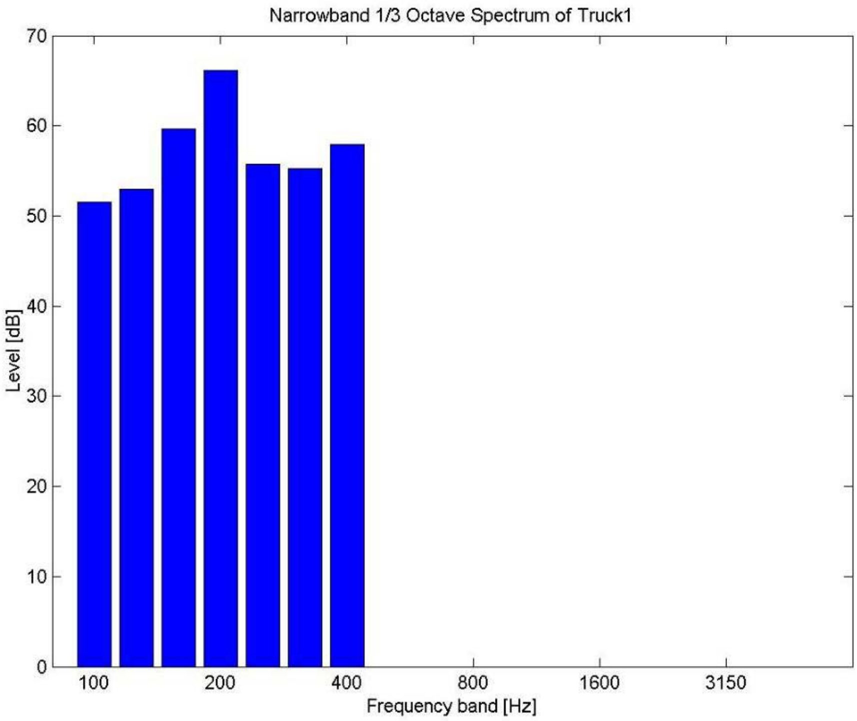


Figure 8.21: Narrowband 1/3 octave spectrum of Truck1 (20ms frames averaged over 125ms)

Once again the 1000ms analysis frame returned the closest groupings (Table 8.6). Apart from Truck5 having an abnormally low SF value of 9.6776, the remainder of the compression braking group have values between the range from 14 to 20. Should this range be used as an identification feature, nine of the eleven additional vehicles would be ruled out. However Truck5 would also be ruled outside this range. The use of this feature in a weighted application rather than a pass or fail standard could be useful.

Table 8.6: Spectrum Flux of 1000ms narrowband waveform

Wave File	SF
Car1	11.5268
Car2	10.518
Car3	11.0771
Car4	13.6865

Car5	10.9686
Car6	12.2338
Truck1	**16.7432**
Truck2	**14.0461**
Truck3	**19.6612**
Truck4	**18.5436**
Truck5	**9.6776**
Truck6	**16.5814**
Truck7	12.2632
Truck8	11.0012
Truck9	17.9045
Truck10	13.8063
Mot1	19.2461

8.9 Single Band Spectrum Flux

In the observance of the recorded signals, it was found that the waveforms would audibly shift in peak frequency over time. As explained in Chapter 2 this can be a result of Doppler shifting, however there are other contributing factors to this effect. The use of compression braking returns expelled pressure on the vehicles engine thus slowing the revolution speed and the engine frequency. It could be expected therefore that a heavy vehicle applying compression brakes would produce a waveform that not only shifted down in frequency due to the velocity of the vehicle but also the decreasing speed of the engine. A similar effect is observed where a vehicle is accelerating. Here the engine speed increases over the recording time and hence a positive shift in frequency is observed. Figures 8.22 and 8.23 below show the behaviour of selected frequency band amplitudes over the 1 second recordings for Car3 and Truck2. The frequency bands chosen for observation were the most dominant band in a 1/3 octave spectrum and the bands directly above and below it. It is seen in the figures that the behaviour if the band magnitudes over time in these cases tends to follow an upward or downward trend. This trend has been approximated using first order curve fitting in Matlab and is shown in the figures as a dashed line for the dominant frequency band, a dotted line for the upper band and a line of stars for the lower band.

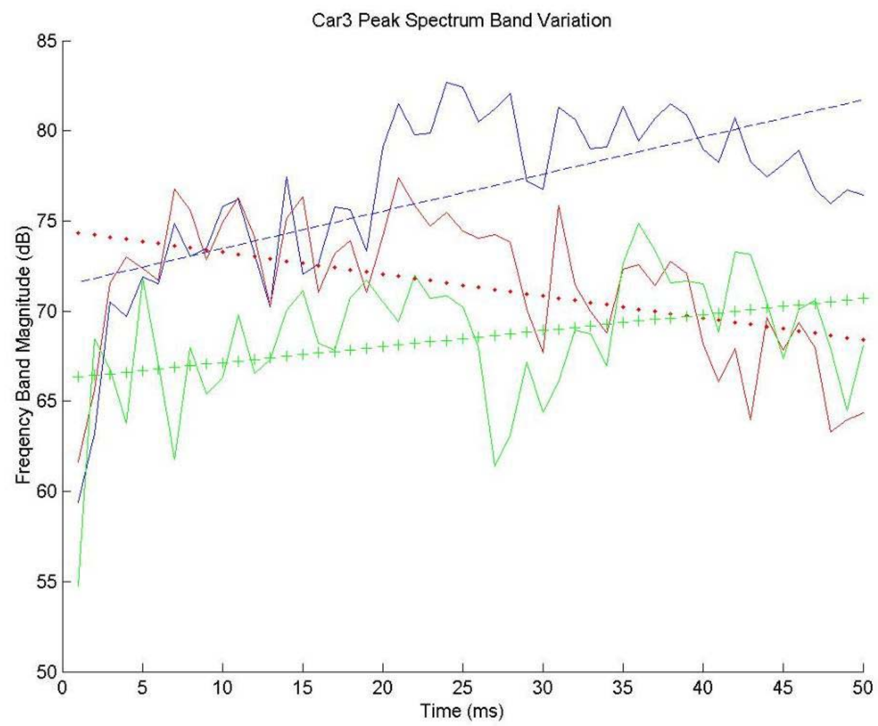


Figure 8.22: Single band flux of Car3

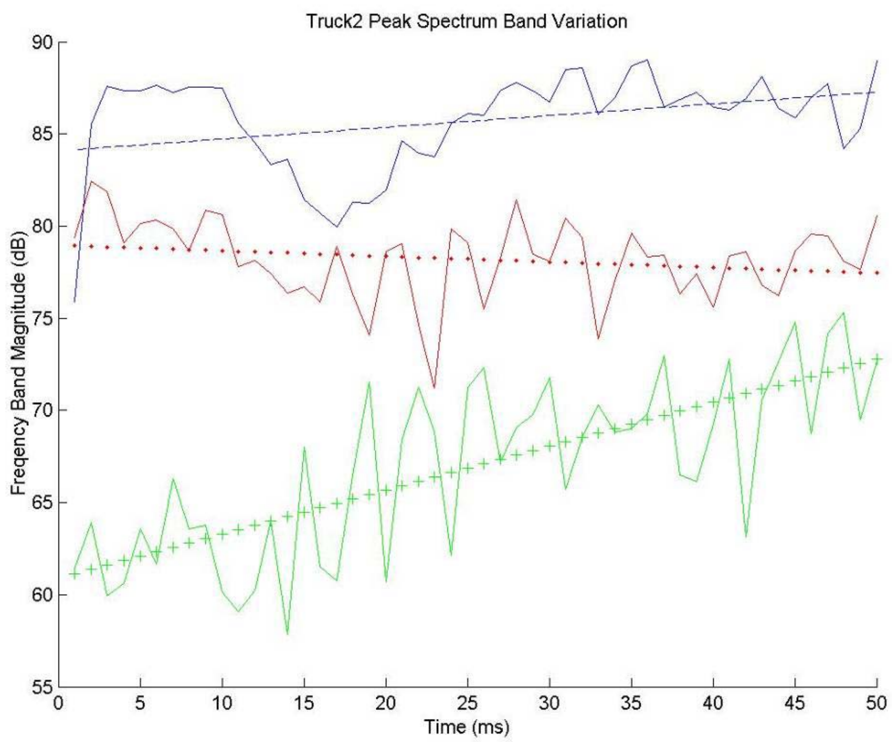


Figure 8.23: Single band flux of Truck2

The idea behind looking at these slopes comes from the observed shift in frequency in the signals over time. As explained in section 7.2.2 the further a vehicle is away from an observation point, the less the frequency of the vehicle signal will appear to shift. As the vehicle approaches the microphones of an array, due to the non-aligned path of travel of the vehicle, the velocity of the vehicle will appear to change with respect to the microphone. This causes a frequency shift. The greatest shift in frequency will be perceived where the relative velocity of the vehicle to the fixed microphone changes the most. This occurs in the region where the vehicle passes the point at which it is travelling perpendicular to the propagation of its sound signal to the microphone (Figure 7.12 section 7.2.2).

Clearly if all frequencies in the sound had the same magnitude, a shift from one frequency bin to another would be compensated by the frequency shifting into the bin losing another frequency and no effect would be perceived.

By focusing on one band in particular (in this case the band that exhibits the peak magnitude throughout the signal) and the neighbouring 1/3 octave bands, the shift in frequency may be observed where frequencies from the peak band shift into the adjacent bands. A shift due to engine operation will shift only the bands related to the engine whereas Doppler shifting affects the entire spectrum. The shift in frequency may be interpreted from the change in magnitudes of the peak (engine) bands over the signal period.

An increase in magnitude in the band either side of the peak band is not enough to indicate a shift in frequency in either the upper or lower directions. A more stringent interpretation would involve the relationship of all three bands. As explained above, all vehicles travelling towards the microphone will produce a signal that effectively shifts down in frequency at an increasing rate until they are travelling perpendicular to the source-microphone propagation path. Then as the vehicles move away from the observation point, the frequency will continue to decrease at a decreasing rate. It is also known that the source signal is dependant upon engine operation and that as compression breaks are applied, the slowing engine will result in a downwards shift in signal frequencies. These shifts have a similar nature and it would therefore be difficult to isolate compression braking based on a downwards shift in frequency. A rapid downwards shift (relative to that possible from Doppler shifting) could indicate the use of compression braking however any vehicle that applies breaks will exhibit a downwards shift in signal frequencies.

Rather than identifying compression braking, this process may be used to rule out signals that exhibit characteristics contrary to heavy vehicles using compression braking. In particular, heavy vehicles that are accelerating may be ruled out were they produce an upwards shifting spectrum due to increasing engine revolutions. This may be identified in the PSF data where the upper adjacent frequency band increases in magnitude accompanied by a decrease in the lower adjacent band. Figure 8.24 below exhibits this characteristic. Listening to the recording confirms that the vehicle is accelerating.

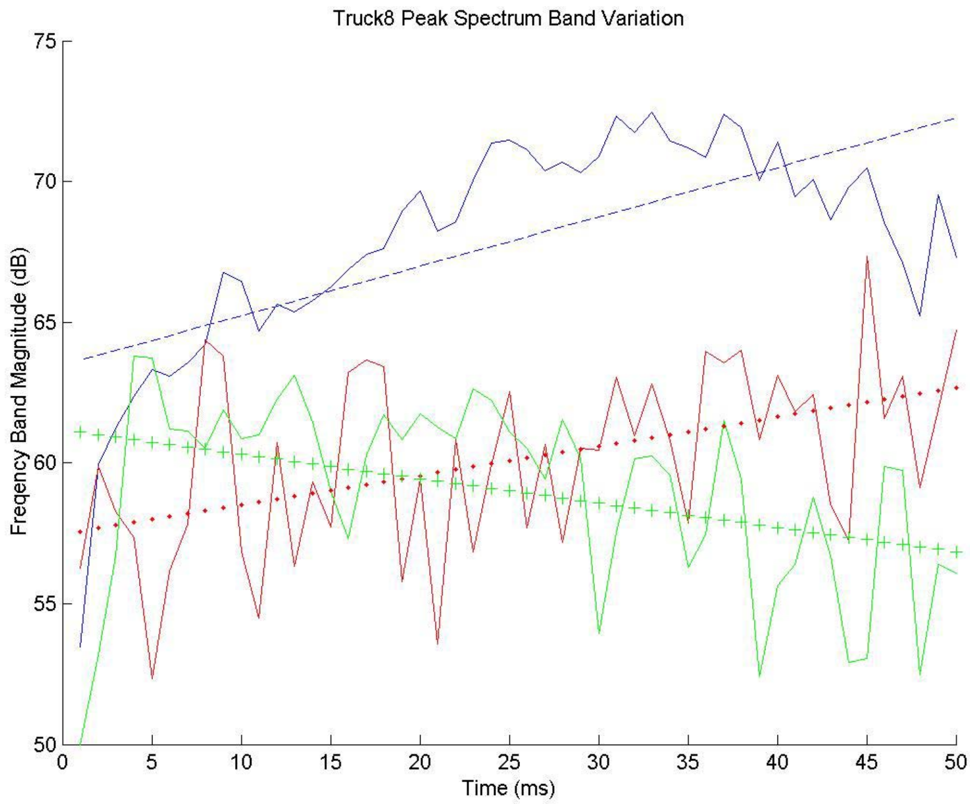


Figure 8.24: Shifts in frequency band magnitudes for Truck8 indicating acceleration

It was seen in many of the PSF plots that the peak magnitude frequency band tended to have a positive slope for most signals. It is believed that this is largely due to the selection of signal recordings where the vehicle was approaching the microphone, thus the total signal volume was increasing. This underlying trend should be taken into account when viewing the data as all frequency bands would have an increase in magnitude for the same reason.

Table 8.7 below lists the slope of the first order lines of best fit for the frequency band magnitudes over each one second recording.

Table 8.7: Single band flux trends and vehicle behaviour

Wave File	Slope			Estimated shift in Freq.	Vehicle Behaviour
	Lower Band	Peak Band	Upper Band		
Car1	-0.2222	-0.0263	0.1829	Up	Accelerating
Car2	-0.0734	0.0644	0.1836	Up	Accelerating
Car3	0.0888	0.2063	-0.1214	Down	Pass
Car4	0.0240	0.1702	0.1082	None	Pass
Car5	0.2385	0.1736	0.0082	Down	Pass
Car6	0.1346	0.1957	0.0365	None	Pass
Truck1	**0.1826**	**0.0696**	**0.0448**	**Down**	**Pass + Break**
Truck2	**0.2381**	**0.0637**	**-0.0304**	**Down**	**Pass + Break**
Truck3	**0.0484**	**0.0282**	**0.0329**	**None**	**Pass + Break**
Truck4	**0.0557**	**-0.1198**	**0.0120**	**None**	**Pass + Break**
Truck5	**0.1427**	**-0.0325**	**0.1285**	**None**	**Pass + Break**
Truck6	**0.0699**	**0.0497**	**0.1215**	**None**	**Pass + Break**
Truck7	0.1976	0.2153	0.0512	None	Accelerating
Truck8	-0.0867	0.1756	0.1041	Up	Accelerating
Truck9	0.0777	0.1139	0.1106	None	Pass
Truck10	0.0661	0.0855	0.1181	None	Pass
Mot1	0.0170	0.0472	-0.0386	None	Pass

Looking at the data

In table 8.7, the vehicle was classed as having frequencies shifting up, down or not at all. For a shift up in frequency, the consideration criterion was that the upper band had to be decreasing and the lower band increasing. This ruled out many vehicles whose upper band was increasing due to the approach of the vehicle to the microphone. For vehicles that had a downward shifting frequency, the lower band slope had to be positive and greater than the slope of the upper band + 1.2.

The data gave successful recognition of the accelerating cars and the accelerating truck Truck8. Truck7 which is also accelerating was not recognized by the technique most probably due to the heavy oscillations in signal magnitude cause by ‘chattering’ gears. All of the trucks using compression braking apart from Truck6 have a greater lower band slope than upper band slope (thus indicating decreasing engine revolutions). This feature used as a selection criterion would therefore select five out of the six heavy vehicles using compression braking.

As predicted, many of the vehicles have similar positive sloping band magnitudes. This indicates unchanging engine operation and a universal increase in frequency bands due to the close proximity.

One perceivable complication with this analysis is its frequency selectivity. Due to the low resolution of the 1/3 octave spectra, some shifts in frequency will pass from one band into another (thus causing a recognizable change in magnitude) where others will move within a frequency band and remain unrecognized. This problem exists only with low resolution frequency analysis. A larger number of frequency bins would allow for a more accurate observation of frequency shifting.

8.10 Noise Frame Ratio

It is seen in the autocorrelation figures (Figures 8.25 and 8.26 below) that the peaks of the waveforms represent various signal periods created by the fundamental frequency components. The frequency characteristics of the recordings are known from the 1/3 octave spectra. Therefore, this characteristic is of little interest with respect to the NFR (Noise Frame Ratio).

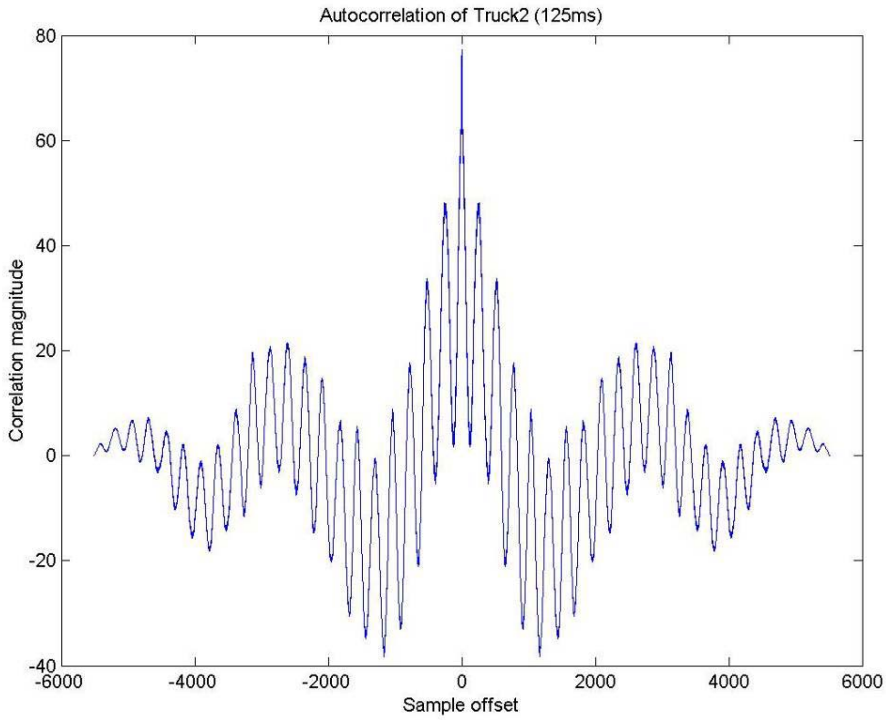


Figure 8.25: Auto-correlation of Truck2 (125ms) waveform

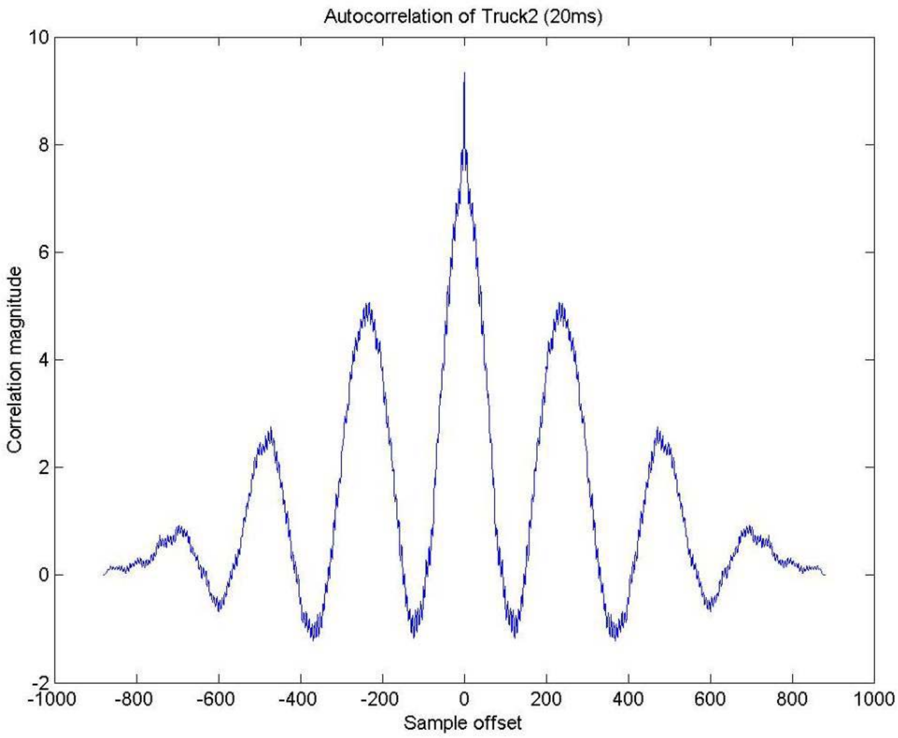


Figure 8.26: Auto-correlation of Truck2 (20ms) waveform

The figures above show that the correlation peak created at a sample lag of zero is not a single point peak. Rather, it has shoulders that extend to the nearest correlation minima. To aid in searching for the local maxima, the majority of the zero lag peak is nullified. The window size of eighty samples from the zero lag point was chosen to be nullified. This restricts the Matlab analysis code from detecting local correlation peaks for fundamental frequencies higher than 550Hz (using a sampling frequency of 44100Hz). Figure 8.27 shows the equivalent to the waveform in Figure 8.26 after the nullification process.

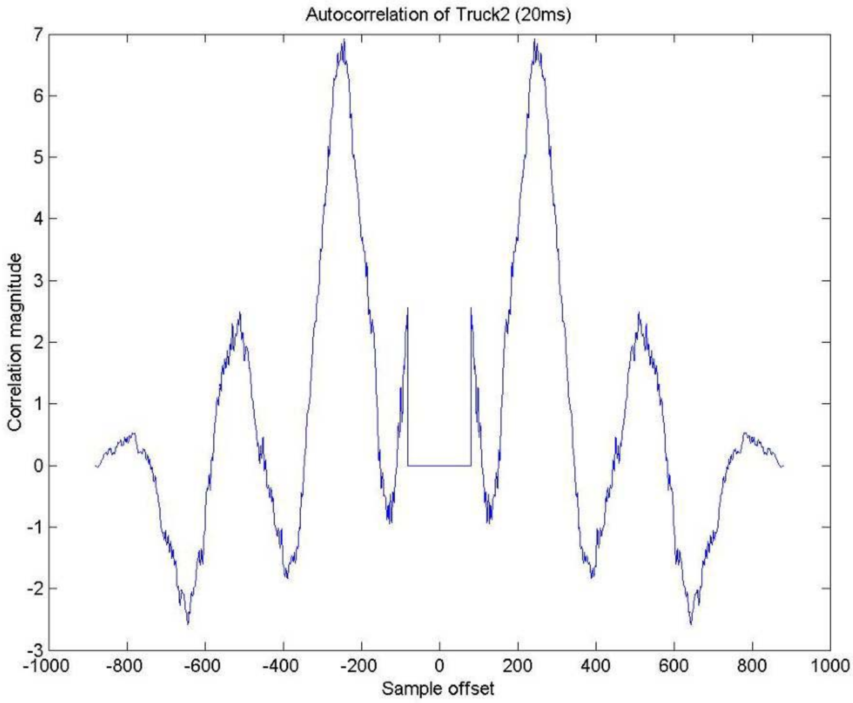


Figure 8.27: Nullification of zero lag peak

The noise frame ratios calculated using fifty auto-correlated 20ms frames in each 1000ms recording and various correlation cut-off magnitudes are given in Appendix E. In the cases of each cut-off value, the NFR of the compression braking group had distributed values across the range that encompassed the values for the other vehicles. The NFR used in such an application does not therefore provide notable class separability and does not appear to be of any use in recognition of compression braking. Using a 125ms frame provided no further class separability. One use for the NFR value may be to classify recordings as clean or noisy, thus giving a gauge for the reliability of the data in the frames.

8.11 Third Octave Spectra

The third octave frequency spectra were obtained for the 125ms recording of each vehicle. The 125ms recording represents the most characteristic acoustic emission for the operation of the vehicle and therefore gives the most accurate 1/3 octave frequency response. The waveforms for each of the spectra are given in Appendix F. From these waveforms it is immediately obvious that vehicles even within specific classes produce highly varying 1/3 octave spectra. The proposed method for direct comparison of spectra for classification or recognition would produce little accuracy. This is not only due to the large number of different spectra that represent one vehicle type and operation (i.e. compression braking of heavy vehicles) but also due to there being vehicles in other classes that produce a similar spectrum to that of vehicles in unrelated classes.

As expected, those vehicle recordings with a notably high content of tyre on road noise and other high frequency sounds have large components in the upper bands of their spectrum. It has been stated previously that the upper bands of the spectrum may be disregarded when we are only interested in the frequencies characteristic to the engine and exhaust operation. It was decided therefore to examine the frequency bands across the 3 octaves 50Hz to 400Hz in the derived spectra. Furthermore, from the possible features listed, the band levels were considered for class separability. Table 8.8 shows the three highest bands within the range of 50Hz to 400Hz for each spectrum.

Table 8.8: Peak frequency bands between 50Hz and 400Hz

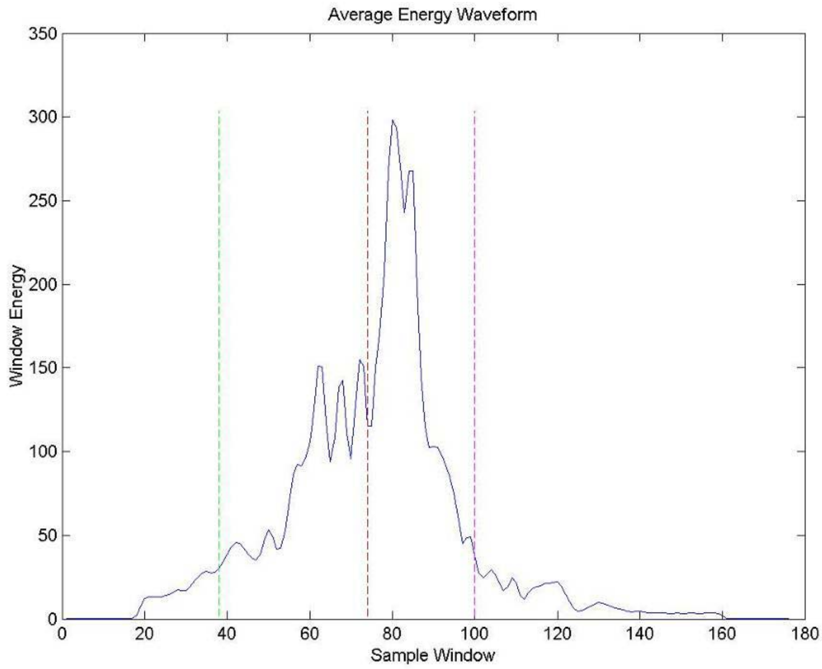
Wave File	3 highest frequency bands			
	1^{st}	2^{nd}	3^{rd}	
Car1	160	125	315	No
Car2	160	125	200	No
Car3	160	200	80	Yes
Car4	160	80	125	No
Car5	160	315	200	No
Car6	160	315	250	No

Truck1	**200**	**160**	**400**	**Yes 2**
Truck2	**160**	**200**	**250**	**Yes 3**
Truck3	**250**	**200**	**160**	**Yes 3**
Truck4	**250**	**160**	**315**	**Yes 2**
Truck5	**200**	**160**	**125**	**Yes 2**
Truck6	**125**	**315**	**63**	**No**
Truck7	160	250	200	Yes 3
Truck8	125	63	200	No
Truck9	80	160	100	No
Truck10	63	80	125	No
Mot1	250	125	160	No

One particular trend in this data is the appearance of two of the three frequency bands 160Hz, 200Hz and 250Hz as the bands having the greatest magnitudes. This is most probably due to the pitch of the applied compression braking being in the range of 160Hz – 250Hz. All trucks using compression braking except for Truck6 were seen to have this feature. It was also noted that some of the other vehicles (Car3 and Truck7) exhibited the same identifying feature. This feature applied along with other techniques may be valuable in the recognition of vehicles using compression braking.

8.12 Energy waveforms

The time domain energy envelopes for two trucks using compression braking are displayed in Figure 8.28.



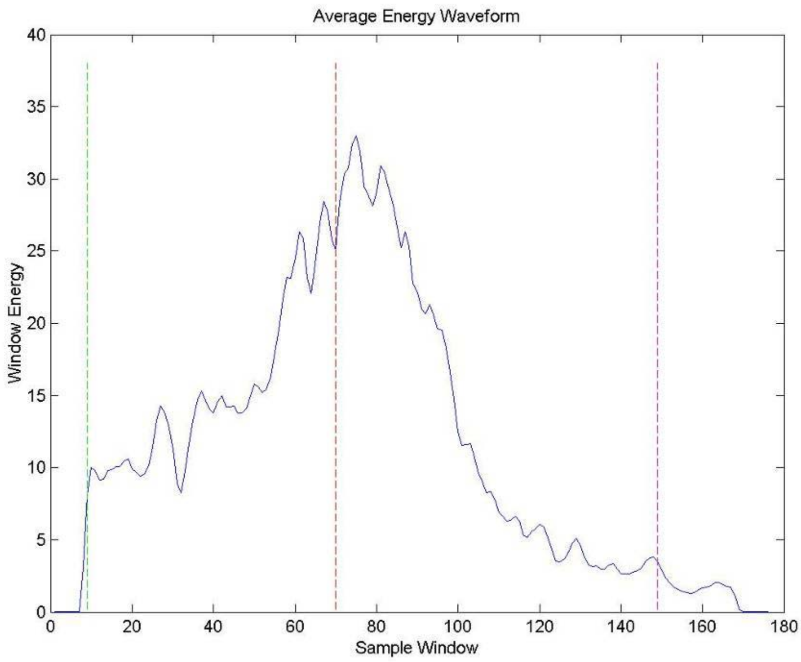


Figure 8.28: energy envelopes for passing trucks using compression braking

The two truck waveforms exhibit a similar characteristic in the fact that the centroid of energy lies behind the peak signal amplitude. This is undoubtedly due to the configuration of a heavy vehicle. As the truck passes, exhaust noise is more easily heard from behind. Apart from this similarity, it can be seen that the waveforms vary a great deal and that correlations will be difficult to make. This is amplified in an environment with varying source speed and operation. Variations in magnitude may result from vehicle operation or unique vehicle noise contributors.

The figures above were created using Matlab function coding. For these particular signals, the start point was specified as the first window with an energy value equal to or grater than 10% of the maximum energy value (Emax). Likewise, the end point is specified as the last window with an energy value greater than or equal to 10% of Emax.

Processing Lag

The energy envelope analysis technique relies upon a single vehicle signal free of reflections and noise. The computation of the features listed above requires that the vehicle has passed by the recording point (in fact 1-3 seconds past). Therefore image capture would have to be pre-emptive or based on a sound level before classification had taken place. However, this then forfeits the ability to detect a vehicle under the noise limit but creating offensive noise. This problem could be overcome however by continually capturing images of the scene in a circular buffer. If a vehicle is detected, the lag in the system could be used to retrieve and save the matching image before it is overwritten. In such a configuration, computation is not necessarily immediately time dependant and therefore time domain processing is a viable option. The benefit of using a complete pass-by signal for classification analysis is that the complete acoustic fingerprint of a vehicle is considered. Therefore a greater scope of information is available concerning each vehicle. This is because the signal is comprised by the noise emitted by the front, middle and rear sections of the vehicle as it passes by. Evidently, the single point noise source assumption is disregarded here. The problem still exists of isolating a single vehicle's acoustic emissions for analysis. This is unlikely without complex filtering

Single vehicle dependence

One clear problem with using the time variant energy envelope for classification is that a recording of a single vehicle pass by is required. Additional vehicles recorded in the signal will alter the energy levels dramatically. Furthermore, the system cannot define where the noise from one vehicle ends and another one starts. This will make the selection of the index points (start point, centroid and end point) for accurate comparison difficult if not impossible. For this reason it is considered that the time domain energy envelope analysis would be more suitable in an environment where only single vehicle pass-bys occur. The technique has not been further examined.

CHAPTER 9 - VEHICLE DETECTION

9.1 Australian Standards

In order to create some alignment with standard Australian sound level limit measurement procedures, a vehicle noise measurement system should be designed with due respect to the considerations outlined in the relevant Australian Standards Documents. The Australian Design Rule 28/01 – "External noise of motor vehicles" was introduced in section 3.2.2 as the Australian compliance standard for all Australian motor vehicle noise emissions. The measurement procedure of the Design Rule specifies the use of a SLM in compliance with 'Type 1' of AS1259-1982 (superseded by AS1259-1990) applying an 'A' weighting and 'fast response' or 'F' time constant.

Some of the sound level limits specified by the ADR are given below in table 9.1

Table 9.1: ADR 28/01 sound level limits

Vehicle Type	Sound Level Limits for Vehicle in motion (dB(A))	
	Diesel	Spark Ignition
Motor Cycle	-	82
Passenger Car	78	77
Forward Control Passenger Vehicles	80	79
Off Road Passenger Vehicles	80	79
Light Omnibuses	85	85
Heavy Omnibuses	86	86
Light Goods Vehicles	82	81
Medium Goods Vehicles	86	86
Heavy Goods Vehicles	87	87

Note that the table shows the maximum sound level assigned for each class of vehicle. Within a class, the maximum value usually belongs to the vehicle make having the highest net engine power or GVM.

The measurement procedure of a passing vehicle's noise emission is also defined by the ADR. The document states that the measurement should be made at an open site where the ambient and wind noise levels are at least 10dB(A) below the noise level being measured. The configuration for measurement of a passing vehicle is illustrated in Figure 9.1. The standard requires that the vehicle be driven full throttle from line AA until it passes completely line BB. Sound level measurements are to be taken at the points marked on the line PP whilst the vehicle is in the designated recording area. The microphone position lies exactly in between the measurement start and finish lines which are 20m apart and set back 7.5m from the line of travel. The microphone must be 1.2m above the track surface and the standard requires that measurements are taken from both sides of the vehicle.

At least two measurements must be taken on both sides of the vehicle with measurements containing spurious measurements repeated. The maximum sound level recording during this procedure constitutes the measurement result.

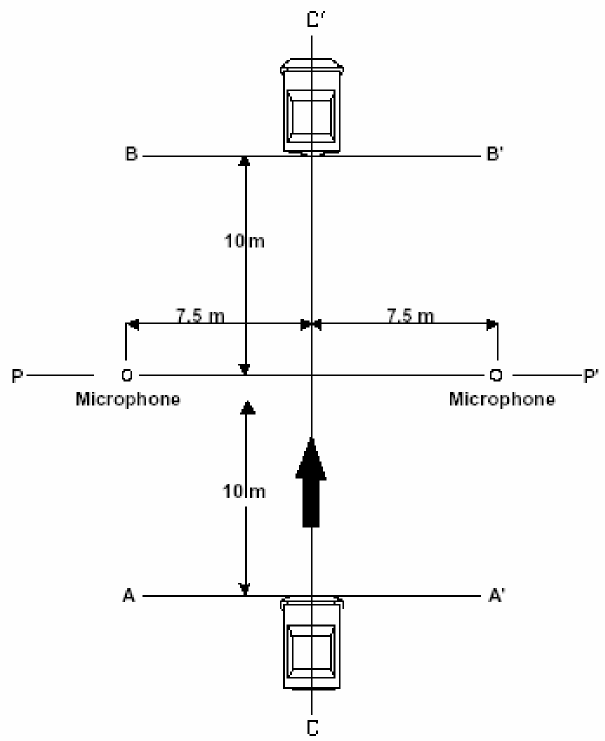


Figure 9.1: ADR recording procedure for a moving vehicle

The ADR measurement procedure adopts a leniency of 1dB(A) from the recorded levels to compensate for possible lack of instrument precision. If a vehicle exceeds the designated noise level limit by more than 1dB(A) and the difference between two consecutive noise level measurements is no greater than 2dB(A), the vehicle is tested again. If when tested again more than one out of four measurements exceed the prescribed limits, the vehicle does not comply with Australian Standards for external noise emissions.

9.2 Compliance with Australian Standards

The capacity for a roadside noise level meter to record sound levels that satisfy the requirements outlined in the Australian Standards discussed is quite low. There are steps that may be taken in system design however to provide measurements as compliant as possible.

The requirement that the ambient and wind noise levels are no greater than 10dB(A) below the noise level measured can be implemented in the system by estimating ambient levels using the noise levels recorded where no vehicles are present. The sound level meter could be positioned 1.2m above the road surface and 7.5m from the nearest carriageway. Positioning the device in an open site will help to reduce additional acoustic interference.

The ADR requires that tested vehicles be operated full throttle in the measurement region thus producing a maximum noise emission. Vehicle operation cannot be controlled on the roadway however it is likely that most cars will not be operating full throttle and will therefore produce a sound level lower than that if measured according to Australian Standards. This suggests that if the recorded levels of a normally operated vehicle exceed the prescribed limits, the noise level at full throttle would most definitely exceed the limit. The measurement procedure requires that more than one measurement of the sound level be taken. One of the reasons for this is to eliminate erroneous measurements form possibly faulty instruments or field conditions. With the use of a number of microphones, it may be possible to compare several recorded sound levels to eliminate faulty levels. A smaller microphone array aperture in this case will yield more similar sound levels.

9.3 Australian Standards for SLM's

Sound level meters indicate the sound pressure level as a short time R.M.S average of the continuous signal. The time constant used for this averaging can be varied depending on the measurement application. As specified in ADR28/01, a fast 'F' time constant is desired where monitoring vehicle noise.

The averaging time used for a fast response time constant is 125ms as specified in the AS IEC 61672.1-2004 meaning that an equivalent continuous sound level is given for every 125ms. This helps to define the processing and measurement rate of the system.

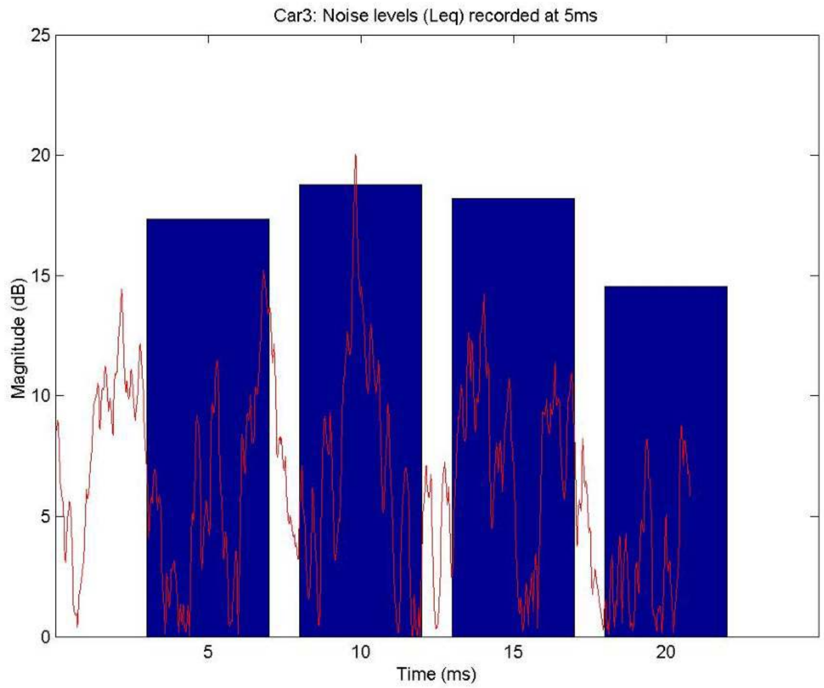


Figure 9.2: 5ms averaged Leq waveform

9.4 Multiple Vehicles

The sound level perceived at any point on a roadway will consist of a combination of sound levels from surrounding noise sources. It cannot be assumed in an omni directional receiving environment that the sound pressure level recorded belongs to only one vehicle. The acoustic properties in section 2.1.1 describe the ways in which multiple sound levels interact. The accurate and isolated measurement of an individual vehicle's radiated sound pressure level will be extremely difficult to attain in the uncontrolled multi-source environment. There are a number of measures that can be taken to avoid the false detection of noisy vehicles.

9.4.1 Methods for avoiding false identification

The addition of incoherent noise levels will occur on any roadway and there is a chance that the contribution from other vehicles could force the noise level assigned to the dominant vehicle to exceed the detection threshold (whatever it may be). The probability of this occurring will be fairly low on roadways that experience un-congested traffic flow. This is due to the sharp decline in signal amplitude over distance as well as the nature of incoherent sound level addition.

It would take 5 cars producing a level of 70dB each to appear as 77dB. This would require that all of the vehicles were in close proximity and that the actual sound level produced was greater than the 70dB received at the sensor due to the propagation attenuation. This scenario is of little consequence as the number of vehicles will prevent the accurate localization of any individual source.

Location Accuracy Factor

It was seen in the investigation of the cross-correlation between two microphones in a multiple vehicle environment in section 7.8 that the correlation values would vary according to the clarity of the dominant signal. It was seen that where more than one dominant signal was present or where the signal was not clearly correlated, the peak correlation value was reduced. The magnitude of the peak cross-correlation value may then be used not only as an indication of the accuracy of the localization process but also as an effective indicator for the presence of multiple vehicles.

By implementing a location accuracy factor, the correlation peak would have to exceed a certain magnitude before a noisy vehicle is analysed. A low correlation peak would indicate that a vehicle has been localized but there is much disturbance present and the location could be erroneous. A high correlation value would indicate the isolated nature of a vehicle. The results in section 7.8 also show how the smooth transition of the TDOA curve can be upset by the presence of multiple sources or reflections. The curve continuity may also be utilized as an accuracy factor.

A detection problem could occur however where a dominant vehicle provides a signal strong enough for accurate localization but is forced over the detection threshold by the presence of other vehicles. As sound pressure levels increase in decibels, this is again highly unlikely. Further leniency in the detection level could be applied to confirm beyond a doubt that the located vehicle is exceeding the allowable limit.

9.5 Extrapolating Sound Levels

One important measurement procedure in achieving compliance with Australian Design Rule values is a measurement distance from the source line of travel to the microphone of 7.5m. This may be accomplished by placing the microphone array at a distance of 7.5m from the centreline of the nearest carriageway. Noise levels recorded from vehicles on adjacent lanes or where a proximity of 7.5m is not possible however will not be represented accurately. Using the location information (or simply the lane number) of the contributing vehicle, the recorded noise level may be approximated using the equation derived below. In this way the noise level is "normalized" to a lane centreline distance of 7.5m.

The sound pressure level of a source may be expressed as (Bies & Hansen):

$$L_{P(r)} = L_{\omega} + 10\log(4\pi r^2) \tag{9.1}$$

where L_{ω} is the sound intensity (in dB) at the source and *r* is the distance from the source to the observer. The "normalized" 7.5m SPL is then found as follows

$$\begin{aligned}
&L_{P(7.5)} = L_{\omega} - 10\log(4\pi \times 7.5^2) \\
&L_{P(d)} = L_{\omega} - 10\log(4\pi \times d^2) \\
&\therefore L_{P(7.5)} = L_{P(d)} + 10\log(4\pi \times d^2) - 10\log(4\pi \times 7.5^2) \\
&\qquad = L_{P(d)} + 10\log\left(\frac{4\pi \times d^2}{4\pi \times 7.5^2}\right) \\
&\qquad = L_{P(d)} + 10\log\left(\frac{d}{7.5}\right)^2 \\
&\qquad = L_{P(d)} + 20\log\left(\frac{d}{7.5}\right)
\end{aligned} \tag{9.2}$$

Where *d* = the distance from the vehicle lane to the microphone in meters.

The adjusted sound level is only available after the acquisition of the source location. This presents the decision then of whether to detect only the initially perceived noisy vehicles or whether to locate all dominant vehicles and then initiate the detection process after obtaining the adjusted noise levels. As the detection process relies fundamentally on the ability to accurately locate the excessively noisy vehicle, it would be appropriate to place the localization process first in the order of detection.

9.6 How and when to trigger a measurement

Location Accuracy

The detection process is dependant on the ability to accurately locate the offending source. The localization process must then be the primary step in the process. The location accuracy factor informs the system of the accuracy of the cross-correlation localization. Where the factor is below the accepted level, further detection processing is not required and the system would continue to monitor the roadway. When the location accuracy factor is above the accepted level, the sound level of the located source could then be approximated using the normalization process discussed above. The normalized sound level would be the level used then for adjudicating the vehicles compliance.

Having established this however, it may be possible to exclude the location accuracy factor from the detection of heavy vehicles using compression braking. It is not likely that such vehicles will be drowned out by other noise sources and will most probably provide coherent and accurate localization. However if the sound is corrupted and the location accuracy factor is below the identification limit, (which exists to ensure correct noisy vehicle localization) an image captured would quite possibly feature only one heavy vehicle capable of using compression braking. In such a case the vehicle could then be verified and identified. It could occur however where two trucks engage compression braking whilst travelling side by side down a hill. If the localization accuracy is compromised here, manual identification of the offending vehicle would not be possible.

Sound Level

The fast (125ms) averaged sound level of a passing truck is shown in Figure 9.3 below. It is seen in the waveform that

- A number of the Leq values exceed the peak level of 87dB.
- The peak exceedance period is relatively flat and lasts for around 1.5 seconds

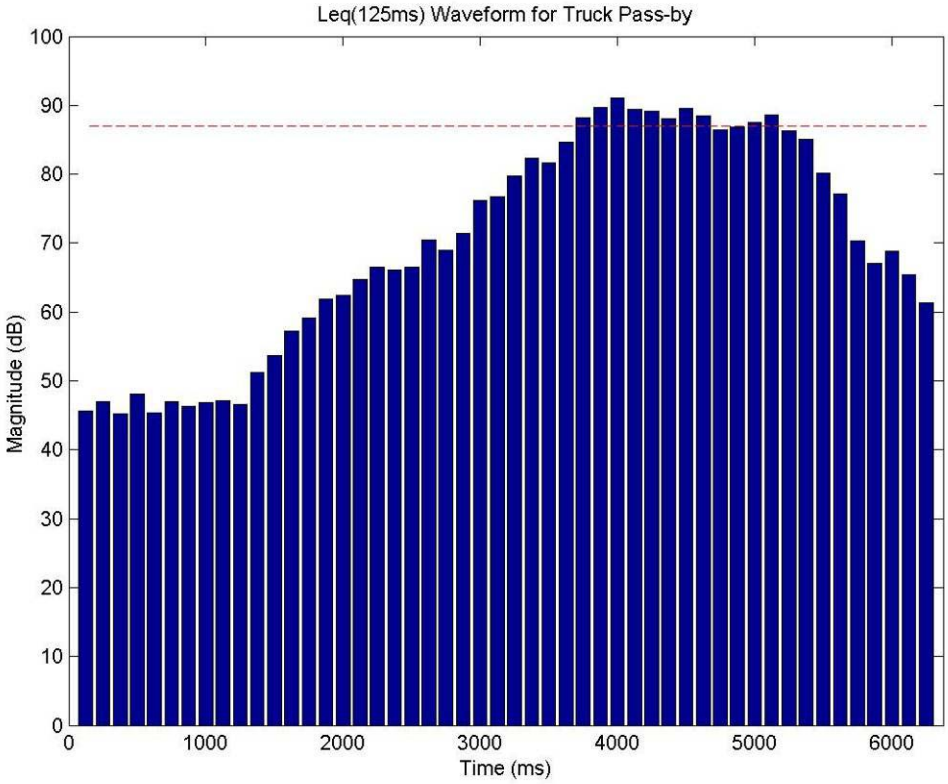


Figure 9.3: 125ms Leq(A) waveform for a truck pass-by

These observations suggest that without precautions, the same vehicle might be detected several times during a pass-by. In order to avoid multiple detections, a simple logic process should be implemented which detects the vehicles responsible for peak levels every few seconds. The location information may also be used to avoid multiple detections of a vehicle. A proposed detection logic process is illustrated in Appendix C.

9.7 Tolerance for different Vehicles

The maximum allowable noise levels specified in the Australian Design Rule ADR2801 state differing acceptable noise levels for various vehicle classes. The maximum noise limit for a moving vehicle is 87dB and the lowest is 77dB. The ability to classify vehicle type by using the acoustic recordings could allow the system to apply the different level restrictions to different vehicle classes. However there are a number of current technologies that effectively classify passing vehicles. These include pneumatic tube counters, infrared sensors, inductive loop counters and even digital video analysis.

Without vehicle class information however, only one detection level can be implemented by the system. This level should be set by the user to reflect their measurement needs or the type of traffic on the roadway.

For example, if the device is used on a roadway that undergoes a large amount of heavy vehicle traffic, the highest priority detection is most likely noisy trucks or those using compression braking and a high threshold would be used. Where noisy vehicles are concerned, the device would be used where heavy vehicle traffic is infrequent and therefore poses little threat to erroneous measurements. The level detection threshold in this case could be set much lower.

9.8 Configuration Considerations

A noise sensor positioned at 1.2m on the roadside (as specified by Australian standards) would be subject to attenuation from other vehicles on the road. As with the noise camera discussed in Chapter 3, mounting the device above the traffic flow will assist in removing the effects of object related attenuation. The placement of electronic equipment (particularly of the law enforcing variety) above accessible reach will also help to reduce the risk of vandalism. With the knowledge of its height and offset from the roadway, the system could derive the 7.5m normalized sound level values.

CHAPTER 10 – CONCLUSION

The aim of this research project has been to investigate the primary concepts behind a noisy vehicle detection system. The primary investigation of the report introduces the theory relative to outdoor traffic noise such as sound pressure levels, propagation and attenuation, the available acoustic measurement and processing equipment and related research and technology. The latter section of the report has detailed the application of noise measurement technology and theory to the determination of the noise type, source location and noise level of vehicles moving on a roadway.

10.1 Background

The outdoor sound propagation theory discussed for the project considers each vehicle as a single point noise source producing spherically propagating sound waves. This introduces an attenuation characteristic of acoustic intensity inversely proportional to the distance from the source squared. The addition of logarithmic sound levels was discussed and defined in equation 2.6 where vehicle signals are considered as incoherent. The effect of Doppler shifting on a moving sound source's frequency components has also been introduced and considered in the applications of the project.

Certain noise source elements such as tyre on road noise were also investigated. These contributing factors are considered to have an effect on the noise produced by passing vehicles and hence the design of a roadside noise monitoring system. It was seen that differences in road surfaces can result in tyre/road noise emission differences of up to 12dB(A). The condition of the site specific road surface should then be considered in the operation of a roadside noise analysis device. In addition to road surface texture, the presence of wet patches on the road and the immediate speed limit are also factors worthy of consideration.

Precision sound level meters were introduced as the industry standard for sound pressure measurement. The use of sound level meters in a roadside measurement device allows for accurate and comparable measurement levels as well as a range of possible frequency analysis functions. Further sections of the project were based on the standard Queensland road width of 3.5m as specified in the document *Road Planning and Design Manual* (2004).

Two existing devices used for roadside traffic noise monitoring were investigated and presented in the report, the Noise Camera and the Acoustic Sensor. The Noise Camera had several features which were considered to be appropriate to the design of a noisy vehicle detection system. These included its mountability on existing roadside structures, wireless access and its mounting position above the traffic flow. The features of the Acoustic Sensor that were considered appropriate to this design topic were once again a mounting position above the traffic flow as well as its use of signals processing to create microphone directivity.

Some of the existing research pertaining to the scope of this project was discussed. One report focused on a low resolution vehicle classifier using the acoustic energy waveforms obtained from passing vehicles. Another report discussed looked at using TDOA measurements from a microphone array to approximate the far field angle of incidence of a passing jet aircraft.

10.2 Localization

A number of signal properties that may be used for the localization of a moving sound source were investigated. The signal time delay was chosen as the most appropriate signal property. Unlike the signal amplitude, the time delay between a source and microphone remains relatively un-affected by the presence of physical objects in the propagation path. The time difference of arrival (TDOA) of a signal between two microphones may be derived using the cross-correlation of the signals from two microphones.

Two approaches to source localization using the TDOA are the far field and the near field techniques. The far field approximation states that the source is far enough from the microphone array for the incident wavefront to be considered as a straight line. The near field approach considers the sound signal to have a true circular wavefront and relies on more complex derivation. The near field localization technique was chosen as the most appropriate for a roadside system. The method provides grater accuracy and allows for a wider array aperture. The microphone array was illustrated in Figure 4.6 as having a greater accuracy with larger aperture sizes. A near field localization technique was designed for a three microphone array (Figure 4.5). An example optimization code for the process is given in Appendix I.

Time domain cross-correlation was selected over the cross-spectral technique for time delay estimation. This decision was made due to the robustness and peak value definition of the time domain technique as well as its compatibility with restricted lag processing. The use of a restricted lag in the correlation process can reduce the number of computations required by more than a factor of 42. The time domain cross-correlation algorithm was tested using simulated vehicle waveforms. These simulations highlighted the need for an array aperture of 1m or less to avoid un-correlative signal properties created by the varying moving source time delay. This phenomena and its reduction with a small array aperture were confirmed from the correlation of recorded dual point field signals. Multiple vehicle sound fields were considered and dual source pass-by simulations were conducted. The dual source simulations were used to derive the sound level difference required between two vehicles for the accurate localization of the farthest source. Simulated results show that a sound level difference of 5dB is great enough to define a noisy vehicle in the far lane of a four lane roadway where another vehicle is travelling in the first lane. Other such differences were derived for alternate lane configurations but are expected to vary in real field applications. Further field testing would refine the signal differences required however knowledge of such would not improve the accuracy of the system but rather define its limitations.

10.3 Recognition

It was decided that the system being developed has a greater need to identify vehicles that belong to a certain sound class rather than classify them even if they are not offending. For this reason, the sound characteristic research was oriented towards vehicle operation recognition rather than vehicle classification. More specifically, the recognition of heavy vehicles using compression braking was chosen, both as a research example and due to the practical priority in the field.

The possible features for recognition of heavy vehicle compression braking that were selected for investigation are:

- Zero-crossing rate (ZCR)
- High zero crossing rate ratio (HZCRR)
- Noise level modulation rate (NLMR)
- Low and high short time energy ratio (LSTER & HSTER)
- Spectrum flux
- Noise frame ratio
- Third octave spectra comparison
- Time domain energy envelope

With the issue of class separability in mind, a wide range of vehicle types and operations were captured in the form of roadside pass-bys. The selected recordings were analysed according to each of the features above. The feature extraction processing frames were restricted to 20ms, 125ms and 1000ms to provide long time and short time operation information. The analysis of the data set isolated some possible signal features that may be used for the acoustic identification of compression braking.

Using these features, a recognition algorithm was developed that best selected the compression braking vehicle group from the test set. The recognition process applies or deducts probability weightings for each of the identifying features. Table 10.1 shows the features and detection limits used in the algorithm which is derived in Appendix B.

Table 10.1: Proposed compression braking recognition features

Test	Condition	Applied probability weighting
1	If >3dB Leq modulation rate is greater than 25 in 1000ms	-2
2	Averaged normalized Leq(5ms) level lies between 0.40 –	+2

	0.55 in a 1000ms window	
3	LSTER (5ms frames in 125ms window) > 0.28	-2
4	HSTER (5ms frames in 125ms window) > 0.16	-2
5	Narrowband Spectrum Flux lies within the limits of 14 – 20 in a 1000ms window	+2
6	Accelerating	-10
7	Two of the three highest frequency bands between 50Hz and 400Hz are 160Hz, 200Hz and 250Hz (125ms frame)	+2

The proposed recognition process successfully identifies 100% of the compression braking group and excludes 100% of the non-compression braking vehicles. The accuracy of the system will not be known however until a large test group is analysed from which the detection limits have not been derived.

10.4 Detection

The final section of this report has investigated the requirements for noisy vehicle detection. It was decided that close compliance with Australian Design Rules sound measurement procedures are required where comparison with Australian sound level limits are to take place. The specifications provided by Australian Design Rule 28/01 for moving vehicle sound level measurement were discussed along with further requirements for compliant sound level instrumentation. Although any monitored roadway will be far from the controlled recording environment specified in the Design Rule, certain steps may be taken to provide reasonable compliance. One such step involves approximating the recorded sound level as a standard 7.5m offset measurement using a vehicle's location information or lane number.

The roadside detection process will involve a logical sequence of tests and precautions to avoid false detection or multiple detections of a single vehicle. It was decided that the localization accuracy of a vehicle's detection takes primary importance in the detection process. In-accurate location information will ultimately lead to false or illegible vehicle detection. Two features identified as location accuracy indicators were the time delay estimation correlation peak magnitude and the TDOA waveform continuity.

It was seen that where more than one dominant signal was present or where the signal was not clearly correlated, the peak time delay estimation correlation value was reduced. The magnitude of the peak cross-correlation value then presents itself as an indicator of the accuracy of the localization process and also an effective indicator for the presence of multiple vehicles. It was also seen thought the pass-by simulations that the peak correlation magnitude is specific to each lane.

The simulation results also showed how the smooth transition of the TDOA curve can be upset by the presence of multiple sources or reflections. The curve continuity or more so, the estimated source location continuity could be used as an indication of location accuracy.

Another precaution necessary in a detection system is the avoidance of multiple detections. The pass-by Leq observations suggest that without precautions, the same noisy vehicle might be detected several times during a single pass-by. In order to avoid multiple detections, a simple logic process should be implemented which detects the vehicles responsible for peak levels every few seconds. Once again, the location information (if accurate) may be used to prevent multiple detections of a single source. The proposed detection system is given in Appendix C.

10.2 Recommendations for future research

As mentioned previously, a large field data set is required for the assessment of the individual detection system algorithms as much of the system parameters have been designed to comply with the original data set.

With a larger and more diverse test group, the recognition feature tolerance values can be refined. Further research may even increase the source recognition capacity to more than one vehicle operation type. A larger dual microphone test group can be used to asses the localization technique and to determine appropriate values for location accuracy factors in various lanes. The robustness of the system can then also be tested in a multiple vehicle environment containing even multiple noisy vehicles.

Further culmination of test data may also be used to determine appropriate noise level thresholds for detection in different traffic environments. This process will require the use of precision sound level equipment.

Positive progress in the development of a noisy vehicle detection system may also involve research into passive acoustic speed detection or integration with existing traffic classifier technologies.

APPENDIX A – PROJECT SPECIFICATION

University of Southern Queensland

FACULTY OF ENGINEERING AND SURVEYING

ENG 4111/4112 Research Project

PROJECT SPECIFICATION

FOR:	**BENJAMIN PHILLIP CURTIS**
TOPIC:	NOISY VEHICLE DETECTION SYSTEM
SUPERVISORS:	Chris Snook, USQ Toowoomba
TECHNICAL ADVISORS:	John Leis, USQ Toowoomba
	Jeshua Brouwer, Main Roads Brisbane
ENROLMENT:	ENG4111 – S1, D, 2006
	ENG4112 – S2, D, 2006
PROJECT AIM:	The Project aims to investigate the acoustical properties of noise emitted by vehicles travelling along a roadway and to develop a method of identifying noisy vehicles using technology that can be mounted on the roadside.
SPONSORSHIP:	Department of Main Roads

PROGRAMME: **Issue A, 24th March 2006**

1 Research available information on the acoustical properties of moving vehicles and any relevant roadside noise measurement test procedures or standards.

2 Critically analyse the possible configurations for roadside noise recording and select the most appropriate

3 Gather data from various test sites covering a range of roadside noise scenarios using the developed testing procedure.

4 Analyse the acoustical properties of the field data and hence determine the possible processing scope provided by the chosen recording procedure. If necessary, reassess and improve the roadside recording configuration.

5 Produce a set of algorithms that may be implemented using on site technology to identify noisy vehicles and their whereabouts on the carriageway.

As time permits:

1. Produce algorithms that can classify vehicle types and noise sources
2. Coordinate the detector with a camera to capture an image of identified vehicles
3. Produce a source file code that implements the algorithms and logs noise data

APPENDIX B – COMPRESSION BRAKING RECOGNITION SYSTEM

The features which provided the greatest class separability were combined into a recognition system using the class dependant levels obtained in the experiments.

Leq Modulation Rate

Exclude if >3dB Leq modulation rate is greater than 25 in 1000ms

Or

Exclude if >5dB Leq modulation rate is greater than 6 in 1000ms

Results

Passes all compression braking trucks

Fails 3 out of 6 cars (car1, car3 and car4)

Fails 2 out of 4 Trucks (truck8 and truck10)

Fails Motorbike

Averaged normalized Leq level

For the 1000ms waveform, include if LeqAvNorm is within range:

0.45 – 0.53 (Tight limits)

Passes compression braking group and Truck7

(100% group selectivity) (9% false selection)

0.40 – 0.53 (Less tight limits)

Passes compression braking group and Truck7

(100% group selectivity) (9% false selection)

0.40 – 0.55 (Loose limits)

Passes compression braking group, Truck7 and Car6

(100% group selectivity) (18% false selection)

0.00 – 0.53 (Low pass limits)

Passes compression braking group, Truck7 and Truck9

(100% group selectivity) (18% false selection)

A detection process that falsely selects cars rather than trucks is more desirable. A car can be easily discarded as using compression braking either by image observation or other feature extraction values. The loose limits of 0.40 – 0.55 then allow for a slightly larger detection group range while biasing detection error towards cars.

Low and High Short Time Energy Ratio (LSTER & HSTER)

For a frame size of 5ms and a window size of 125ms, using a low pass cut-off of 0.24 for the LSTER and of 0.16 for the HSTER will identify the compression braking group while excluding five of the additional vehicles (Car4, Trucks7, 8, 10 and Mot1).

Using slightly looser limits

LSTER lowpass of 0.28 excludes Car4, Truck8, Truck10 and Mot1

(100% group selectivity, 36% exclusion)

HSTER lowpass of 0.19 excludes Car4, Truck8, Truck10 and Mot1

(100% group selectivity, 36% exclusion)

Narrowband Spectrum Flux

From the 1000ms narrowband spectrum flux data

83% of the compression braking group and 18% of the non-compression braking group lies within the limits of 14 – 20. Use range therefore as an incremental probability factor.

Single Band Spectrum Flux

Exclude vehicle if classed as accelerating (shifting up in frequency)

For a shift up in frequency, the consideration criterion is that the upper band has to be decreasing and the lower band increasing. This selection criterion detects 3 out of the four accelerating test recordings.

Third Octave Spectra

A selection criterion of 125ms third octave spectra having at least two of the three highest frequency bands between 50Hz and 400Hz as 160Hz, 200Hz and 250Hz selects all trucks using compression braking except for Truck6. Car3 and Truck7 are also falsely detected.

As with the narrowband spectrum flux 83% of the compression braking group and 18% of the non-compression braking group are selected. The criterion can therefore be as an incremental probability factor.

Detection system

The detection features and limits from above are arranged into a weighting system

Table B.1: Recognition weighting system

Test	Condition	Applied probability weighting
1	If >3dB Leq modulation rate is greater than 25 in 1000ms	-2
2	Averaged normalized Leq level lies between 0.40 – 0.55	+2
3	LSTER > 0.28	-2
4	HSTER > 0.16	-2
5	Narrowband Spectrum Flux lies within the limits of 14 – 20	+2
6	Accelerating	-10
7	Two of the three highest frequency bands between 50Hz and 400Hz are 160Hz, 200Hz and 250Hz	+2

Table B.2: Recognition Results

Vehicle	1	2	3	4	5	6	7	Score
Car1	-2	0	0	0	0	-10	0	-12
Car2	0	0	0	0	0	-10	0	-10
Car3	-2	0	0	0	0	0	+2	0
Car4	-2	0	-2	-2	0	0	0	-6
Car5	0	0	0	0	0	0	0	0
Car6	0	+2	0	0	0	0	0	+2
Truck1	0	+2	0	0	+2	0	+2	**+6**
Truck2	0	+2	0	0	+2	0	+2	**+6**
Truck3	0	+2	0	0	+2	0	+2	**+6**
Truck4	0	+2	0	0	+2	0	+2	**+6**
Truck5	0	+2	0	0	0	0	+2	**+4**
Truck6	0	+2	0	0	+2	0	0	**+4**
Truck7	0	+2	0	-2	0	0	+2	+2
Truck8	-2	0	-2	-2	0	-10	0	-16
Truck9	0	0	0	0	+2	0	0	+2
Truck10	-2	0	-2	-2	0	0	0	-6
Mot1	-2	0	-2	-2	+2	0	0	-4

Using a selection limit of greater than or equal to +4 provides 100% selection success and 100% exclusion success. These results prevail in favour of the recognition process due to the selection thresholds being based on the tested waveforms. In order to test the wider performance of the recognition algorithm, a larger test group of vehicle waveforms must be analysed. This task lies beyond the time available for this project and is recommended for further attention in Chapter 10.

APPENDIX C – VEHICLE DETECTION PROCESS

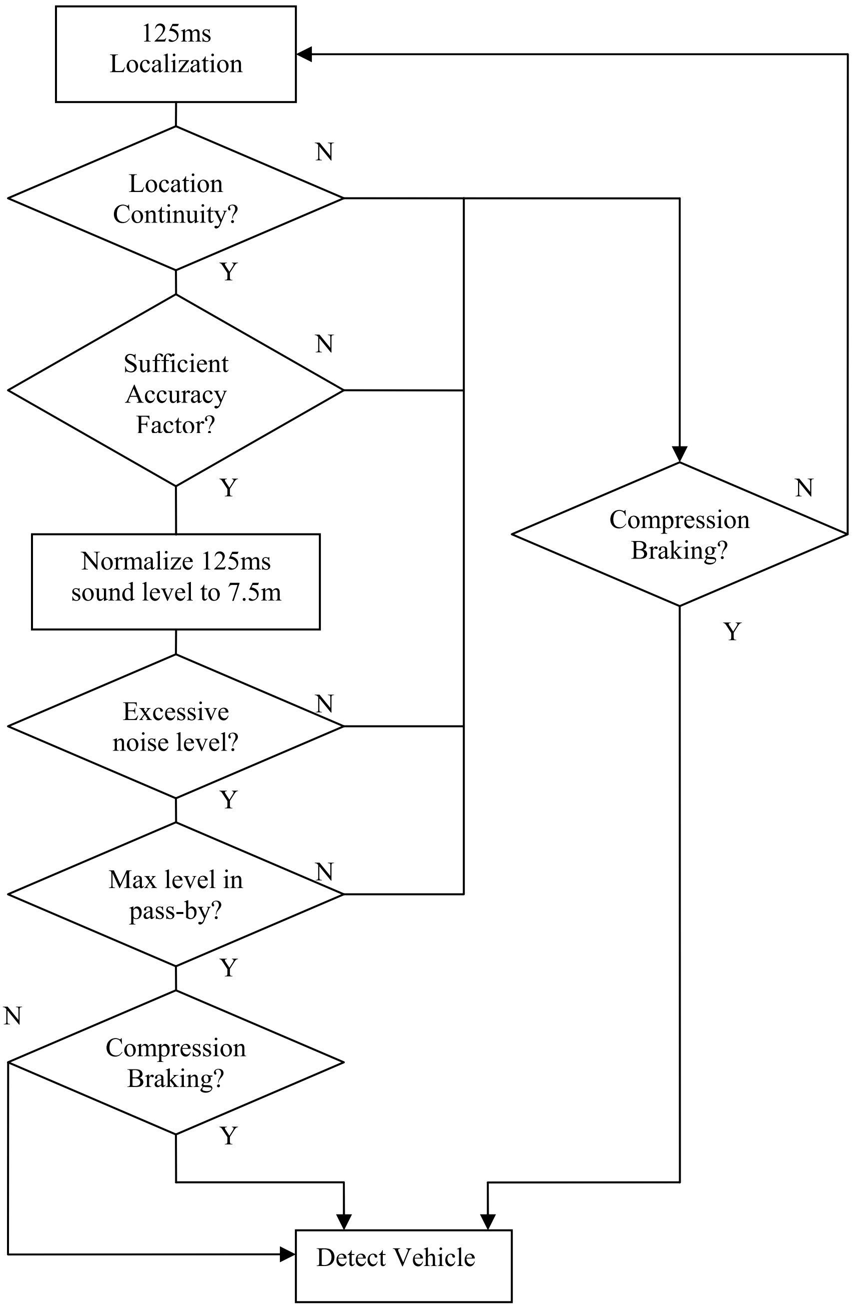


Figure C.1: Vehicle detection 125ms frame processing

Detection Process:

1. Localization algorithm is used to locate the dominant sound source on the roadway. Only locations within the predefined detection area are passed on to the system.
2. Location continuity requires that the location is part of a series (4-5)* of consecutively derived locations that reflect possible speed and position for a vehicle.
3. The location accuracy factor is determined by the magnitudes of the TDOA cross-correlation peaks. The peak values must exceed a pre-defined threshold to guarantee location accuracy.
4. With assurance of the location accuracy, the 125ms Leq may be normalized to the equivalent 7.5m sound level.
5. The derived sound level is now compared to the pre-defined detection threshold. For detection, as well as exceeding the threshold, the sound level must be 10dB(A) above the ambient level. The ambient level is the minimum Leq recorded over the last minute*.
6. If the excessive noise level is greater than all of the other levels in the surrounding 1.5* seconds of Leq's, the vehicle is processed for detection.
7. If the vehicle is also using compression braking, this is recorded in the detection information.
8. Detection consists of an image capture of the roadway and the provision of the available derived vehicle information (Location, lane number, direction of travel, peak sound level, operation type)
9. Where signal frames do not pass the previous detection steps, they are tested for the presence of compression braking before processing the following frame. If the operation is identified and there has not been detection of the operation type for the past 2-3* seconds, the image capture is triggered.

*Note: the processing time values mentioned in the detection steps are only suggested times. Implementation and analysis of the system would assist in providing the most suitable time periods. The speed limit on the section of roadway observed will also have an effect on the time periods used.

Additional considerations in the detection process will be:

1. The recorded sound level will be derived from the sound levels at each of the microphones. The closer the microphones in the array are the closer the individual levels will be. If one or more of the meters exhibit spurious measurements, the recording may be disregarded.
2. The specific carriageway dimensions will be passed to the system upon initial configuration. Other information such as the immediate speed limits, the road gradient and the position of the device with respect to the road surface will also be useful.

APPENDIX D – FIELD RECORDED SOUND LEVELS

Table D.1: Field recorded sound pressure levels

Vehicle type	Leq (dB(A))	Lane	Adjusted Leq (dB(A))	Noise level offence (Y/N)
Truck accelerating uphill	81.8	1	82.4	N
Falcon uphill	72.9	2	76.4	N
Compression braking downhill	82.6	4	90.0	Y
Truck accelerating uphill	81.9	1	82.5	N
Mazda sedan uphill	70.5	1	71.1	N
Truck accelerating uphill	81.4	1	82.0	N
Toyota Camry uphill	72.8	2	76.3	N
Quiet cars uphill	67.7-69	1	69.6	N
Cement truck uphill	82.2	1	82.8	N
Many cars downhill	75.8	3-4	81.5	N
Cement truck downhill	80.2	3	85.9	N
Many passing cars	75.7	1-4	75.7	N
Truck uphill	76.9	1	77.5	N
Commodore sedan downhill	72.2	4	79.6	Y
Ambient on windy day	61.9-63			

Lane 1 (8m) – add 0.6dB

Lane 2 (11.2m) – add 3.5dB

Lane 3 (14.4m) – add 5.7dB

Lane 4 (17.6m) – add 7.4dB

APPENDIX E – FEATURE EXTRACTION DATA

Zero Crossing Rate and ZCRR

Table E.1: Zero Crossing Rate data for the test group

	ZCR		
	20ms	125ms	1000ms
Car1	1670.45	1425.15	1622.20
Car2	5609.21	5921.02	6186.46
Car3	1056.86	1269.81	1591.66
Car4	816.67	1117.67	1115.71
Car5	431.88	566.41	527.98
Car6	3803.3	3875.26	3657.50
Truck1	**1958.94**	**1985.77**	**2390.47**
Truck2	**2343.71**	**3107.77**	**4632.79**
Truck3	**5787.53**	**6407.28**	**5605.71**
Truck4	**7285.24**	**6663.68**	**6102.73**
Truck5	**5892.88**	**4297.58**	**4852.78**
Truck6	**2047.84**	**1938.81**	**2170.01**
Truck7	4910.70	4834.75	5134.95
Truck8	4719.89	3603.91	4177.27
Truck9	8000.86	9112.89	8494.37
Truck10	1952.59	1376.38	1719.7
Mot1	860.95	1646.06	1187.92

Table E.2: High and Low Zero Crossing Rate Ratios for the test group

	HZCRR		**LZCRR**	
Window size	20ms	125ms	20ms	125ms
Car1	0.12	0.13	0.18	0.13
Car2	0.00	0.00	0.00	0.00
Car3	0.10	0.00	0.18	0.00
Car4	0.08	0.00	0.14	0.00

Car5	0.10	0.00	0.12	0.00
Car6	0.00	0.00	0.06	0.00
Truck1	**0.06**	**0.00**	**0.08**	**0.00**
Truck2	**0.04**	**0.00**	**0.20**	**0.13**
Truck3	**0.06**	**0.00**	**0.16**	**0.13**
Truck4	**0.00**	**0.00**	**0.06**	**0.00**
Truck5	**0.02**	**0.00**	**0.02**	**0.00**
Truck6	**0.02**	**0.00**	**0.04**	**0.00**
Truck7	0.06	0.00	0.08	0.00
Truck8	0.00	0.00	0.02	0.00
Truck9	0.02	0.00	0.08	0.13
Truck10	0.10	0.00	0.14	0.00
Mot1	0.12	0.13	0.20	0.13

Filtered signals

Table E.3: Zero Crossing Rate data for the test group after lowpass filtering

	ZCR		
	20ms	125ms	1000ms
Car1	429.55	437.89	538.07
Car2	1547.37	1989.65	1994.50
Car3	768.63	790.64	921.23
Car4	36.27	710.52	712.18
Car5	431.88	462.70	418.19
Car6	2503.49	2578.20	2425.06
Truck1	**764.46**	**921.40**	**1199.71**
Truck2	**813.12**	**687.07**	**1001.52**
Truck3	**1530.59**	**1837.50**	**1376.44**
Truck4	**1797.03**	**1755.70**	**1792.28**
Truck5	**1835.49**	**1682.35**	**1510.51**
Truck6	**1285.85**	**1144.22**	**1128.80**

Truck7	1251.75	1222.77	1092.71
Truck8	1764.00	1403.29	1395.08
Truck9	2190.71	1894.21	1871.15
Truck10	523.87	477.3	539.72
Mot1	382.65	559.34	530.96

Table E.4: ZCR's of the 1 second waveforms using a lowpass filter

	ZCR				
Cut-off Freq	400Hz	800Hz	1200Hz	2000Hz	No Filter
Car1	318.45	334.42	401.31	538.07	1622.20
Car2	344.91	727.82	991.75	1994.50	6186.46
Car3	327.73	400.66	551.54	921.23	1591.66
Car4	352.59	388.55	526.39	712.18	1115.71
Car5	317.39	311.40	347.33	418.19	527.98
Car6	255.85	770.52	1804.86	2425.06	3657.50
Truck1	**429.39**	**625.20**	**889.59**	**1199.71**	**2390.47**
Truck2	**290.86**	**408.81**	**564.73**	**1001.52**	**4632.79**
Truck3	**159.93**	**369.85**	**688.72**	**1376.44**	**5605.71**
Truck4	**426.35**	**857.70**	**1160.24**	**1792.28**	**6102.73**
Truck5	**352.19**	**594.62**	**946.81**	**1510.51**	**4852.78**
Truck6	**377.26**	**608.20**	**864.02**	**1128.80**	**2170.01**
Truck7	276.67	381.55	605.29	1092.71	5134.95
Truck8	281.81	483.32	916.26	1395.08	4177.27
Truck9	163.31	358.50	770.77	1871.15	8494.37
Truck10	153.35	187.21	348.53	539.72	1719.7
Mot1	323.98	353.98	415.97	530.96	1187.92

Table E.5: HZCRR and LZCRR Using a ZCR ratio of 1.5

	HZCRR	**LZCRR**
Window size	20ms	20ms

Car1	0.08	0.20
Car2	0.00	0.04
Car3	0.06	0.16
Car4	0.02	0.06
Car5	0.10	0.00
Car6	0.00	0.06
Truck1	**0.00**	**0.02**
Truck2	**0.08**	**0.16**
Truck3	**0.22**	**0.22**
Truck4	**0.00**	**0.04**
Truck5	**0.00**	**0.06**
Truck6	**0.06**	**0.14**
Truck7	0.08	0.20
Truck8	0.00	0.06
Truck9	0.00	0.10
Truck10	0.12	0.24
Mot1	0.08	0.18

Table E.6: HZCRR and LZCRR Using a ZCR ratio of 1.2

	HZCRR	**LZCRR**
Window size	20ms	20ms
Car1	0.26	0.36
Car2	0.20	0.20
Car3	0.18	0.32
Car4	0.20	0.28
Car5	0.12	0.36
Car6	0.00	0.08
Truck1	**0.24**	**0.22**
Truck2	**0.24**	**0.28**
Truck3	**0.36**	**0.42**
Truck4	**0.12**	**0.16**
Truck5	**0.12**	**0.20**

Truck6	**0.16**	**0.24**
Truck7	0.32	0.32
Truck8	0.14	0.16
Truck9	0.24	0.20
Truck10	0.28	0.32
Mot1	0.18	0.22

Leq Modulation Rate

Table E.7: Number of modulations in the 125ms sound recording

Wave File	Modulations > 3dB	Modulations > 5dB	Modulations > 7dB
Car1	4	0	0
Car2	2	0	0
Car3	3	1	0
Car4	6	2	1
Car5	1	0	0
Car6	0	0	0
Truck1	**3**	**0**	**0**
Truck2	**1**	**0**	**0**
Truck3	**2**	**0**	**0**
Truck4	**1**	**0**	**0**
Truck5	**1**	**0**	**0**
Truck6	**2**	**0**	**0**
Truck7	1	0	0
Truck8	7	4	0
Truck9	0	0	0
Truck10	7	4	0
Mot1	3	2	2

Table E.8: Number of modulations in the 1000ms sound recording

Wave File	Modulations	Modulations	Modulations

	> 3dB	> 5dB	> 7dB
Car1	33	13	2
Car2	5	1	0
Car3	30	10	1
Car4	41	19	4
Car5	8	0	0
Car6	11	1	0
Truck1	**19**	**5**	**0**
Truck2	**20**	**2**	**0**
Truck3	**17**	**2**	**1**
Truck4	**8**	**1**	**0**
Truck5	**14**	**2**	**0**
Truck6	**19**	**5**	**0**
Truck7	17	4	0
Truck8	34	8	0
Truck9	8	2	1
Truck10	43	21	6
Mot1	33	25	19

Table E.9: Averaged level of normalized Leq(A) waveform over 125ms

Wave File	Average Leq
Car1	0.55
Car2	0.44
Car3	0.47
Car4	0.56
Car5	0.47
Car6	0.53
Truck1	**0.45**
Truck2	**0.54**
Truck3	**0.62**
Truck4	**0.49**

Truck5	**0.37**
Truck6	**0.57**
Truck7	0.47
Truck8	0.63
Truck9	0.52
Truck10	0.66
Mot1	0.53

Table E.10: Averaged level of normalized Leq(A) waveform over 1000ms

Wave File	Average Leq
Car1	0.58
Car2	0.58
Car3	0.66
Car4	0.58
Car5	0.59
Car6	0.54
Truck1	**0.46**
Truck2	**0.52**
Truck3	**0.45**
Truck4	**0.53**
Truck5	**0.45**
Truck6	**0.53**
Truck7	0.50
Truck8	0.56
Truck9	0.33
Truck10	0.60
Mot1	0.56

Low and High Short Time Energy Ratio (LSTER & HSTER)

Table E.11: STER's using small window of 20ms and long window of 1000ms

Wave File	LSTER	HSTER

Car1	0.02	0.02
Car2	0.08	0.04
Car3	0.18	0.08
Car4	0.14	0.06
Car5	0.20	0.14
Car6	0.06	0.06
Truck1	**0.12**	**0.10**
Truck2	**0.28**	**0.18**
Truck3	**0.20**	**0.36**
Truck4	**0.00**	**0.02**
Truck5	**0.08**	**0.00**
Truck6	**0.00**	**0.02**
Truck7	0.26	0.18
Truck8	0.00	0.00
Truck9	0.14	0.08
Truck10	0.04	0.02
Mot1	0.28	0.20

Table E.12: STER's using small window of 5ms and long window of 1000ms

Wave File	LSTER	HSTER
Car1	0.20	0.12
Car2	0.220	0.115
Car3	0.335	0.180
Car4	0.385	0.205
Car5	0.225	0.155
Car6	0.2289	0.1144
Truck1	**0.2587**	**0.1493**
Truck2	**0.3650**	**0.1750**
Truck3	**0.4250**	**0.1550**
Truck4	**0.1800**	**0.1350**
Truck5	**0.2450**	**0.1200**
Truck6	**0.1990**	**0.1244**

Truck7	0.3600	0.1850
Truck8	0.2050	0.085
Truck9	0.3284	0.0746
Truck10	0.2637	0.1343
Mot1	0.4850	0.2050

Table E.13: STER's using small window of 5ms and long window of 125ms

Wave File	LSTER	HSTER
Car1	0.16	0.16
Car2	0.12	0.12
Car3	0.200	0.08
Car4	0.32	0.24
Car5	0.04	0.08
Car6	0.08	0.04
Truck1	**0.200**	**0.12**
Truck2	**0.240**	**0.120**
Truck3	**0.08**	**0.00**
Truck4	**0.200**	**0.16**
Truck5	**0.200**	**0.08**
Truck6	**0.12**	**0.04**
Truck7	0.1667	0.1667
Truck8	0.3600	0.200
Truck9	0.240	0.160
Truck10	0.320	0.200
Mot1	0.440	0.280

Spectrum Flux

Table E.14: Spectrum Flux of 1000ms waveform

Wave File	SF
Car1	7.6487
Car2	8.0238

Car3	6.9744
Car4	9.6847
Car5	7.1942
Car6	8.5681
Truck1	**9.8848**
Truck2	**8.7554**
Truck3	**11.5716**
Truck4	**11.4794**
Truck5	**6.5138**
Truck6	**9.9481**
Truck7	7.5249
Truck8	6.9564
Truck9	9.9228
Truck10	8.8766
Mot1	14.2682

Table E.15: Spectrum Flux of 125ms waveform

Wave File	SF
Car1	11.1532
Car2	11.6483
Car3	9.5692
Car4	15.2050
Car5	5.8131
Car6	9.7776
Truck1	**8.5966**
Truck2	**10.6369**
Truck3	**8.7569**
Truck4	**9.9809**
Truck5	**15.2797**
Truck6	**14.1541**
Truck7	11.3818
Truck8	9.7060

Truck9	9.7736
Truck10	8.7083
Mot1	18.2288

Table E.16: Spectrum Flux of 1000ms narrowband waveform

Wave File	SF
Car1	11.5268
Car2	10.518
Car3	11.0771
Car4	13.6865
Car5	10.9686
Car6	12.2338
Truck1	**16.7432**
Truck2	**14.0461**
Truck3	**19.6612**
Truck4	**18.5436**
Truck5	**9.6776**
Truck6	**16.5814**
Truck7	12.2632
Truck8	11.0012
Truck9	17.9045
Truck10	13.8063
Mot1	19.2461

Table E.17: Spectrum Flux of 125ms narrowband waveform

Wave File	SF
Car1	17.2416
Car2	20.0555
Car3	17.8414
Car4	27.3225
Car5	9.8730
Car6	17.1399

Truck1	**12.0733**
Truck2	**18.2898**
Truck3	**12.8306**
Truck4	**14.3056**
Truck5	**31.7991**
Truck6	**28.3914**
Truck7	20.6131
Truck8	17.5686
Truck9	16.5232
Truck10	14.5687
Mot1	28.3438

Single Band Spectrum Flux

Table E.18: Single band flux trends and vehicle behaviour

Wave File	Slope			Estimated shift in Freq.	Vehicle Behaviour
	Lower Band	Peak Band	Upper Band		
Car1	-0.2222	-0.0263	0.1829	Up	Accelerating
Car2	-0.0734	0.0644	0.1836	Up	Accelerating
Car3	0.0888	0.2063	-0.1214	Down	Pass
Car4	0.0240	0.1702	0.1082	None	Pass
Car5	0.2385	0.1736	0.0082	Down	Pass
Car6	0.1346	0.1957	0.0365	None	Pass
Truck1	**0.1826**	**0.0696**	**0.0448**	**Down**	**Pass + Break**
Truck2	**0.2381**	**0.0637**	**-0.0304**	**Down**	**Pass + Break**
Truck3	**0.0484**	**0.0282**	**0.0329**	**None**	**Pass + Break**
Truck4	**0.0557**	**-0.1198**	**0.0120**	**None**	**Pass + Break**
Truck5	**0.1427**	**-0.0325**	**0.1285**	**None**	**Pass + Break**
Truck6	**0.0699**	**0.0497**	**0.1215**	**None**	**Pass + Break**
Truck7	0.1976	0.2153	0.0512	None	Accelerating
Truck8	-0.0867	0.1756	0.1041	Up	Accelerating

Truck9	0.0777	0.1139	0.1106	None	Pass
Truck10	0.0661	0.0855	0.1181	None	Pass
Mot1	0.0170	0.0472	-0.0386	None	Pass

Noise Frame Ratio

Table E.19: Noise frame ratio using 20ms frames in a 1000ms waveform

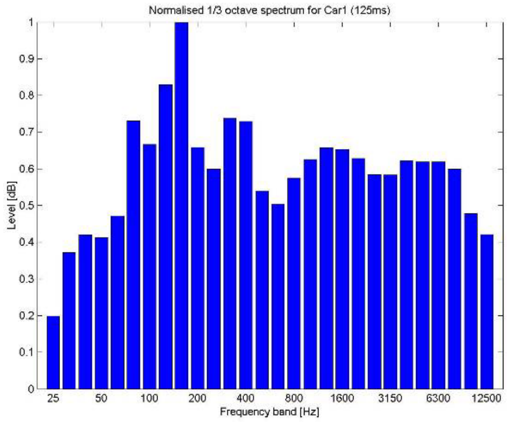
Wave File	NFR			
	0.2	0.3	0.4	0.5
Car1	0	0.08	0.36	0.68
Car2	0.66	0.96	1	1
Car3	0.06	0.46	1	1
Car4	0.1	0.44	0.94	1
Car5	0	0	0	0.06
Car6	0.24	0.76	0.9	0.94
Truck1	**0.16**	**0.52**	**0.76**	**0.92**
Truck2	**0**	**0.24**	**0.70**	**0.84**
Truck3	**0.04**	**0.20**	**0.40**	**0.62**
Truck4	**0.62**	**0.88**	**0.94**	**0.98**
Truck5	**0.28**	**0.86**	**1**	**1**
Truck6	**0.40**	**0.80**	**0.96**	**1**
Truck7	0.22	0.44	0.62	0.86
Truck8	0.40	0.94	1	1
Truck9	0.04	0.44	0.76	0.94
Truck10	0	0.02	0.14	0.60
Mot1	0.16	0.40	0.86	0.98

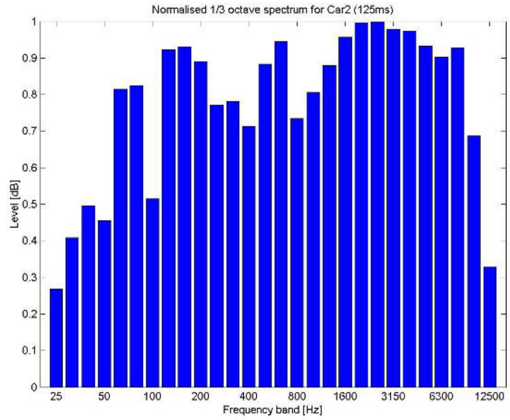
Third Octave Spectra

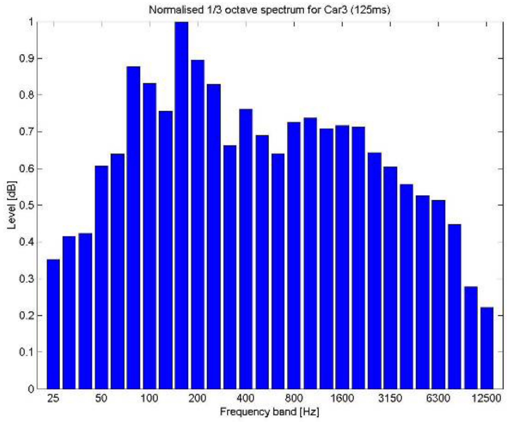
Table E.20: Peak frequency bands between 50Hz and 400Hz

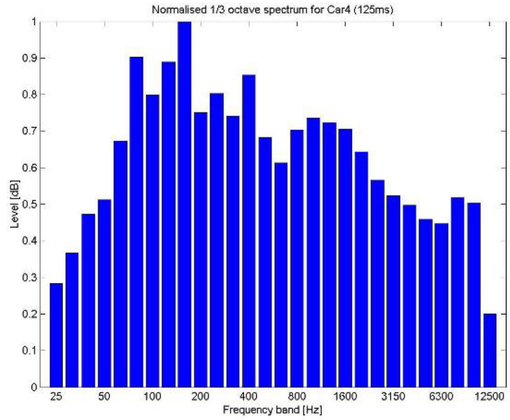
Wave File	3 highest frequency bands			
	1st	2nd	3rd	
Car1	160	125	315	No
Car2	160	125	200	No
Car3	160	200	80	Yes
Car4	160	80	125	No
Car5	160	315	200	No
Car6	160	315	250	No
Truck1	**200**	**160**	**400**	**Yes 2**
Truck2	**160**	**200**	**250**	**Yes 3**
Truck3	**250**	**200**	**160**	**Yes 3**
Truck4	**250**	**160**	**315**	**Yes 2**
Truck5	**200**	**160**	**125**	**Yes 2**
Truck6	**125**	**315**	**63**	**No**
Truck7	160	250	200	Yes 3
Truck8	125	63	200	No
Truck9	80	160	100	No
Truck10	63	80	125	No
Mot1	250	125	160	No

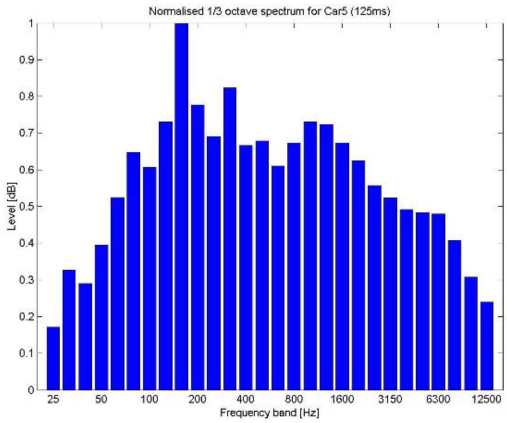
APPENDIX F – THIRD OCTAVE SPECTRA

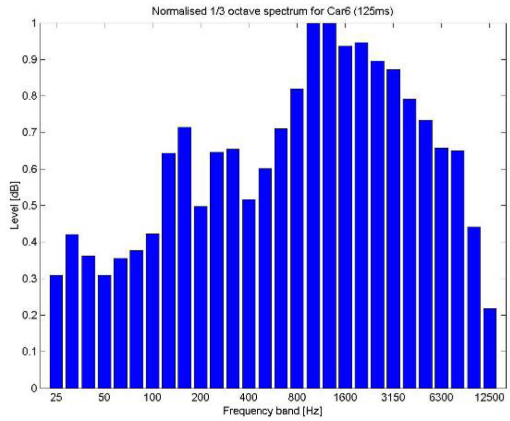


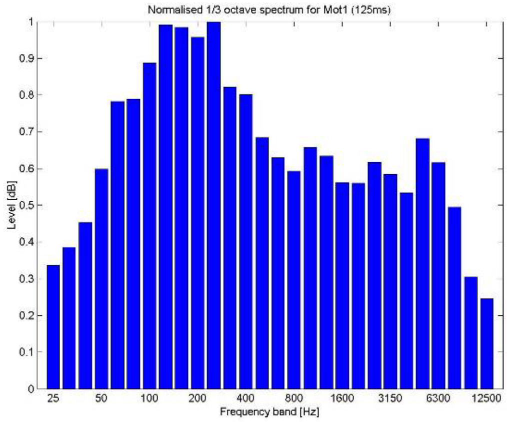


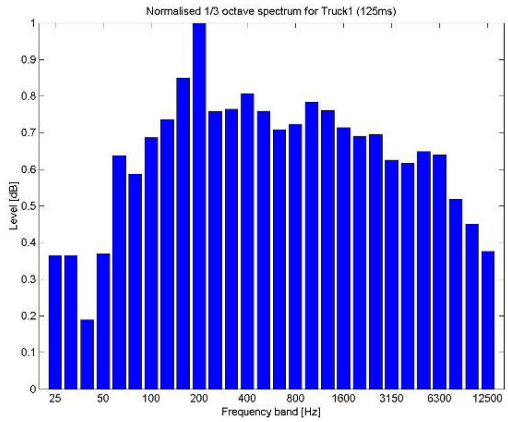


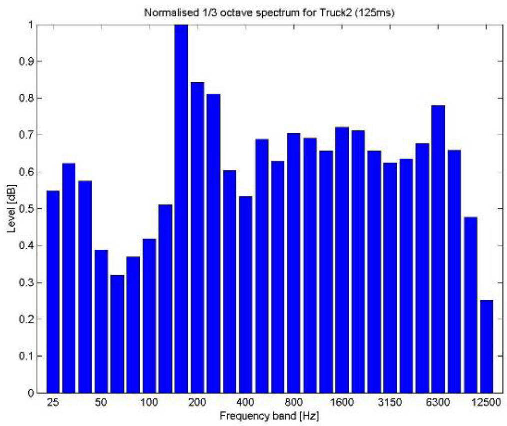


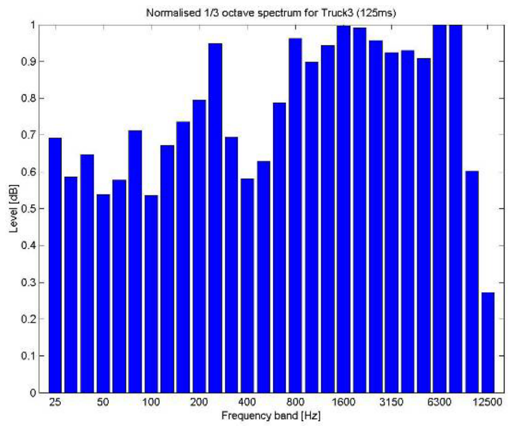


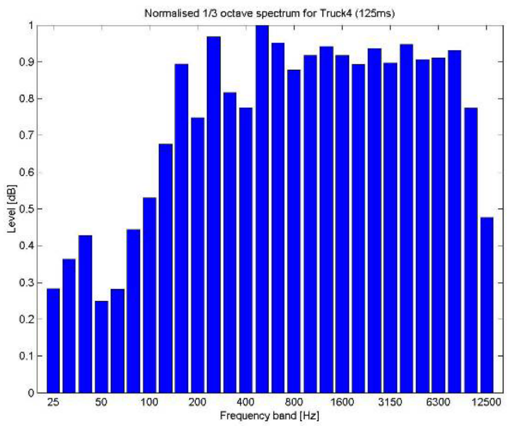


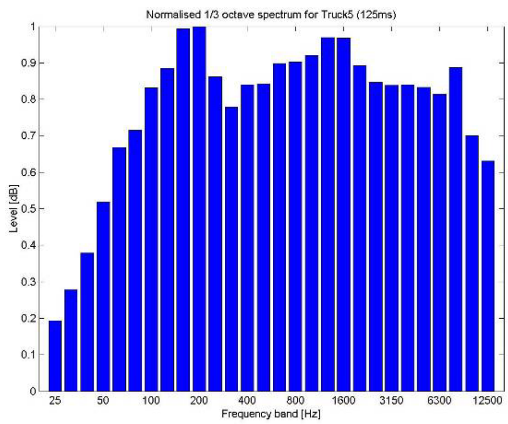

Normalised 1/3 octave spectrum for Mot1 (125ms)
Level [dB]
Frequency band [Hz]

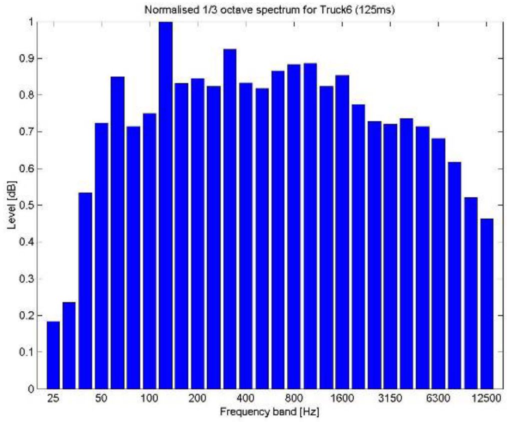

Normalised 1/3 octave spectrum for Truck1 (125ms)
Level [dB]
Frequency band [Hz]

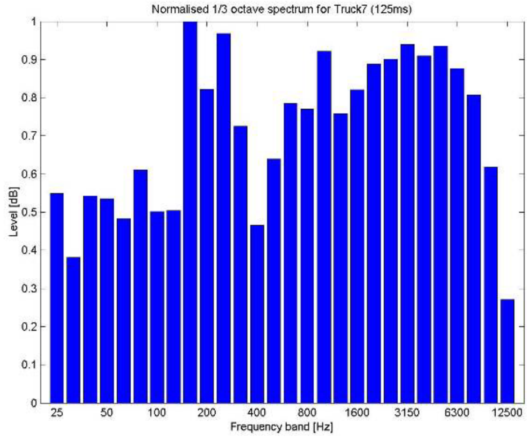

Normalised 1/3 octave spectrum for Truck2 (125ms)
Level [dB]
Frequency band [Hz]

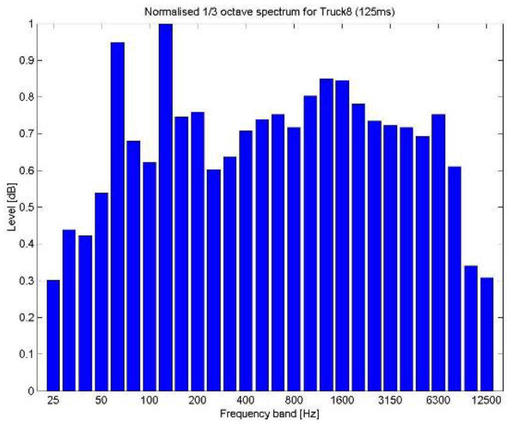

Normalised 1/3 octave spectrum for Truck3 (125ms)
Level [dB]
Frequency band [Hz]

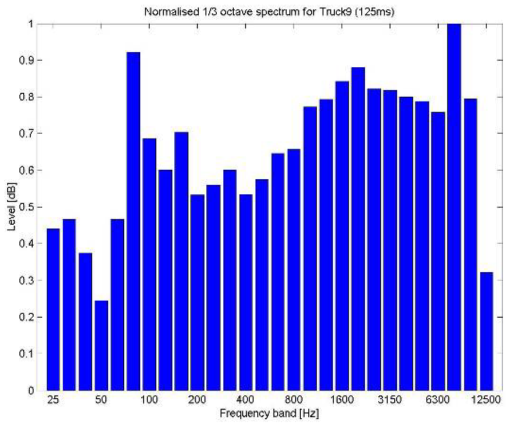

Normalised 1/3 octave spectrum for Truck4 (125ms)
Level [dB]
Frequency band [Hz]

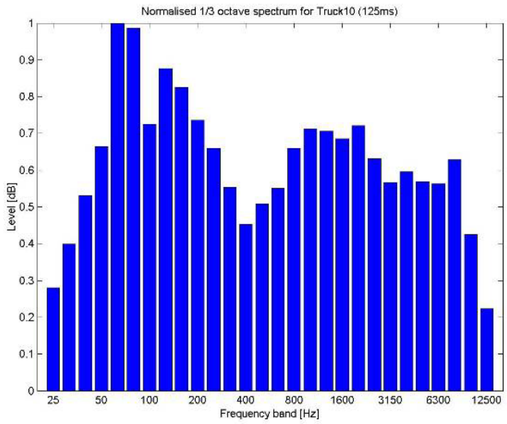

Normalised 1/3 octave spectrum for Truck5 (125ms)
Level [dB]
Frequency band [Hz]


Normalised 1/3 octave spectrum for Truck6 (125ms)
Level [dB]
Frequency band [Hz]


Normalised 1/3 octave spectrum for Truck7 (125ms)
Level [dB]
Frequency band [Hz]


Normalised 1/3 octave spectrum for Truck8 (125ms)
Level [dB]
Frequency band [Hz]


Normalised 1/3 octave spectrum for Truck9 (125ms)
Level [dB]
Frequency band [Hz]


Normalised 1/3 octave spectrum for Truck10 (125ms)
Level [dB]
Frequency band [Hz]

APPENDIX G – TDOA ESTIMATION

Single Vehicle Pass-by Simulations

Group A = -20m to 20m, offset = 7.5m

Group B = -20m to 20m, offset = 11m

Group C = -20m to 20m, offset = 14.5m

Group D = -20m to 20m, offset = 18m

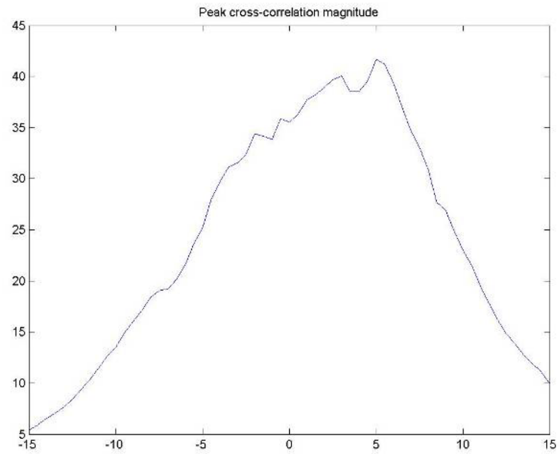
Time Domain Cross-Correlation

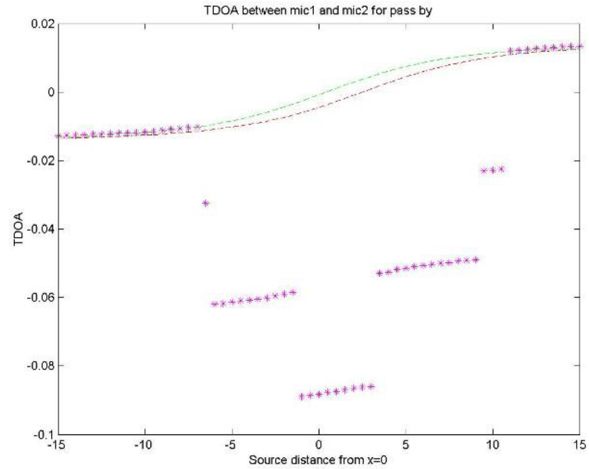
For complete delay length correlation

Correlation peak magnitude and estimated TDOA between the correlated microphone pair during pass-by.

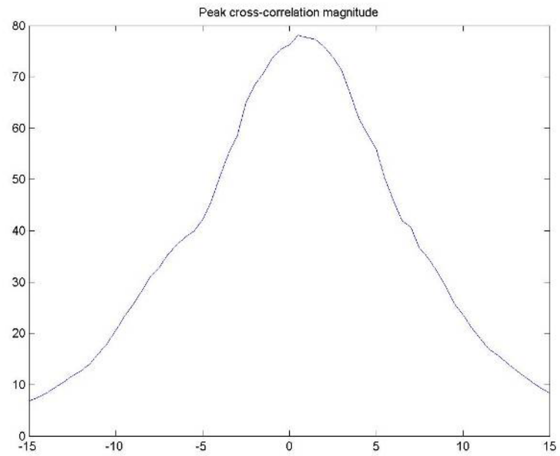
Group A

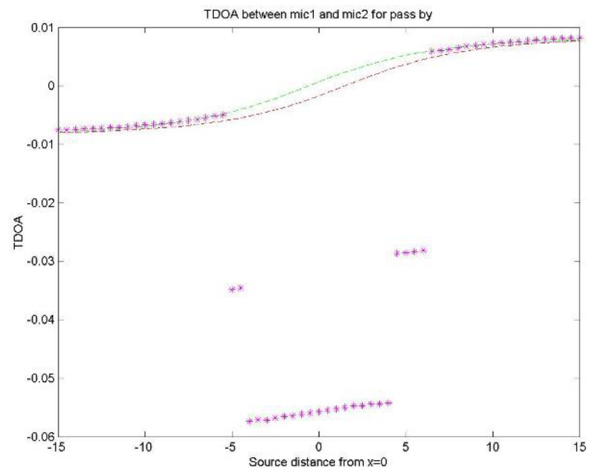
Aperture = 5m



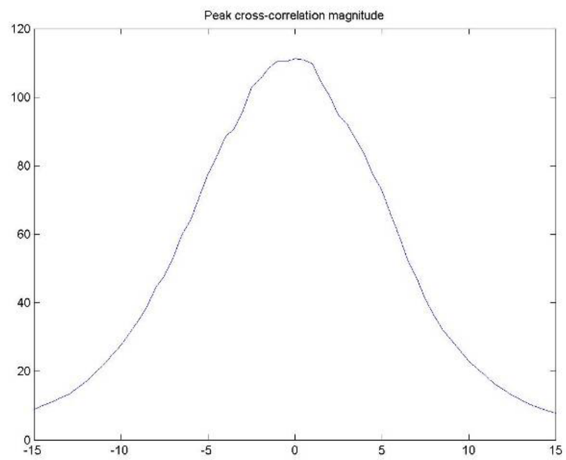


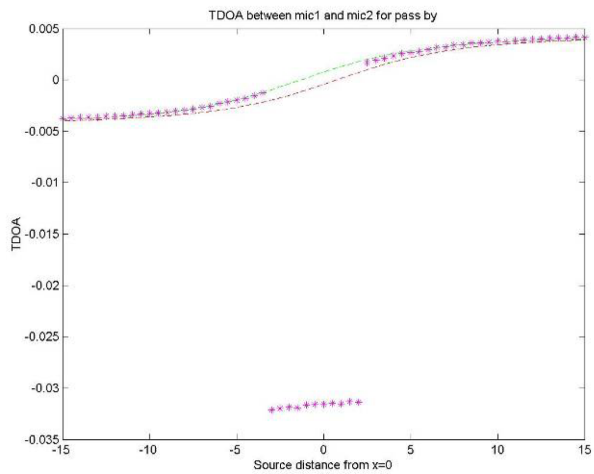
Aperture = 3m



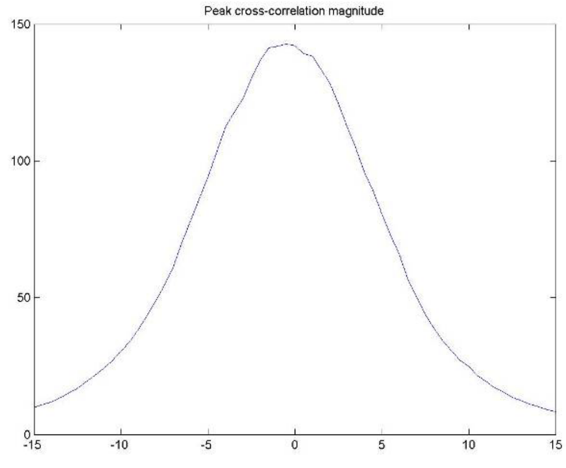


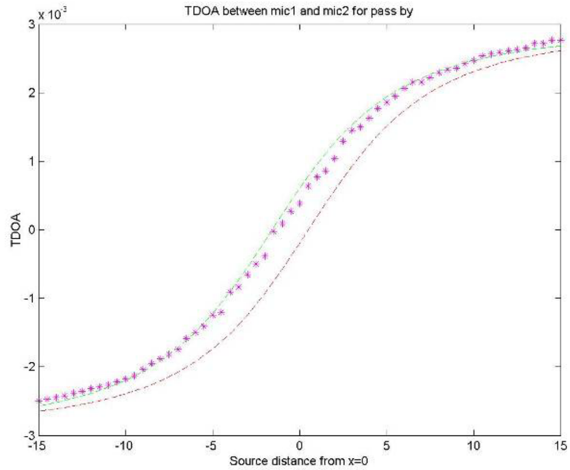
Aperture = 1.5m



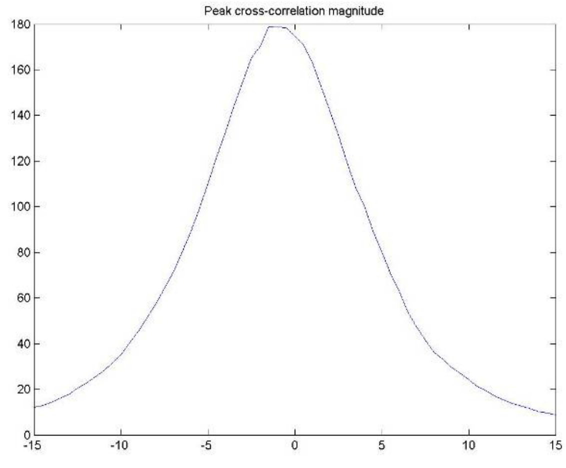


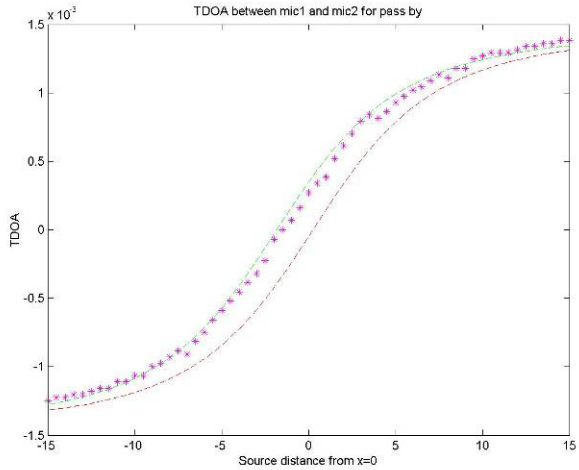
Aperture = 1m





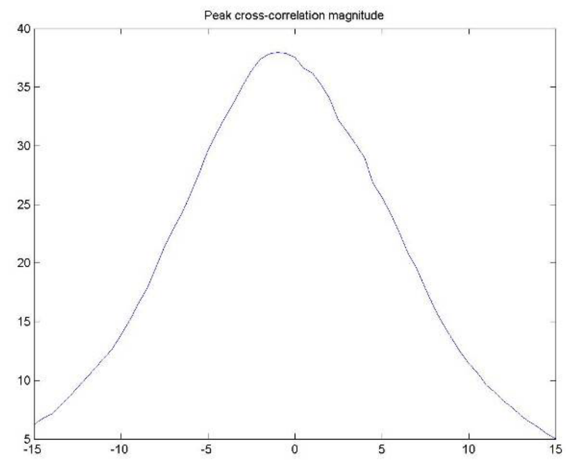
Aperture = 0.5m

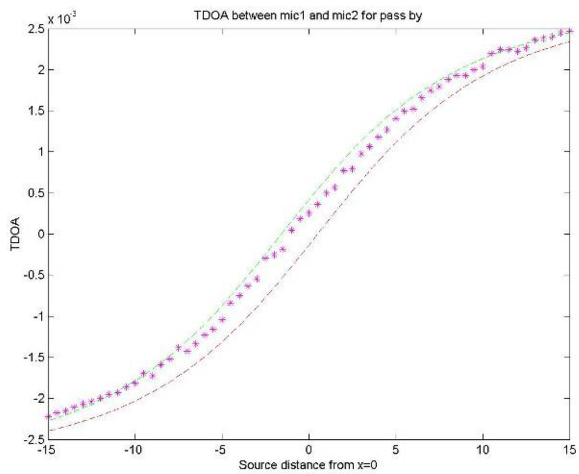




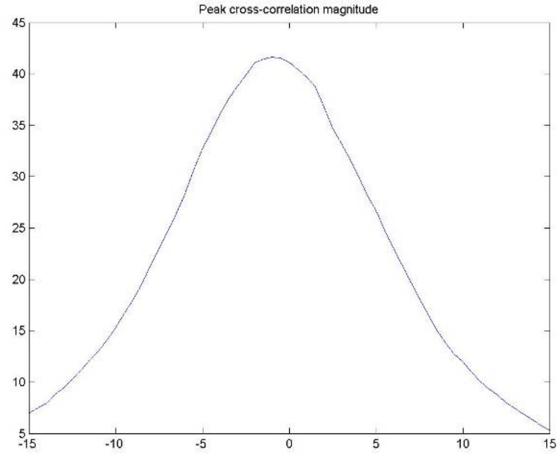
Group B

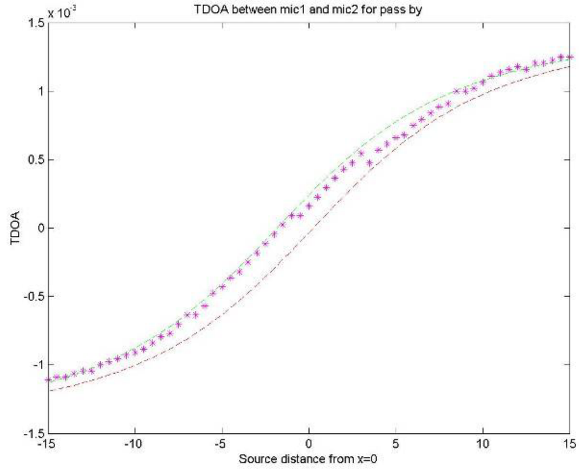
Aperture = 1m





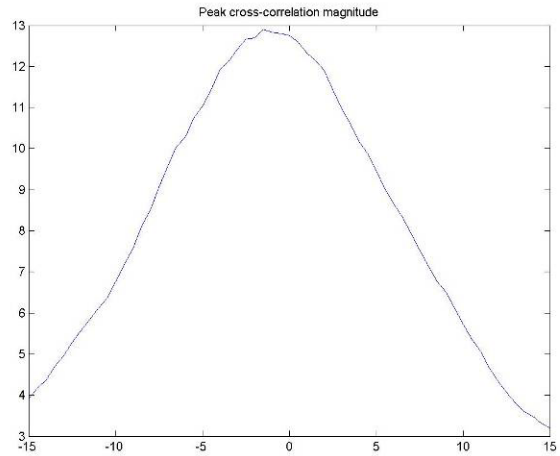
Aperture = 0.5m

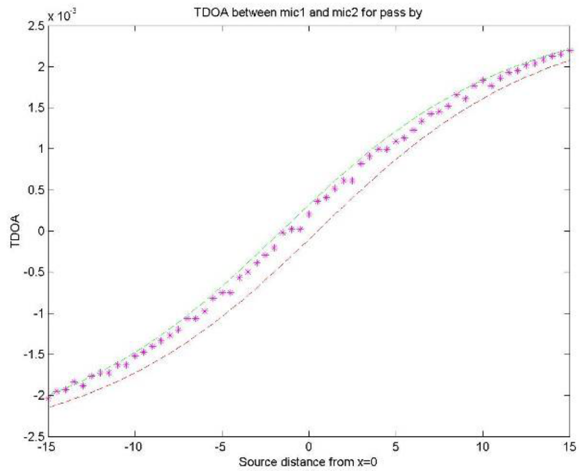




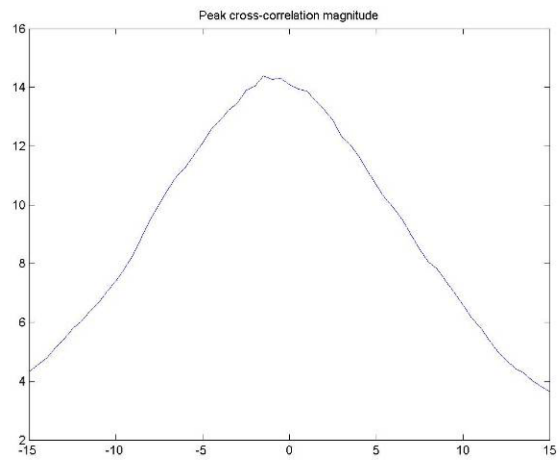
Group C

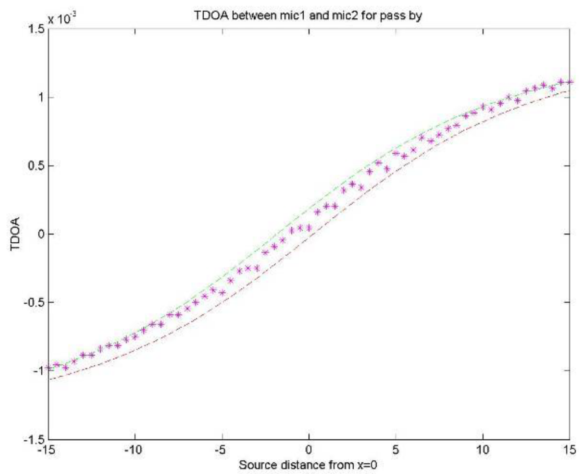
Aperture = 1m





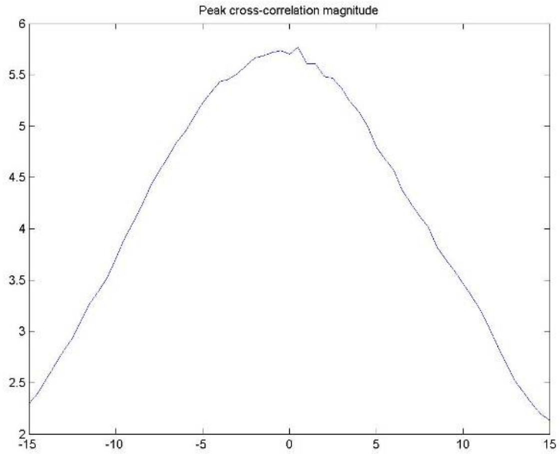
Aperture = 0.5m

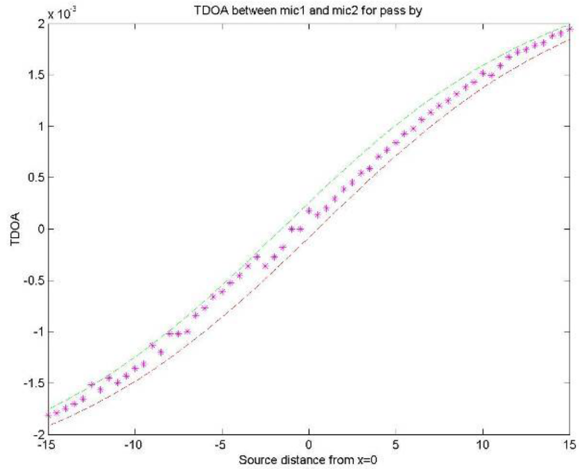




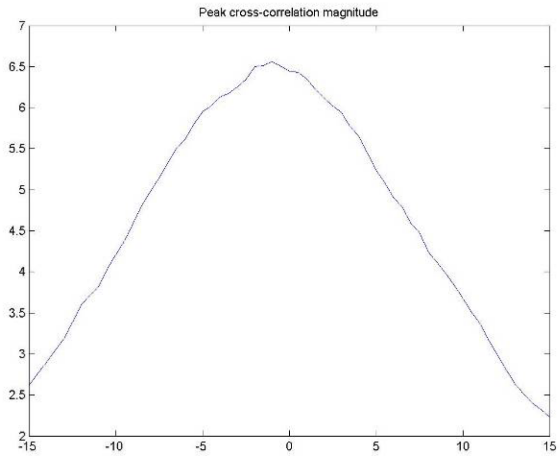
Group D

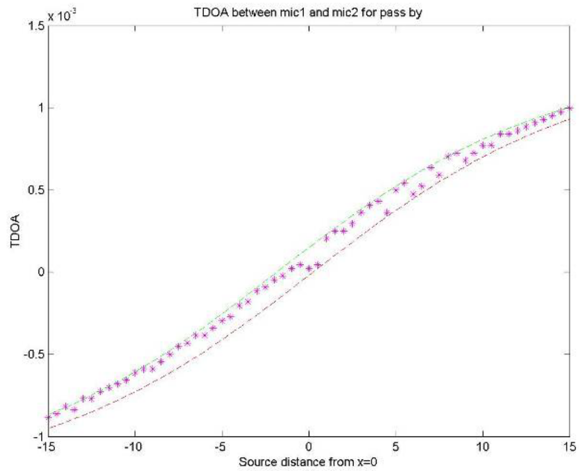
Aperture = 1m





Aperture = 0.5m



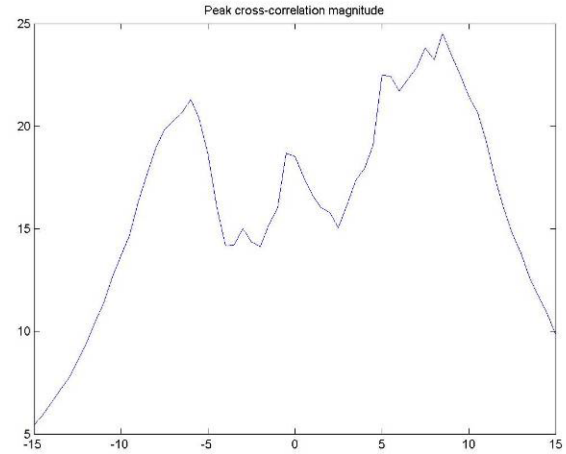


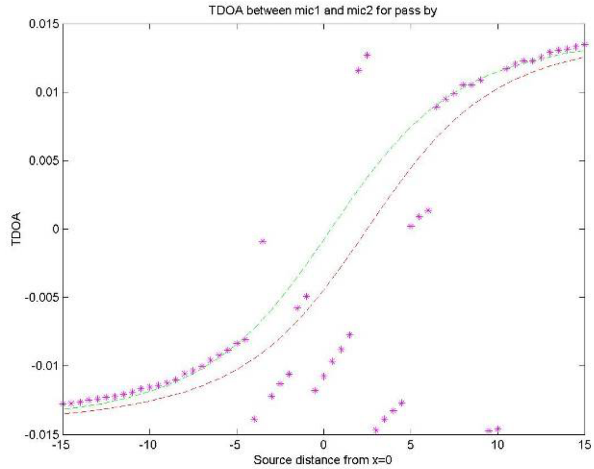
Time Domain Cross-Correlation

For Restricted delay length correlation

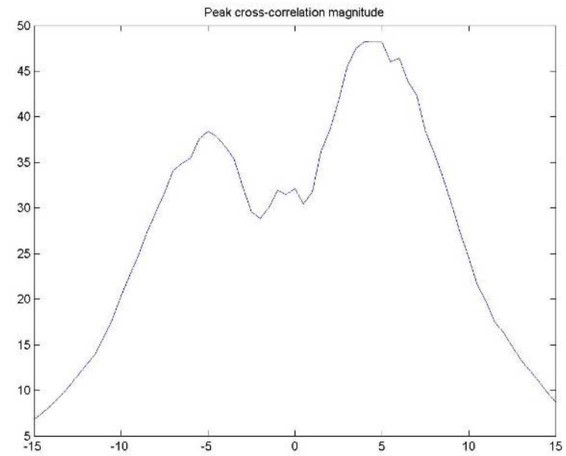
Group A

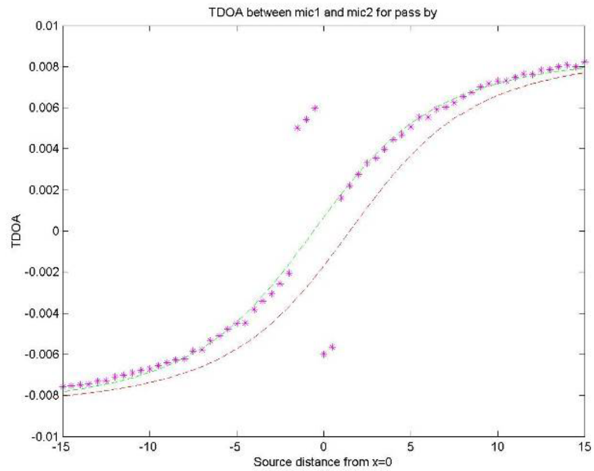
Aperture = 5m



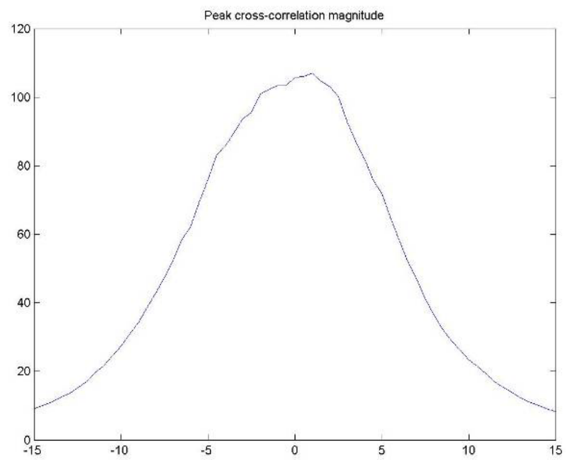


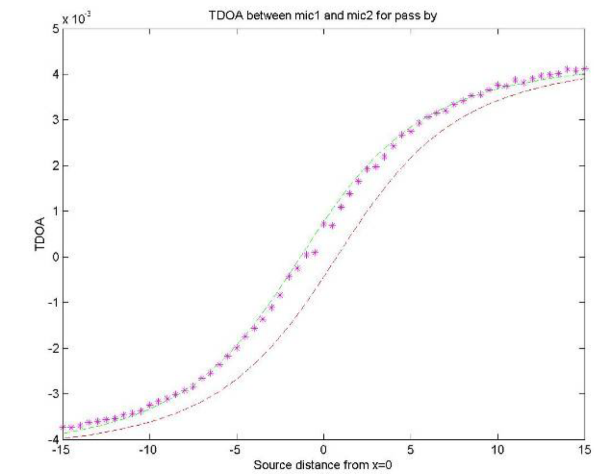
Aperture = 3m





Aperture = 1.5m

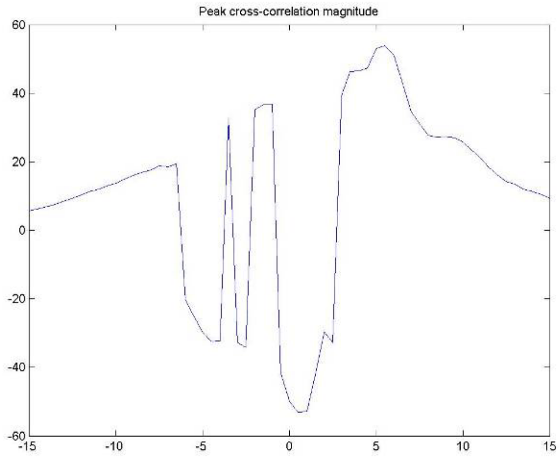


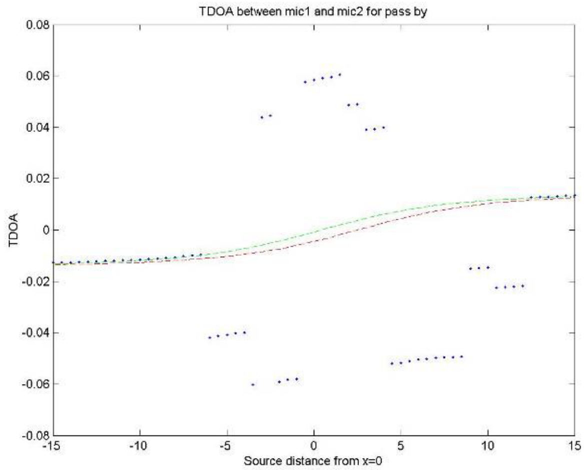


Cross-Spectrum Correlation

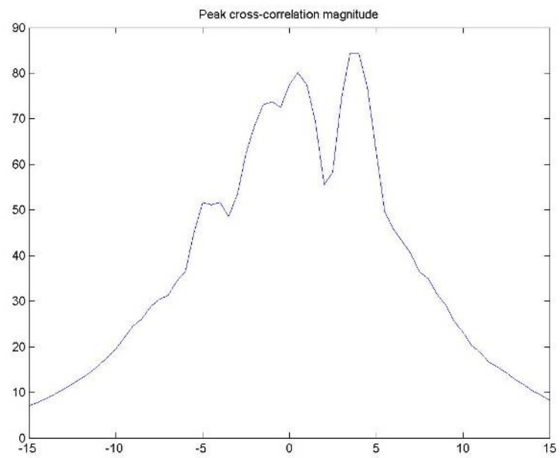
Group A

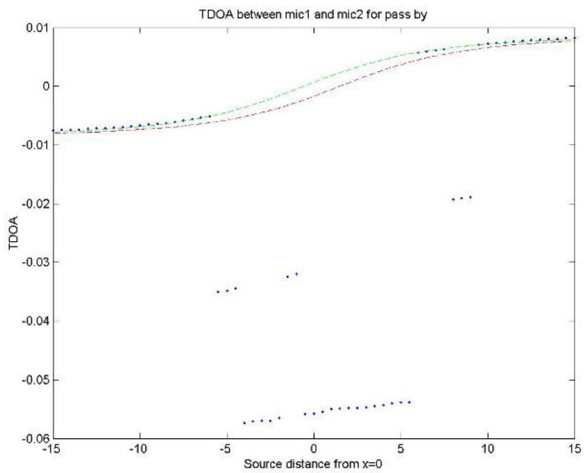
Aperture = 5m



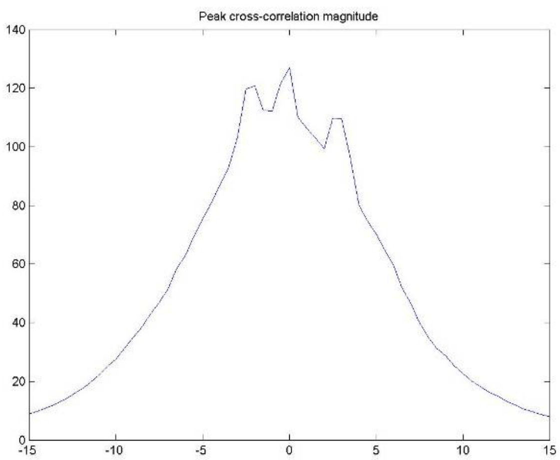


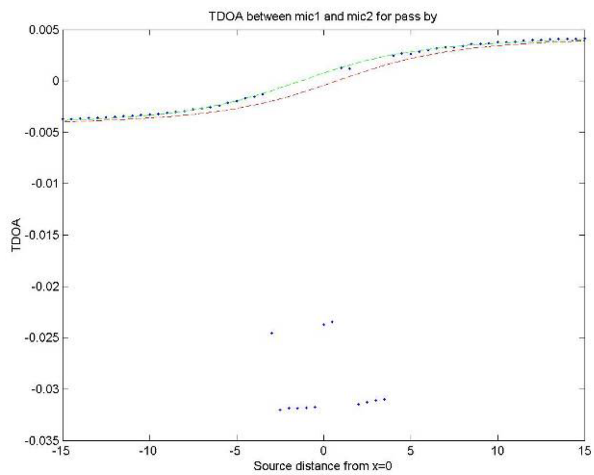
Aperture = 3m



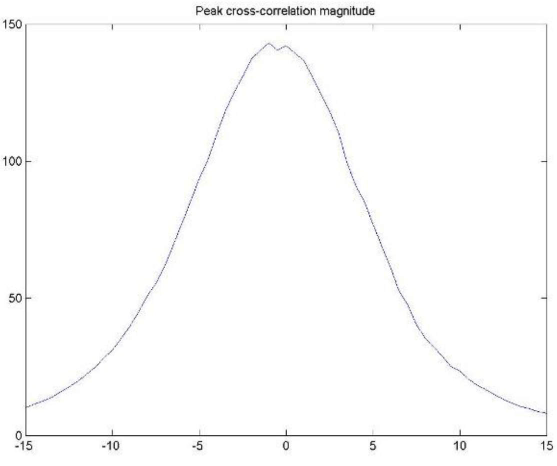


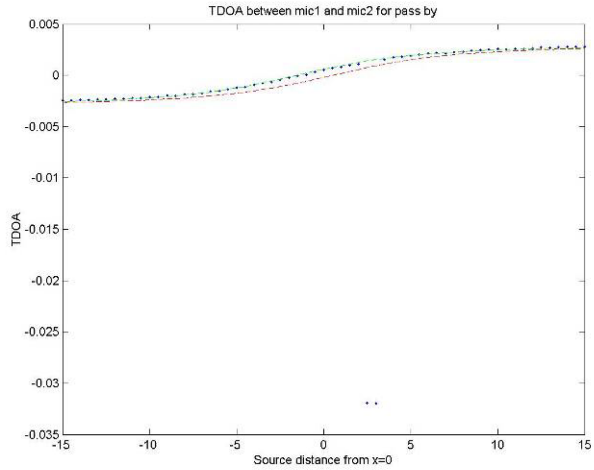
Aperture = 1.5m



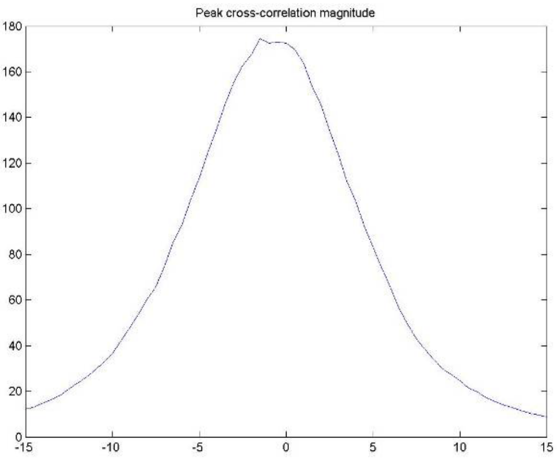


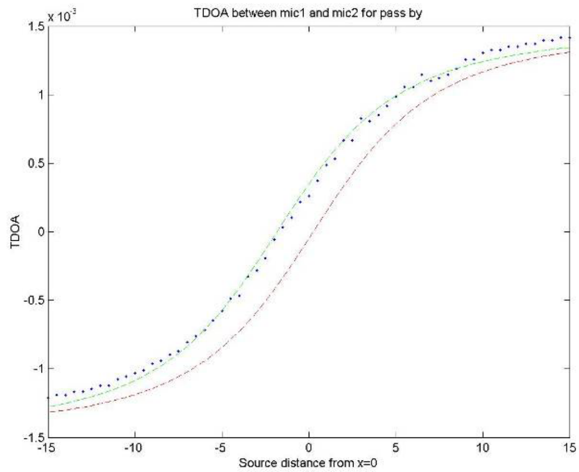
Aperture = 1m





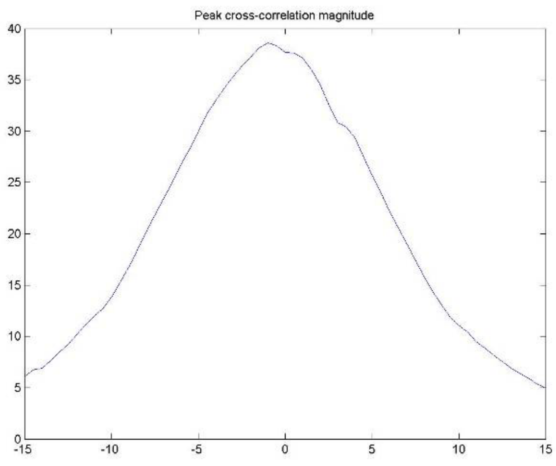
Aperture = 0.5m

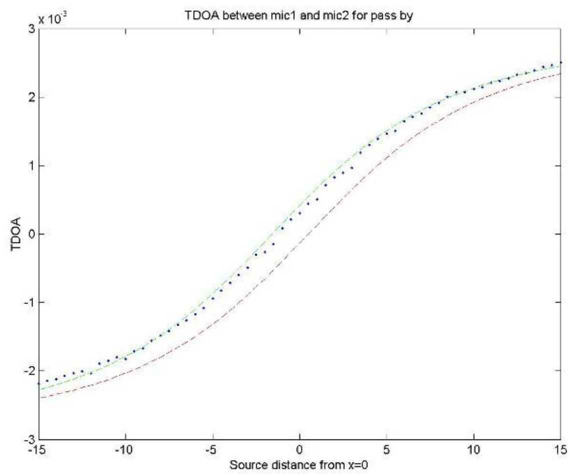




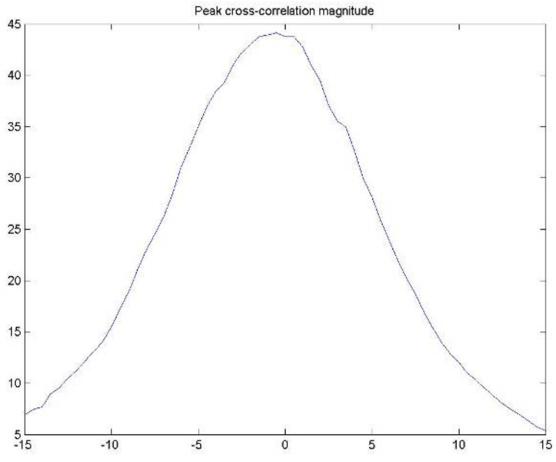
Group B

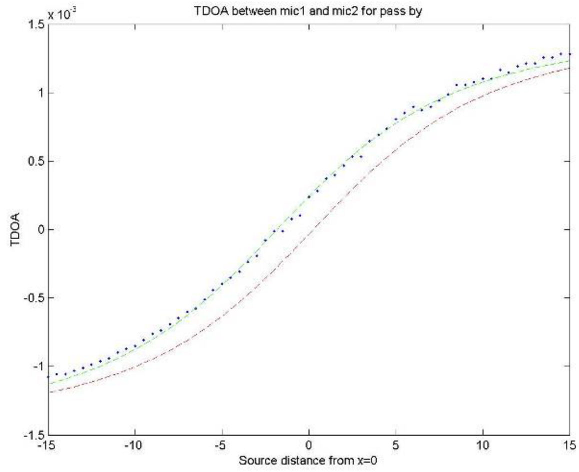
Aperture = 1m





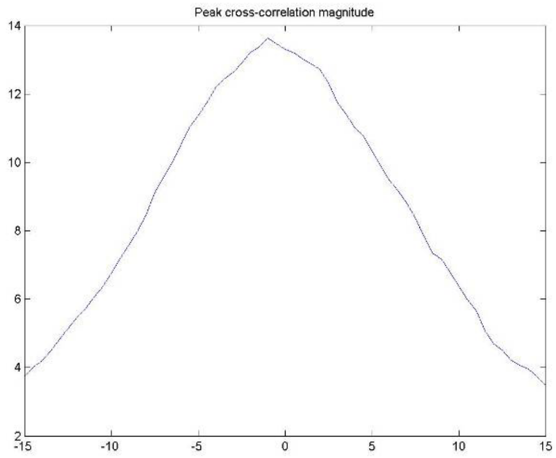
Aperture = 0.5m

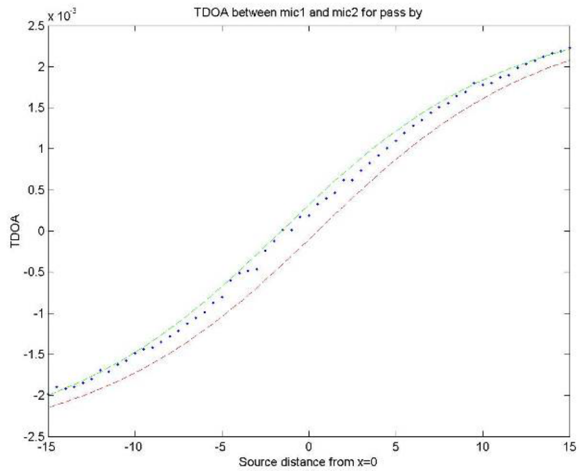




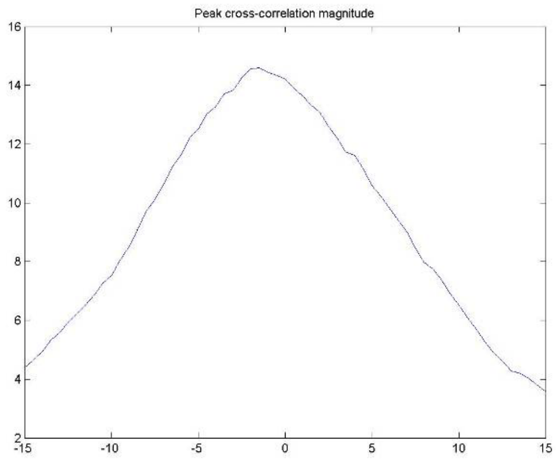
Group C

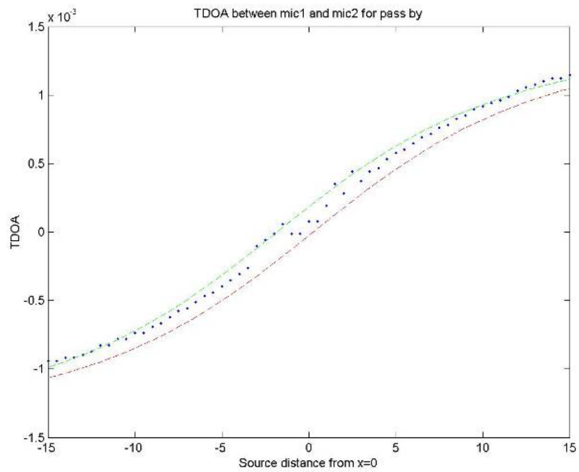
Aperture = 1m





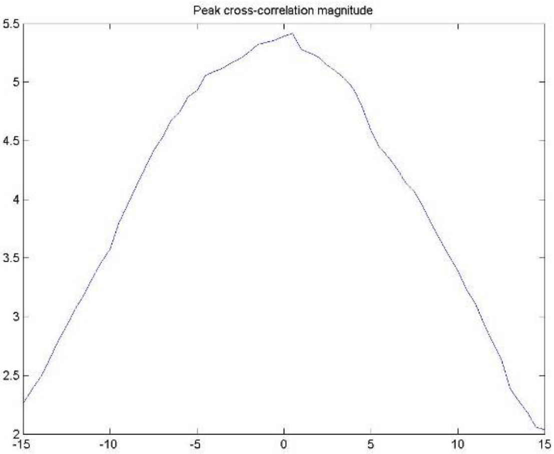
Aperture = 0.5m

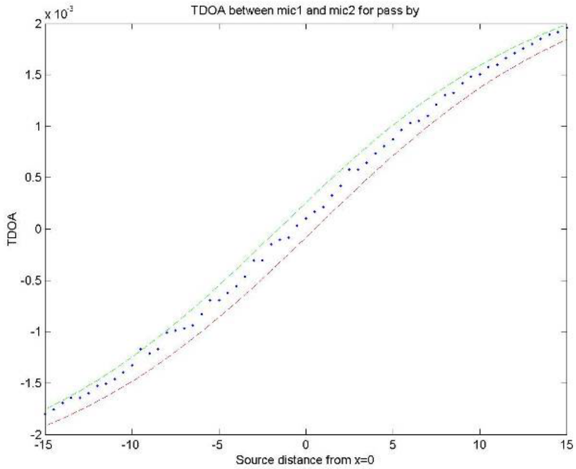




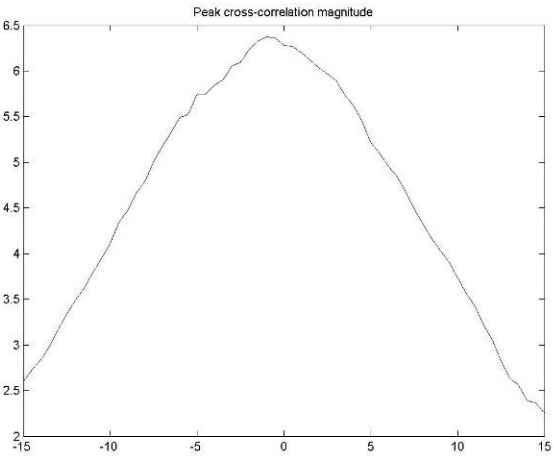
Group D

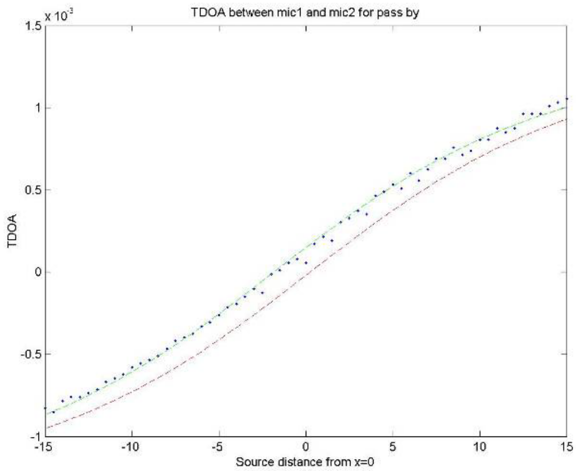
Aperture = 1m





Aperture = 0.5m





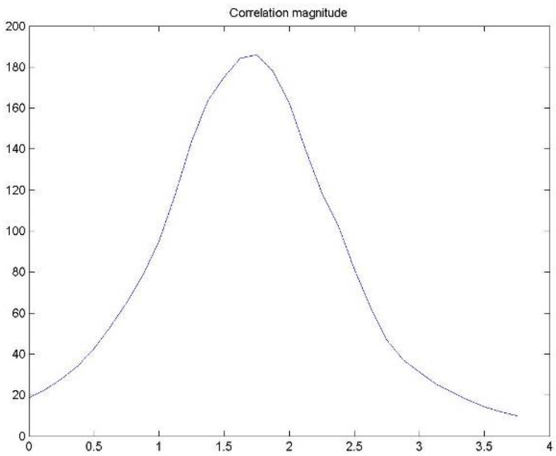
APPENDIX H - MULTIPLE VEHICLE PASS-BYS

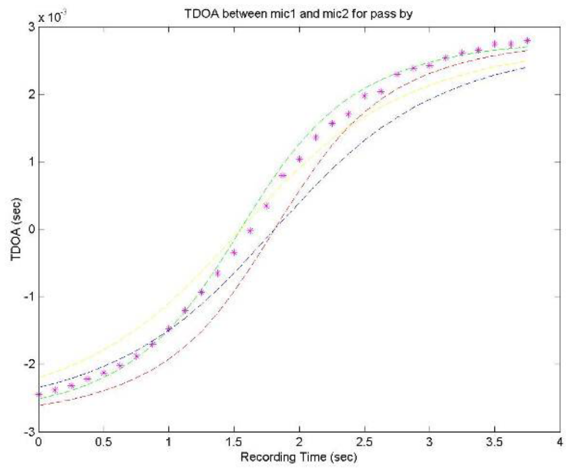
Calculated using restricted lag cross-correlation

DualSim1

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 11, Magnitude factor = 1

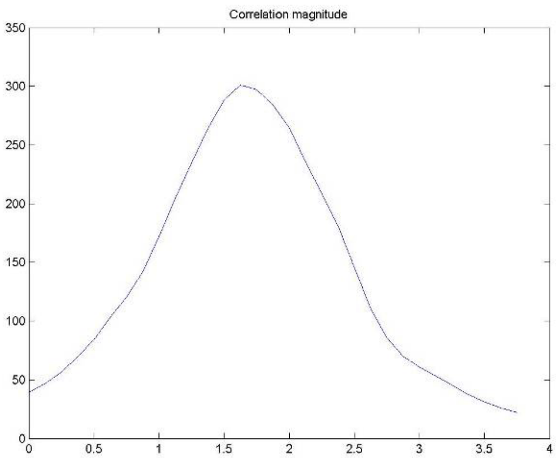


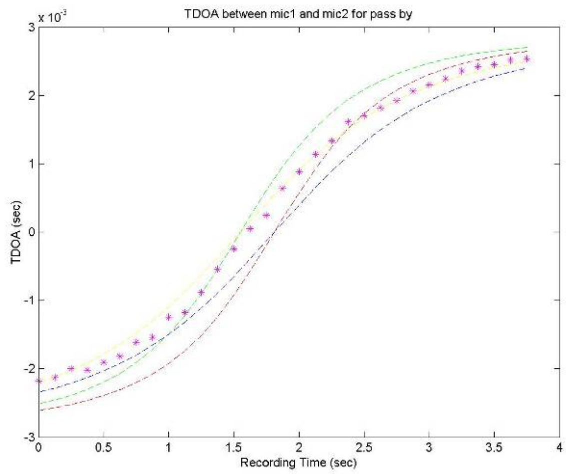


DualSim2

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 11, Magnitude factor = 2

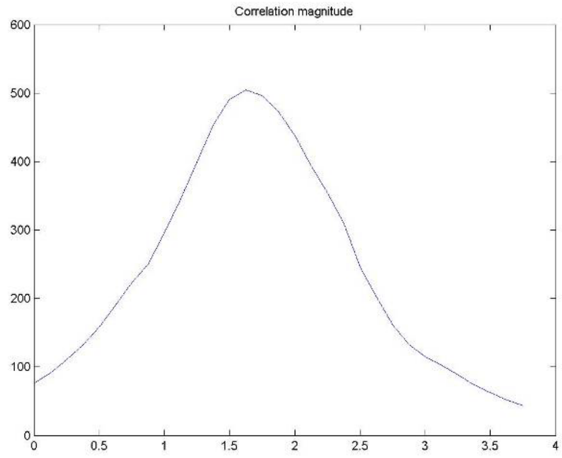


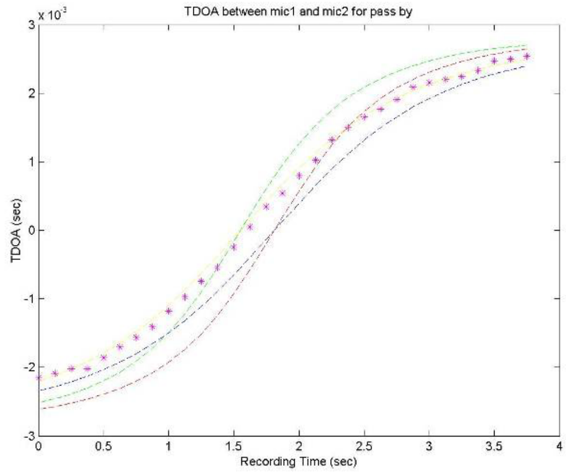


DualSim3

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 11, Magnitude factor = 3

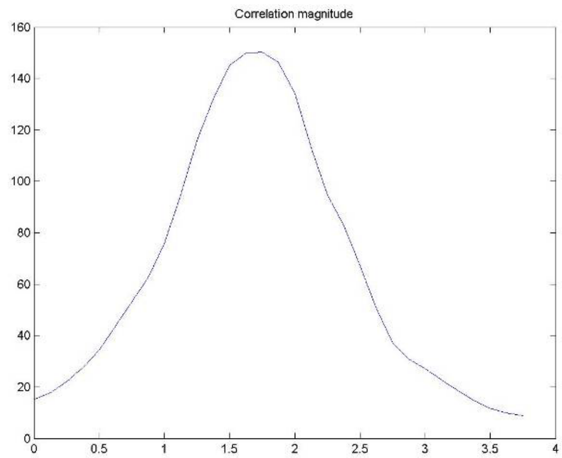


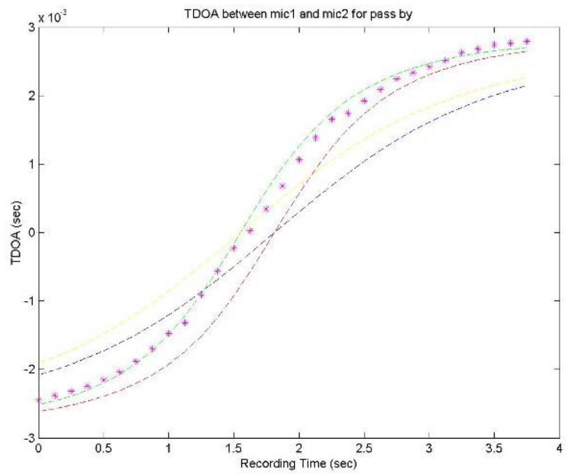


DualSim4

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 14.5, Magnitude factor = 1

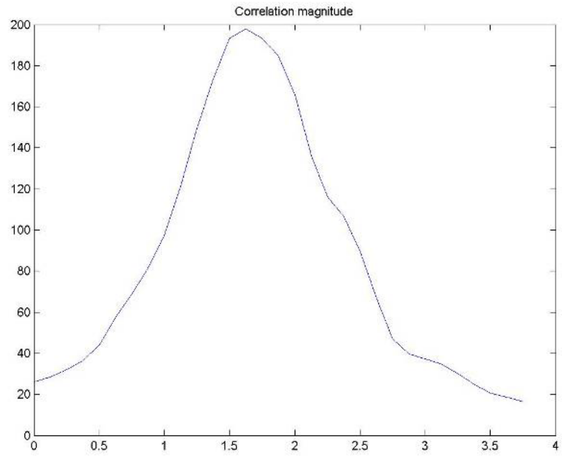


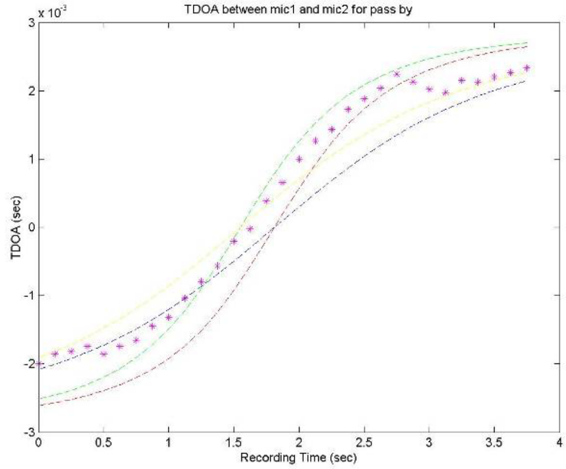


DualSim5

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 14.5, Magnitude factor = 2

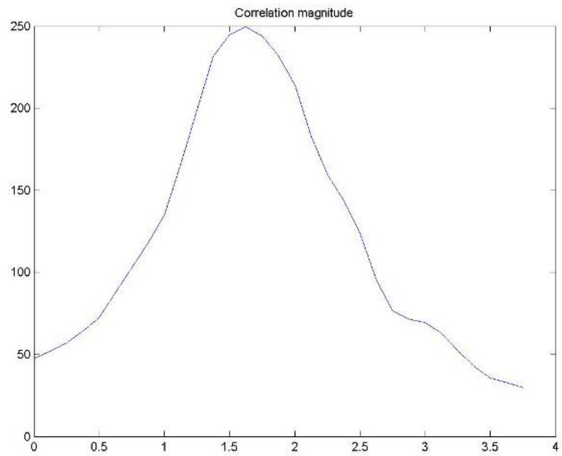


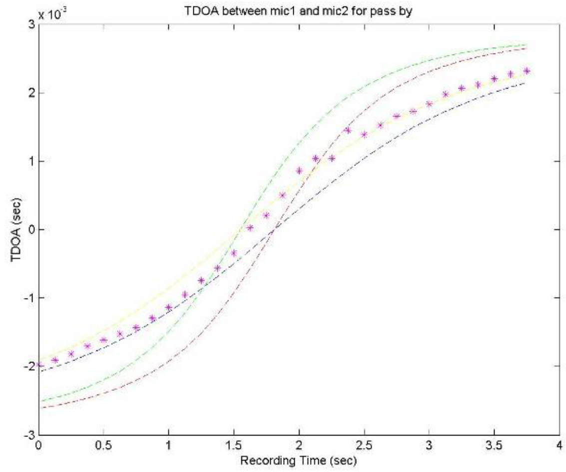


DualSim6

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 14.5, Magnitude factor = 3

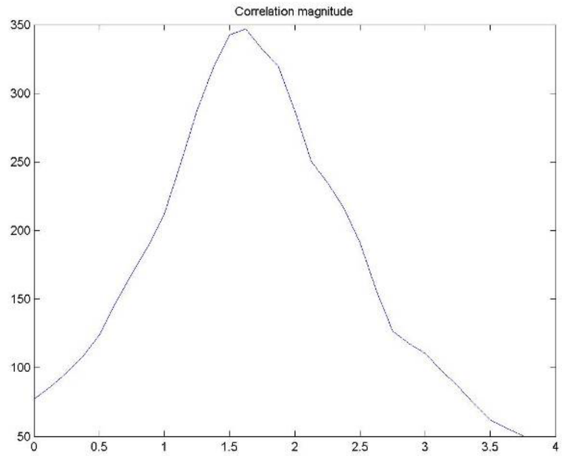


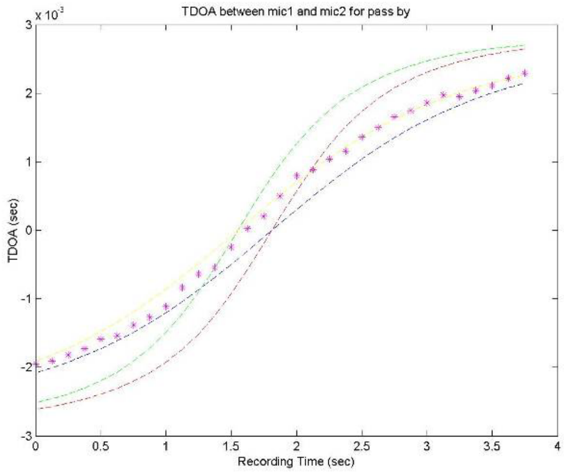


DualSim7

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 14.5, Magnitude factor = 4



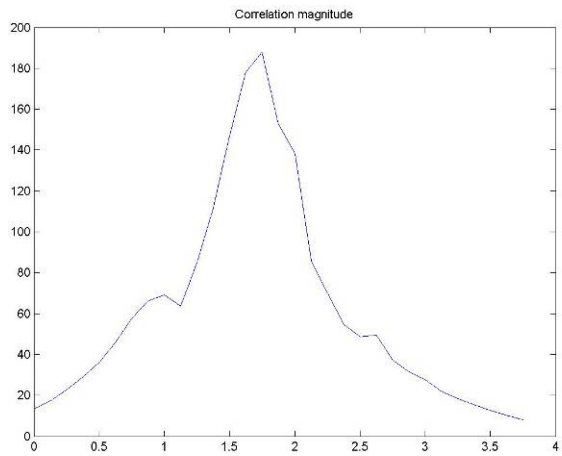


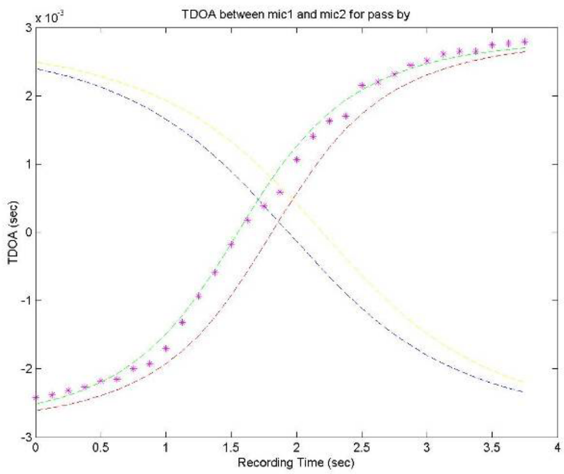
Two Directional

DualSim8

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 11, Magnitude factor = 1

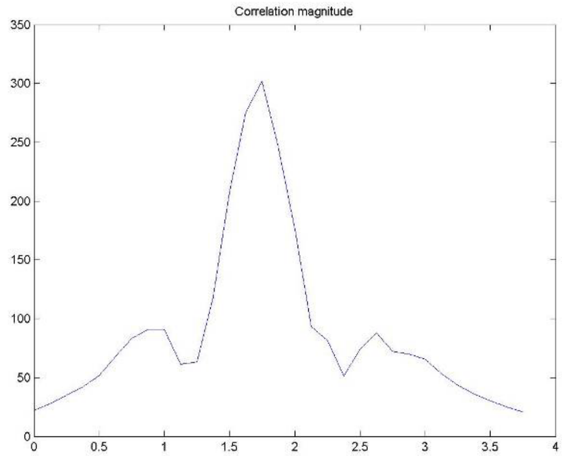


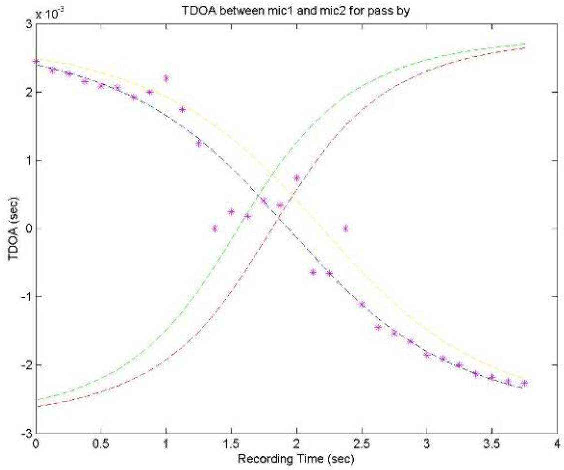


DualSim9

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 11, Magnitude factor = 2

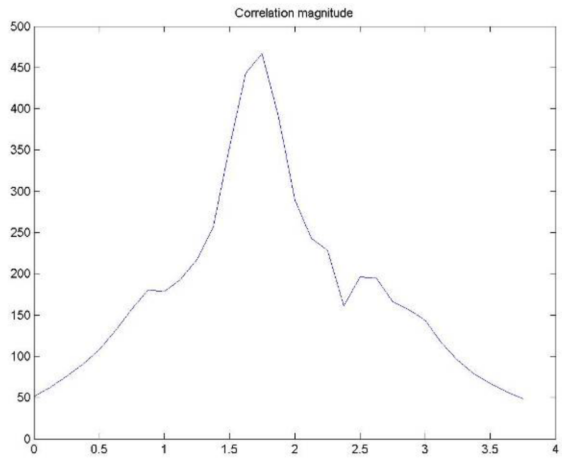


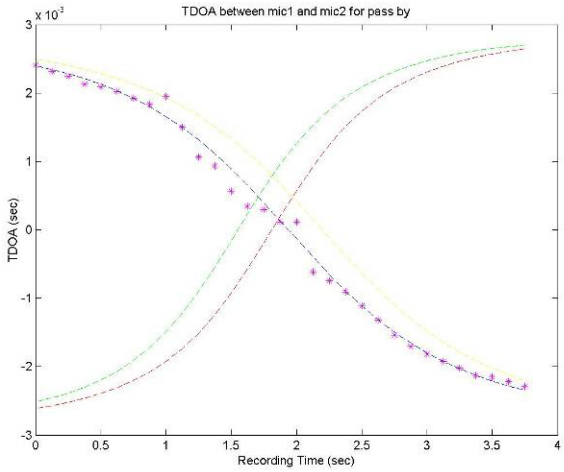


DualSim10

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 11, Magnitude factor = 3

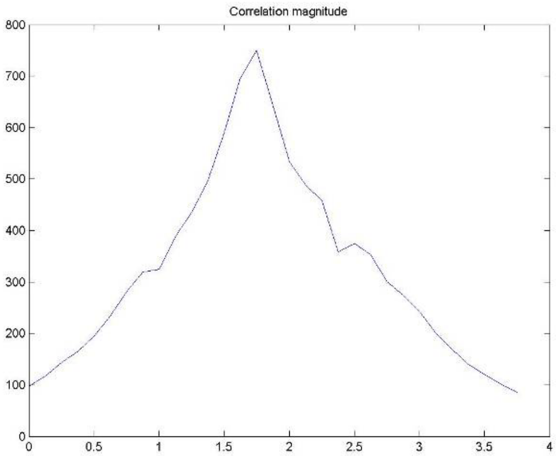


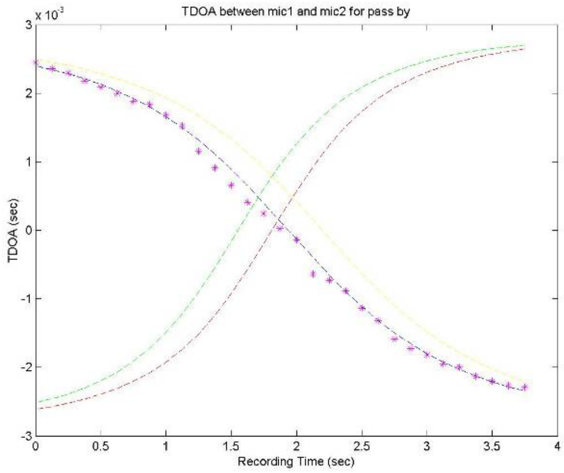


DualSim11

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -20m to 20m, Offset = 11, Magnitude factor = 4

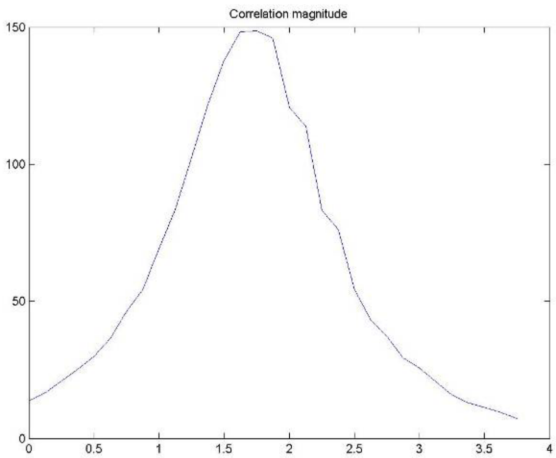


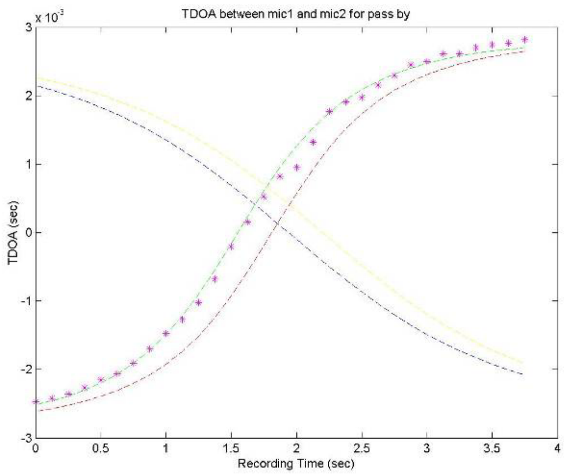


DualSim12

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 14.5, Magnitude factor = 1

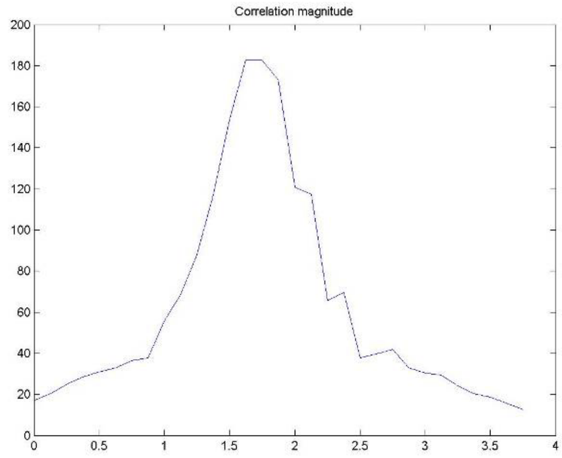


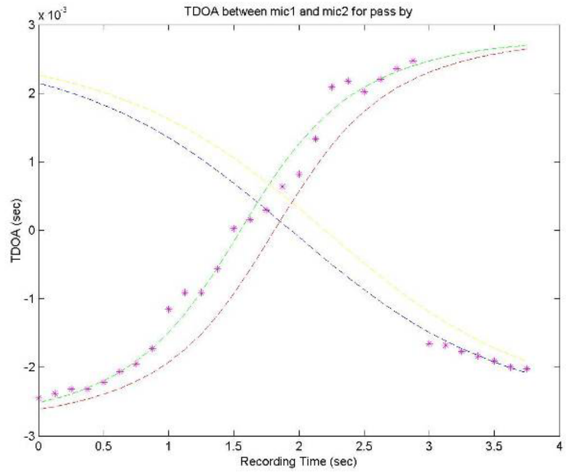


DualSim13

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 14.5, Magnitude factor = 2

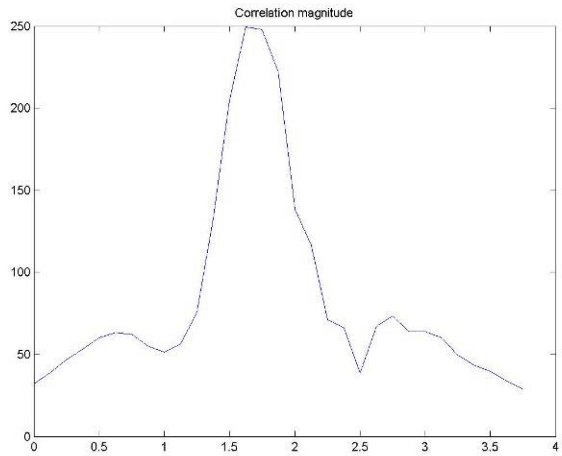


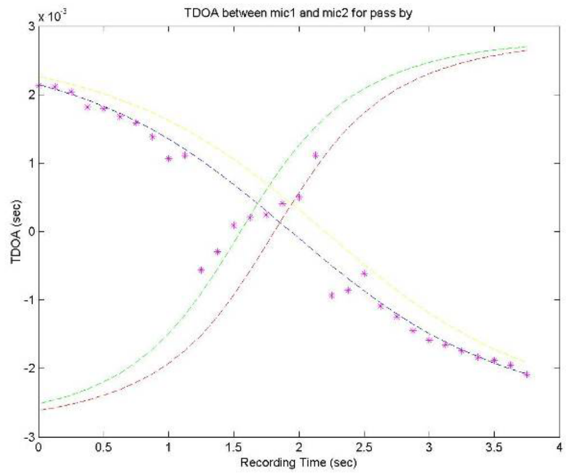


DualSim14

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 14.5, Magnitude factor = 3

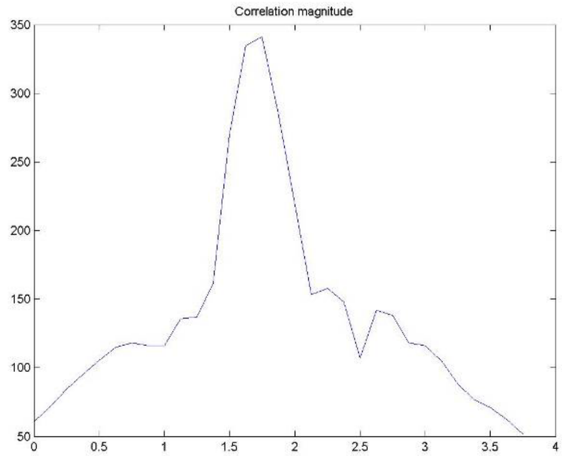


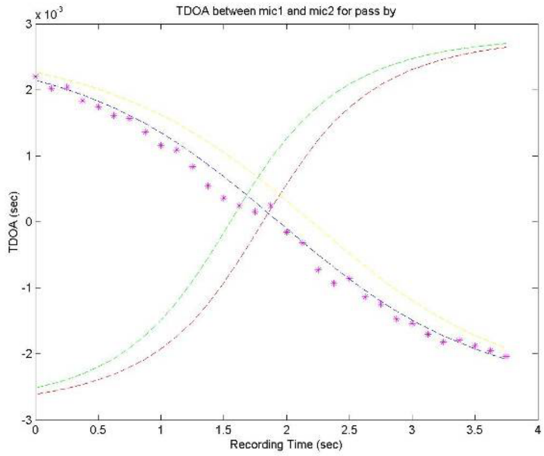


DualSim15

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 14.5, Magnitude factor = 4

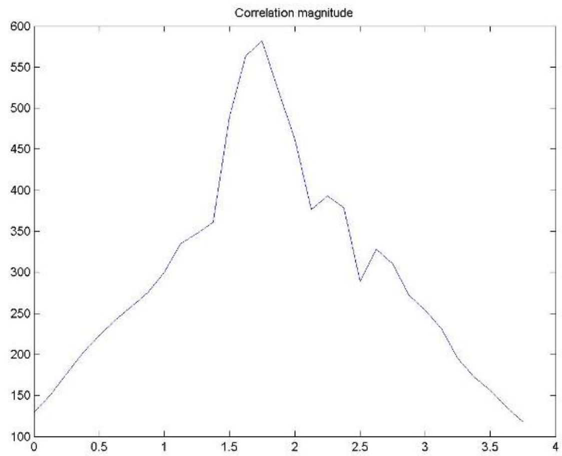


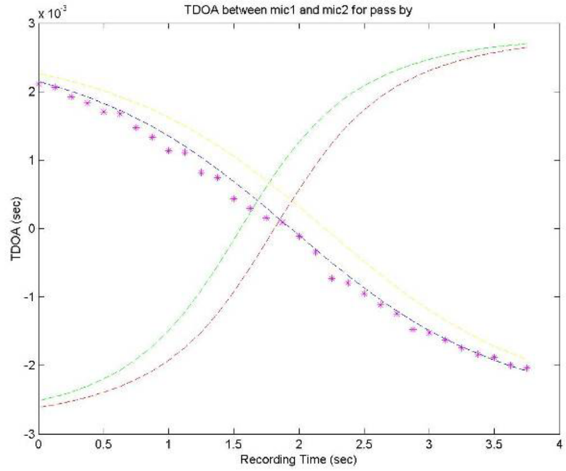


DualSim16

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 14.5, Magnitude factor = 6

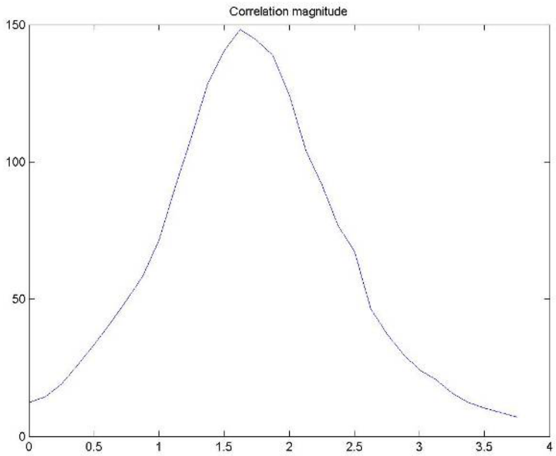


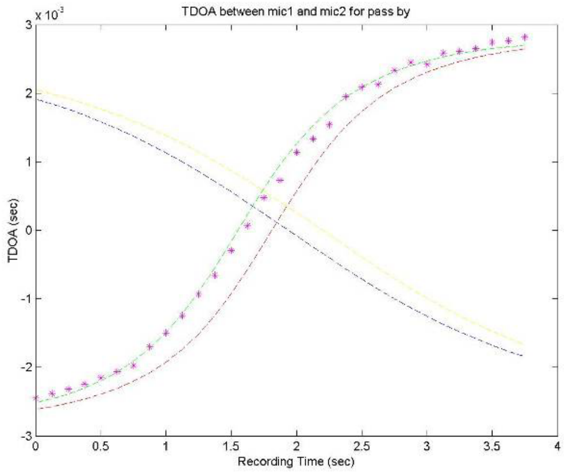


DualSim17

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 18, Magnitude factor = 1

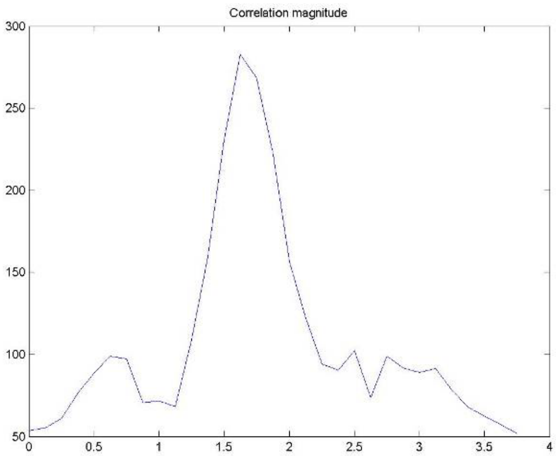


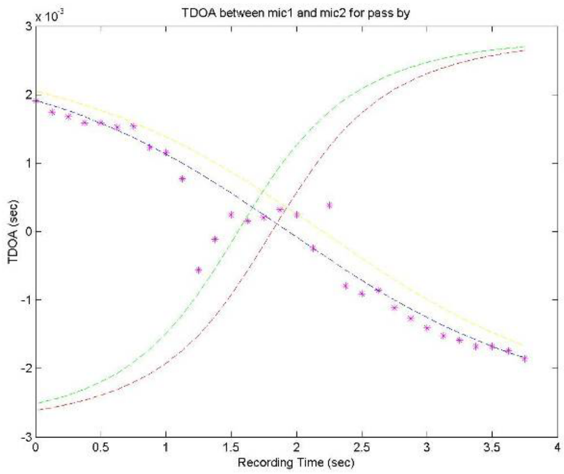


DualSim18

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 18, Magnitude factor = 5

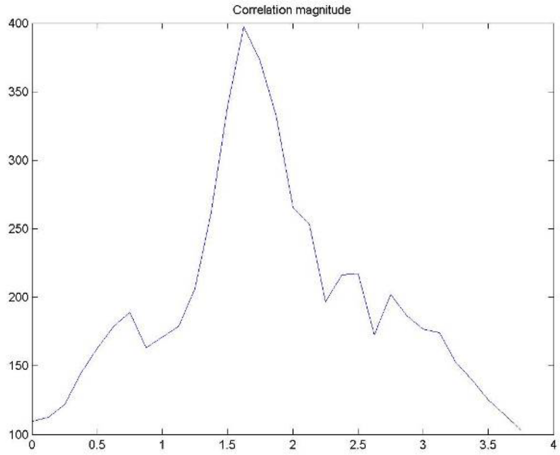


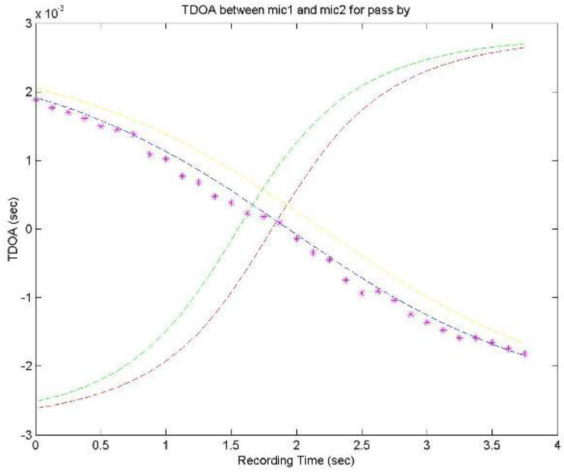


DualSim19

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 20m to -20m, Offset = 18, Magnitude factor = 7

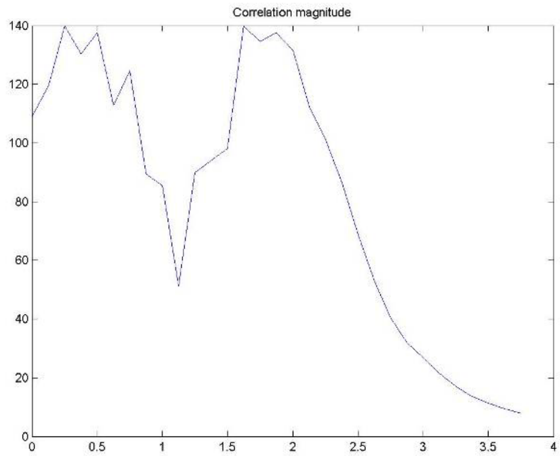


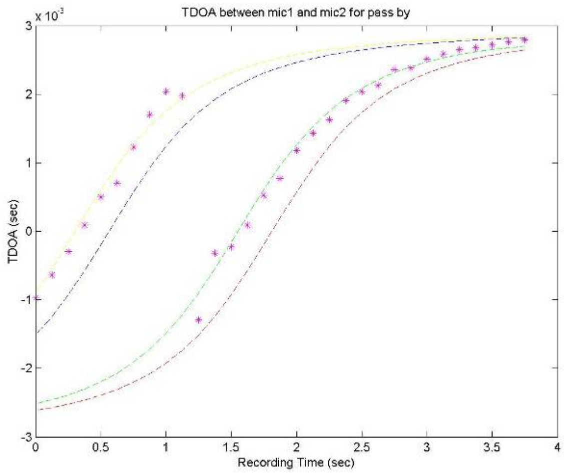


DualSim20

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -10m to 30m, Offset = 7.5, Magnitude factor = 1

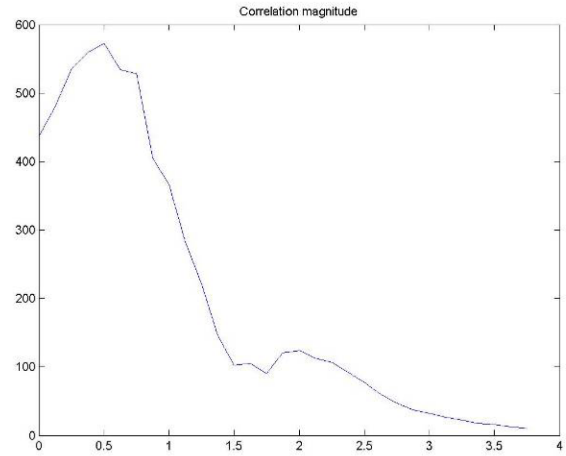


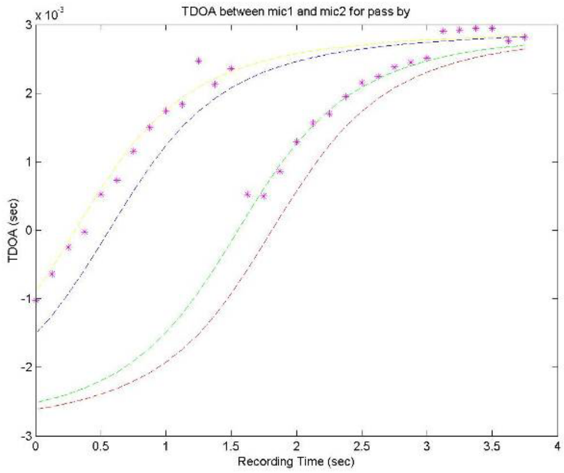


DualSim21

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -10m to 30m, Offset = 7.5, Magnitude factor = 2

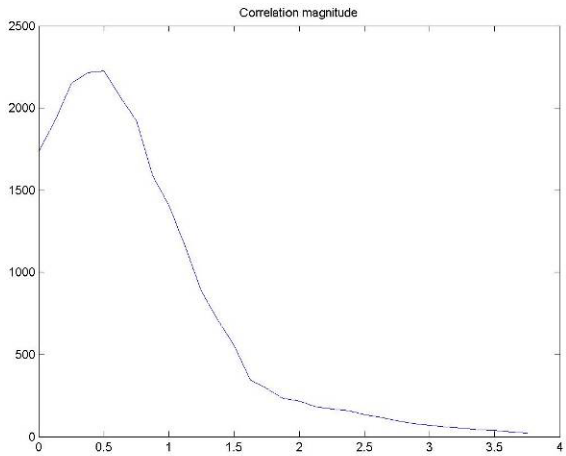


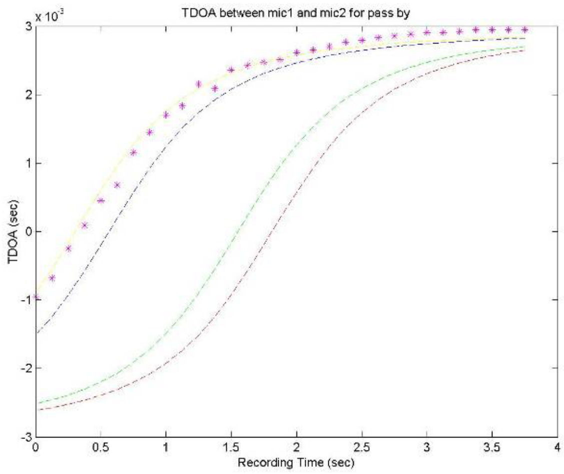


DualSim22

Source 1 -20m to 20m, Offset = 7.5, Magnitude factor = 1

Source 2 -10m to 30m, Offset = 7.5, Magnitude factor = 4





APPENDIX I – MATLAB CODE

Dual Vehicle Simulation Code

```
%-----------------------------------------------------------------
% DelaySimPass.m
% Ben Curtis 2006
%-----------------------------------------------------------------
clear all
%-----------------------------------------------------------------
% Speed of sound (m/s)
c = 340;
% Sampling frequency
Sfreq = 44100;
% Sample time
Stime = 1/Sfreq;
% Running length of simulation (sec)
RunTime = 0.125;
% Speed of vehicle km/h
speed = 60;
vel = speed*1000/3600;
% Array aperture
ap = 1;
% Maximum time and sample delay
maxlag = ap/c; maxlagS = round(maxlag*Sfreq);
% Source noise level factors
F1=1;
F2=1;
% Travel path of vehicles
StartPosR = -20:1:20;
StartPosR2 = -10:1:30;
% Offset of line of travel from microphone axis
Off = 7.5;
Off2 = 7.5;
% A vehicle travelling at 60km/h will travel 100m in 6 seconds
t=[0:Stime:RunTime];
% Generate unique signal
noise1 = randn(1,length(t)*2);
% ------------------ First Vehicle ----------------------------
%-----------------------------------------------------------------
ap1 = -ap*StartPosR(1)/abs(StartPosR(1));
for n = 1:1:length(StartPosR)
%-----------------------------------------------------------------
StartPos = StartPosR(n);
% Model amplitude variance using the varying
% distance from source to observer at a speed of 60km/h

% Position relative to observer
Xpos = StartPos:vel*Stime:vel*RunTime + StartPos;
% Distance between source and observer
Hyp = sqrt(Xpos.^2+Off^2);
% Sound intensity decrases proportionally to 1/d^2
Varamp = Hyp.^2;
%-----------------------------------------------------------------
% Model frequency shift (Doppler shift) of signal as vehicle
% travels through line offset from observer by 7.5m at a speed of 60km/h

% Extra distance required for differentiation
HypEx = sqrt(((Xpos(length(Xpos))+Stime*vel)^2+Off^2));
% Velocity of vehicle relative to microphone
RelVel = diff([Hyp,HypEx])/Stime;
% Doppler shift factor
Shift = (1./(1+(RelVel/c)));
%-----------------------------------------------------------------
```

```
% Model Signal Delay
tau = Hyp./c;
%tau=tau(1);
starttau(n) = tau(1);
endtau(n) = tau(length(tau));
% %---------------------------------------------------------------
% Freq. components
freqA = 125*Shift;
freqB = 160*Shift;
freqC = 250*Shift;
freqD = 500*Shift;
freqE = 800*Shift;
freqF = 1250*Shift;
% Create component sine waves
wavA=sin(2*pi*freqA.*(t-tau));
wavB=sin(2*pi*freqB.*(t-tau));
wavC=sin(2*pi*freqC.*(t-tau));
wavD=sin(2*pi*freqD.*(t-tau));
wavE=sin(2*pi*freqE.*(t-tau));
wavF=sin(2*pi*freqF.*(t-tau));
%---------------------------------------------------------------
% Simulate background noise
noise = noise1(floor((t-tau)*Sfreq+5000));
%---------------------------------------------------------------
% Create simulated vehicle sound waveform
waveF=wavA*6+wavB*8+wavC*9.5+wavD*5+wavE*2+wavF*2+8*noise;
% Vary amplitude of vehicle signal
simA(n,:) = F1*waveF./Varamp;
end
% ----------------------- Second Vehicle ------------------------
% Adjust directivity
ap2 = -ap*StartPosR2(1)/abs(StartPosR2(1));
noise2 = randn(1,length(t)*2);
for a = 1:1:length(StartPosR2)
%---------------------------------------------------------------
StartPos = StartPosR2(a);
Xpos2 = StartPos:vel*Stime:vel*RunTime + StartPos;
Hyp2 = sqrt(Xpos2.^2+Off2^2);
Varamp = Hyp2.^2;
HypEx2 = sqrt(((Xpos2(length(Xpos2))+Stime*vel)^2+Off2^2));
RelVel2 = diff([Hyp2,HypEx2])/Stime;
Shift2 = (1./(1+(RelVel2/c)));
tau2 = Hyp2./c;
starttau1(a) = tau2(1);
endtau1(a) = tau2(length(tau2));
freqA = 110*Shift2;
freqB = 180*Shift2;
freqC = 230*Shift2;
freqD = 550*Shift2;
freqE = 870*Shift2;
freqF = 1150*Shift2;
wavA=sin(2*pi*freqA.*(t-tau2));
wavB=sin(2*pi*freqB.*(t-tau2));
wavC=sin(2*pi*freqC.*(t-tau2));
wavD=sin(2*pi*freqD.*(t-tau2));
wavE=sin(2*pi*freqE.*(t-tau2));
wavF=sin(2*pi*freqF.*(t-tau2));
noise = noise2(floor((t-tau2)*Sfreq+5000));
waveF2=wavA*6+wavB*8+wavC*9.5+wavD*5+wavE*2+wavF*2+8*noise;
simB(a,:) = F2*waveF2./Varamp;
end
Len = maxlagS;
scale= -Len:1:Len;
% Add vehicle waveforms
sim1 = simA(2:a-1,:) + simB(2:a-1,:);
sim2 = simA(2-ap1:a-1-ap1,:) +simB(2-ap2:a-1-ap2,:);
d = 6:1:length(StartPosR2)-5;
```

```
for f=1:length(d)
    Ac(f,:) = xcorr(sim1(d(f),:),sim2(d(f),:),maxlagS);
    maxcorr(f,:) = max(Ac(f,:));
    maxloc = find(Ac(f,:)==maxcorr(f,:));
    samdel(f,:) = scale(maxloc);
    del(f) = samdel(f)/Sfreq;
    y1 = starttau(d+ap1)-starttau(d);
    y2 = endtau(d+ap1)-endtau(d);
    y1a = starttau1(d+ap2)-starttau1(d);
    y2a = endtau1(d+ap2)-endtau1(d);
end
% plot TDOA waveform
dist = 0:0.125:0.125*(length(d)-1);
plot(dist,del,'m*');
hold on;
xlabel('Recording Time (sec)')
ylabel('TDOA (sec)')
title('TDOA between mic1 and mic2 for pass by')
% plot real delays
plot(dist,y1,'r--');
plot(dist,y2,'g--');
plot(dist,y1a,'b--');
plot(dist,y2a,'y--');
% plot correlation magnitude
figure(2);
plot(dist, maxcorr);
title('Correlation magnitude')
```

Location Optimization Code

```
%--------------------------------------------------------------------------
% Ben Curtis 2006
% Function for localization of a noise source using microphone co-ordinates
%--------------------------------------------------------------------------
% Optimization of circle equations for localization
clear;clc;
% TDOA distances
r1=3.7334;
r3=-0.2763;
% Initialise minimum error value
min_err=10;
% Mic coords
x1=0;y1=0;
x2=5;y2=0;
x3=10;y3=0;
% Location precision
div=0.5;

 for n=0:1:10/div
     for m=0:1:10/div
        xb=n*div;
        yb=m*div;
        ra=sqrt(xb^2+yb^2);
        xa=xb*r1/ra;
        ya=yb*r1/ra;
        rg=sqrt((x3-xb)^2+(yb-y3)^2);
        xg=x3-(x3-xb)*r3/rg;
        yg=yb*r3/rg;
        rb=sqrt((x2-xb)^2+(yb-y2)^2);

        err1(n+1,m+1)=abs((xa-x1)^2+(ya-y1)^2-r1^2);
        err2(n+1,m+1)=abs((xg-x3)^2+(yg-y3)^2-r3^2);
        err3(n+1,m+1)=abs((xa-xb)^2+(ya-yb)^2-rb^2);
        err4(n+1,m+1)=abs((xg-xb)^2+(yg-yb)^2-rb^2);
        err5(n+1,m+1)=abs((x2-xb)^2+(y2-yb)^2-rb^2);

        errTn=err1(n+1,m+1)+err2(n+1,m+1)+err3(n+1,m+1)+err4(n+1,m+1)+err5(n+1,m+1);
```

```
        if errTn < min_err
            min_err=errTn;
            xmin=xb;row=n;
            ymin=yb;col=m;
        end

    end
end

errT=(err1+err2+err3+err4+err5);
xmin
ymin
min_err
row*div
col*div
```

REFERENCES

Acoustic Research Laboratories, *Noise Camera: Automated Traffic Noise Analysis System*.

Australian Design Rule 28/01: External Noise of Motor Vehicles (1989).

Austroads Inc (2004), *AP-G11.3-04: Guide to Traffic Engineering Practice, Part 3 – Traffic Studies.*

Austroads Inc (2005), *AP-R277-05: Modelling, Measuring and Mitigating Road Traffic Noise*.

Bendat, J. S. & Piersol, A. G. (1993), *Engineering Applications of Correlation and Spectral Analysis – 2nd edition*, John Wiley & Sons, New York.

Bies, D. A. & Hansen, C. H. (2003), *Engineering Noise Control: Theory and Practice – 3rd edition,* Spoon Press, London.

National Road Transport Commission (2003), *Engine Brake Noise: Development of a Roadside Test Procedure*.

Harris, R. W. & Ledwidge, T. J. (1974), *Introduction to Noise Analysis*, Pion Ltd, London.

Ifeachor, E. C. & Jervis, B. W. (1993), *Digital Signals Processing: A Practical Approach*, Addison-Wesley Publishers Ltd, Wokingham.

Kirkwood, B. C. (2003), ‘Acoustic Source Localization Using Time-Delay Estimation’, PhD thesis, Technical University of Denmark.

Kolawole, M. O. (2002), *Radar Systems, Peak Detection and Tracking*, Newness, Oxford.

Leis, J. (2002), *Digital Signals Processing – A MATLAB Based Tutorial Approach*, Research Studies Press Ltd, Hertfordshire.

Lu L, Jiang H & Zhang H, *A Robust Audio Classification and Segmentation Method*, Microsoft Research, China.

Lynn, P. A. & Fuerst, W. (1998), *Introductory Digital Signals Processing with Computer Applications – 2nd edition*, John Wiley & Sons Ltd, London.

Model BETA 58A User Guide, (2005), Shure Incorporated.

Queensland Department of Main Roads (2004), *Engine Compression Braking: Guidelines for the Use of Signs to Reduce Heavy Vehicle Noise.*

Queensland Department of Main Roads (2004), *Road Planning and Design Manual.*

Richards, M. A. (2005), *Fundamentals of Radar Signal Processing*, McGraw-Hill, New York.

Sampan, S. (1997), 'Neural Fuzzy Techniques In Vehicle Acoustic Signal Classification', PhD thesis, Virginia Polytechnic Institute and State University.

SmarTek Systems Inc. (2003), *SmarTek Acoustic Sensor – Version 1: Installation and Setup Guide*.

Snook, C. (2005), *MEC3303 System Design: Workbook 1, Engineering Noise Control*, University of Southern Queensland, Toowoomba.

Wolfram Research (2006), Wolfram Mathworld, <http://mathworld.wolfram.com>