Remote Sensing Of Road Surface Conditions

By

Mohammad Abbas

A thesis submitted to the School of Electronic, Electrical and Systems Engineering of The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

2015

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

Abstract

The remote real time identification of road surfaces is an increasingly important task in the automotive world. The development of automotive active safety system requires a remote sensing technology that alerts drivers to potential hazards such as slippery surfaces caused by water, mud, ice, snow etc. This will improve the safety of driving and reduce the road accidents all over the world. This thesis is dedicated to the experimental study of the feasibility of an affordable short-range ultrasonic and radar system for road surface recognition ahead of a vehicle. It introduces a developed novel system which can recognize the surfaces for all terrains (both on-road and off-road) based on the analysis of backscattered signals. Fundamental theoretical analysis, extensive modelling and practical experiments demonstrated that the use of pattern recognition techniques allows for reliable discrimination of the surfaces of interest. The overall classification system is described, including features extraction and their number reduction, as well as optimization of the algorithms. The performance of 4 classification algorithms was assessed and evaluated to confirm the effectiveness of the system. Several aspects like the complexity of the classification algorithms and the priori knowledge of the environment were investigated to explore the potential of this research and the possibility of introducing the surface classification system into the automotive market in the nearest future.

Acknowledgements

Table of Contents

List of Figures

Chapter 1: Introduction

Chapter 2: Literature Review

Chapter 3: Clutter

Chapter 4: Surface Identification Data Collection

Chapter 5: Classification Systems

Chapter 6: Evaluation and Results

Chapter 7: Conclusion and Future Work

List of Tables

Chapter 6: Evaluation and Results

Glossary of Abbreviations

ABS   Antiblock Brake System

BSD   Blind Spot Detection

ACC   Adaptive Cruise Control

LIDAR   Laser Illuminated Detection and Ranging

EM   Electromagnetic Wave

ENDURA   Energy-Duration-Range

KNN   K Nearest Neighbors

FFT   Fast Fourier Transform

HRA   Hot Rolled Asphalt

SMA   Stone Mastic Asphalt

KAM   Kirchhoff Approximation Method

MLE   Maximum Likelihood Estimation

MLP   Multilayer Perceptron

SVM   Support Vector Machines

MDC   Minimum Distance Classification

ANN   Artificial Neural Network

FNN   Feedforward Neural Network

BP   Backpropagation

RCS   Radar Cross Section

ISM   Industrial, Scientific, and Medical radio bands

rms   Root Mean Square

Chapter 1

Introduction

1.1   Introduction to Remote Sensing

The emergence of Technology as an integral part of our lives left its stamp on all and every field. The automotive industry is one example of the industries that has been greatly touched by technology; in fact, it has changed forever. Motor vehicles are recently becoming a platform for a sophisticated and complex electronic systems that makes the driver's journey safer [1]. Amongst many technological systems, remote real-time identification of road surfaces is an increasingly important task and the development of remote surface recognition system is an important step in ensuring road safety.

Remote sensing is the technique of obtaining reliable information about a distant target. This can be achieved by using sensors that emit electromagnetic radiation towards the object and then measure the reflections which have information about the properties of that object. It might use different sensors with different force fields like electromagnetic field, and mechanical vibrations (acoustic). In its broadest, remote sensing includes the use of satellites, spacecrafts and aerial observations of the surface or atmosphere of the Earth and the other solar planets. These observations can be used for several purposes like military

intelligence, planimetric maps, and weather forecasting. Nevertheless, most modern remote sensing is restricted to investigate the surface of earth as the object of interest [2].

The energy reflected by the surface is monitored by the remote sensing system. These systems can be categorized into three categories depending on the source of the energy and the way it interacts with the surface:

1. **Reflected solar radiation sensors:** Relying on the sunlight as a source, these sensors detect the material and the physical structure of the surface through the solar radiation reflected or scattered from the surface.

2. **Thermal infrared sensors:** These sensors provide thermal information about the material of the surface by detecting the thermal radiation that the surface emits.

3. **Active sensors:** These sensors broadcast the surface with a microwave, sound wave, or laser, and then detect the surface properties (roughness, surface material, shape, etc.) through the reflected signal. Examples of active sensor are imaging radar sensors, LIDAR, and Ultrasonic sensors.

Reflected solar radiation sensors and thermal infrared sensors are passive systems that rely on the sunlight as their energy source. These sensors wouldn't be useful during certain times of the day when there is no sunlight or when the temperature of the surface changes. The active sensors are common systems used in remote sensing of surfaces and in automotive industry, and will be the primary focus of this research. This chapter will introduce remote sensing in its broadest, and then highlights the application we are aiming to use. It follows up with a motivation for the present work, and the contribution of this research towards solving the problem.

1.2 Brief History of Radar Systems for Remote Sensing

Though remote sensing emerged since the 1600s when Galileo used optical equipment to observe ships, modern remote sensing began with Gaspard-Felix Tournachon who took photographs of Paris from a hot air balloon in 1858 [3]. In the present remote sensing, data are generally obtained from satellites, high flying aircraft, radars, etc. Although each has its own advantages and disadvantage, the focus in this section will be on radars. In parallel with the growth of remote sensing, a new technique was developed for radio detection. In 1886, Heinrich Hertz showed the ranging capabilities by detecting reflected radio waves from solid objects. Later in 1903, detection of radio waves reflected from ships was experimented by Hulsmeyer. Shortly after, investigations about the development of radar (radio detection and ranging) for the detection of ships and aircrafts began in the United States and Great Britain [4].

The World War II brought huge efforts to the development of radars. The years just before the war, there was an enormous requirement to achieve accurate detection and ranging which became available during the 1940's [2]. Remote sensing continued during the World War II and played a big role in detection of nuclear missiles (Cuba 1962). Vegetation remote sensing started by the discovery of W.M. Stinton between 1956 and 1958. He found out the presence of absorption peaks in the spectra of reflected signals that he linked with chlorophyll. It was later explained that these features results from absorption due to deuterated water [3].

Nowadays, radar's applications in remote sensing are very diverse and are used in many industries. Ocean's surface is observed by radars, not only to detect ships, but to sense the

wave structure and detect wind patterns. Lower atmosphere is observed by radars, not only to detect aircrafts, but to observe the weather conditions. Radars are also being used to observe the other planets in the solar systems.

1.3 Motivation for the present work

In developed countries like the UK, USA, Germany, etc., cars are a fact of daily life. Even in the developing countries, cars became the essential aid to everyday' s life. The huge number of cars on the roads leads to traffic accidents and casualties year after year [1]. Vehicle accidents had caused the death of 2,605 people in the UK in 2009. It accounts for 13% of all external causes of deaths. Young people aged 15-24 years are four times more likely to be killed in a car accident than by shooting, alcohol, or self-harm **[5]**. In 2010, the official statistical publication of the UK Department of Transport (DfT), entitled Reported Road Casualties Great Britain (RRCGB), showed that 1,850 cases of the fatalities that took place in Great Britain are caused by road accidents, and a total number of 206,798 people were injured due to vehicle accidents. Figure 1.1 [6] shows the statistics of the road accident casualties(including fatalities) in Great Britain from 1942 to 2012. The casualties increased from 148000 in 1942 to 398000 in 1965 due to traffic increase on roads. After 1999, there was a significant decrease in the number of casualties to reach 19600 in 2012.

The number of fatalities involved in road deaths is a major issue that many countries are investigating. The state of the art safety systems deployed in vehicles like automatic braking system, airbags, and automatic cruise control (ACC) couldn't change the death rates due to vehicle accidents in a significant way. These technologies address the

Reported road accident casualties. Great Britain 1942 to 2012

Casualties (thousands)
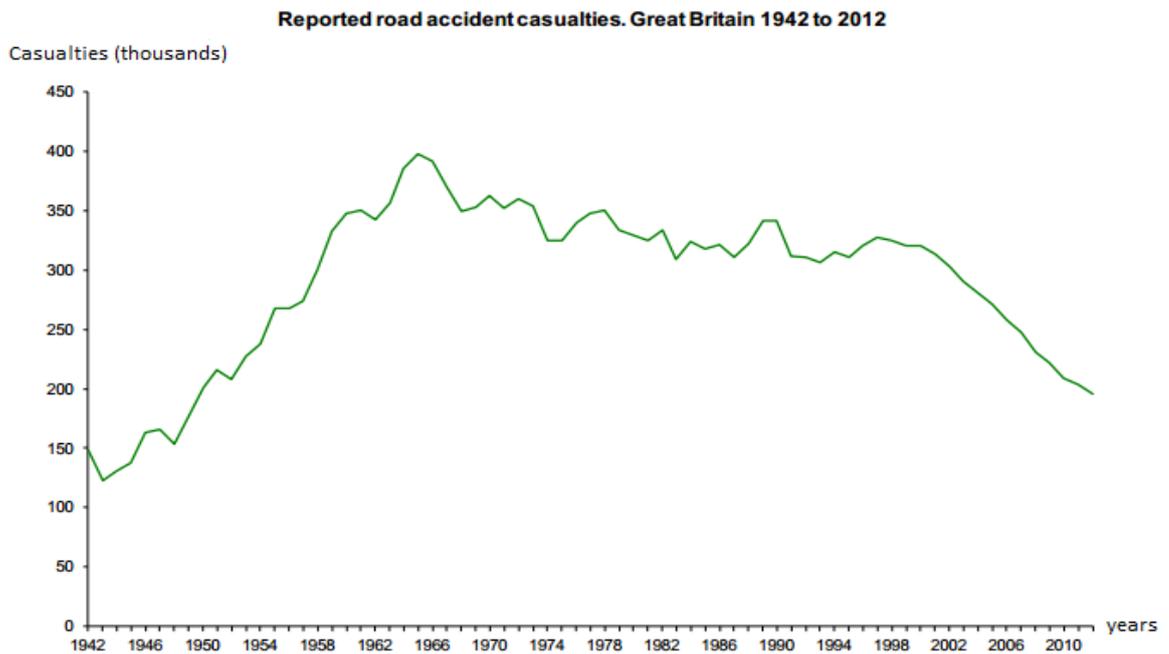
Figure 1.1 Road accident casualties (thousands) vs. years in Great Britain[6].

survivability of the driver, however it doesn't consider the cause of the accident and how to stop it. Most of the vehicle's manufacturers started to employ short range sensors to address a large number of applications. Besides the parking assistance system which use ultrasonic technology, companies like Cadillac STS begun to offer Blind Spot Detection (BSD) which is an alarm system used to warn the driver of the presence of vehicles on its blind spot side. MmWave radar at 77 GHz is being used by the Adaptive Cruise Control (ACC) application in the automotive industry, to automatically increase or decrease the speed of a vehicle on a highway maintaining a fixed time delay with the vehicle in front. In 2005, Mercedes introduced the "Brake Assist Plus" and "Pre-Safe break" in which short range radar at 24 GHz and long range sensor are used to alert the driver when there is a high possibility of collision and apply braking in case he did not attempt to prevent the crash. Many firms produced other systems that basically deliver the same job as Pre-Safe break; for instance, Collision Mitigation System. An advantage of such short range sensors is that the same hardware, with certain specifications, can address many applications [1]. Jaguar Land Rover

had recently revealed a new sensing technology used to interpret the driver's face and eyes to reduce distracted driving and alert the driver when he becomes inattentive [8]. A number of applications that is being used in the automotive market and its description are summarized in table 1.1.

Application	Description	Sensors	Aim
Adaptive Cruise Control	Maintains a safe distance from the vehicles in front.	Radar/laser	Safety
Automatic Parking	Detects the parking space size and distance from the roadside, and then drives the car into the parking space.	Radar/camera	Comfort
Collision Avoidance system	Detects an imminent crash and alerts the driver.	Radar/Camera/laser	safety
Back Up Aide	Indicates objects close to the rear of the car while parking.	Radar/ultrasonic	comfort
Blind Spot Detection	Detects other vehicles on the driver's side and rear.	Radar/camera	safety
Brake assist	Applies braking when the driver fails to react.	Radar	Safety
Lane Change Assistant Side Assist	Monitors the areas to the left and right of the car and warns the driver of a potentially hazardous situation	Radar/camera	safety

Table 1.1 Safety applications in the automotive industry

Most of these applications provide comfort or safety to the drivers. ACC, BSD, Collision Avoidance System, and Brake Assist warn the driver or the vehicle before the crash takes place. This includes alerting the driver, pre-arming airbags, altering vehicle suspension, etc.

6

Nevertheless, there is no commercially available safety system in the automotive market that can detect the road surface condition or type of the surface ahead of the vehicle. Such an application can provide safety in semi automatic or fully automatic systems. In semi automatic systems, the application alerts the driver about slippery surfaces, due to weather conditions or deposit on the road, which helps him to avoid risky situations and gives him around 2 seconds of reaction time. On the other hand, fully automatic system will brake, slow down, change gear or activate vehicle system to optimize traction and progress.

Such a safety system is essential as the biggest factor contributing to the cause of road accidents in the UK is the "driver error/reaction" (68% of all crashes). Travelling too fast for the conditions causes 10.2% of all accidents, while slippery road due to weather causes 10.1%. Surprisingly, road environment in general causes 15.1% of all crashes. This is much higher than the vehicle defect factor which only causes 2% of all accidents in the UK [7]. Figure 1.2 [1] shows that if the driver is given two seconds of reaction time, the probability of collision decreases from 100 to nearly 0%. Our aim is to develop a remote sensing safety system capable of alerting the driver to potential hazards, improving the driver reaction time and identifying the surfaces ahead of a vehicle.
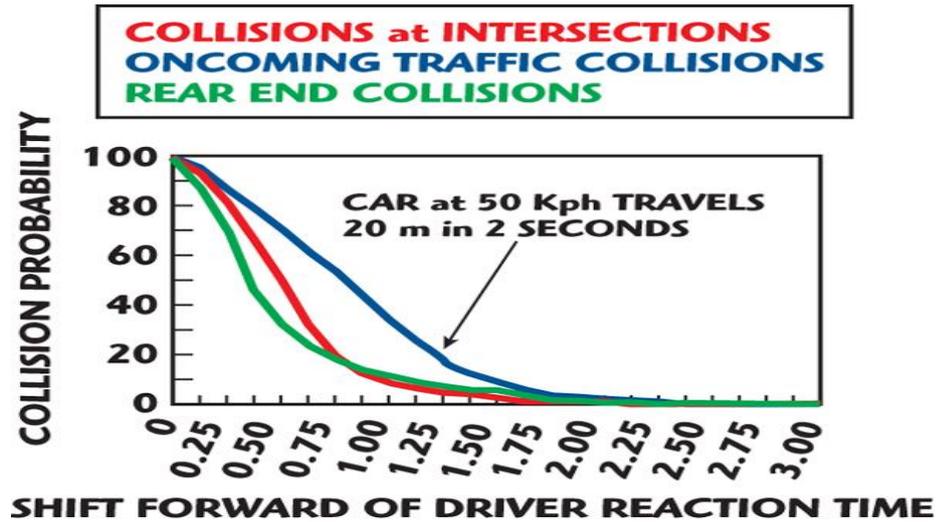
Figure 1.2 Collision probabilities with respect to driver reaction time (seconds) [1].

1.4 Problem Setting and Research Contribution

As shown in Table 1.1, each application uses one or more sensors. There is no one single technology that can cover all the applications. The challenge is to find the right technology that can fulfil the aim of this research. The main objective is to develop an affordable automotive active safety system that is able to remotely identify potential hazards from an upcoming terrain such as slippery surfaces caused by water, wet grass, ice, snow, etc. in order to alert the driver and increase the safety of driving.

A basic feasibility investigation of short-range sensing systems was carried out. Radar along with other sensor technologies such as infrared, LIDAR and visible light and ultrasonic sensors are widely used in the automotive applications. The use of different sensors for surface roughness recognition and detection of low-friction spots caused by water, ice or snow has been investigated in many studies that will be shown in the next chapter. Table 1.2

[1] shows the advantages and disadvantages of different technologies that can be used for road surface identification.

	SHORT-RANGE RADAR (24 GHz UWB)	LONG-RANGE RADAR (77 GHz)	LIDAR	ULTRASONIC	VIDEO
RANGE MEASUREMENT <2 m	++	–	0	++	–
RANGE MEASUREMENT 2 to 30 m	++	++	++	–	–
RANGE MEASUREMENT 30 to 100 m	–	++	+	– –	–
ANGLE MEASUREMENT <10 DEGREES	+	+	++	–	++
ANGLE MEASUREMENT >30 DEGREES	0	–	++	0	++
ANGULAR RESOLUTION	0	0	++	–	++
DIRECT VELOCITY MEASUREMENT	++	++	– –	0	– –
OPERATION in RAIN	++	+	–	0	0
OPERATION in FOG or SNOW	++	++	–	+	–
BLOCKAGE (DIRT on SENSOR)	++	++	0	+	– –
OPERATION at NIGHT	++	++	++	++	– –

++	IDEALLY SUITED	– FEASIBLE, ONLY WITH LARGE EFFORT
+	GOOD PERFORMANCE	– – IMPOSSIBLE
0	POSSIBLE, MODEST PERFORMANCE	

Table 1.2 Different automotive sensor technologies and its applications [1]

The table depicts the trade space among the competing technological solutions. Although video and LIDAR technologies have high resolution, its performance in rain, fog, snow, and severe weather conditions is unfeasible if not impossible in many cases. Radar has excellent performance during bad weather conditions, while ultrasonic sensor is feasible with modest performance. The advantage of microwave sensors is that, in contrary to optical sensors, forward looking radars could detect low friction spots from a longer range at any weather conditions[1]. Moreover, the influence of vehicle body vibrations when using optical methods is quite considerable [9]. Despite the fact that radars and ultrasonic systems are widely used mainly for obstacle detection and avoidance, road profiling, there is no information on the practical fusion of these technologies in the automotive industry for

terrain recognition. Our approach was to use information from two sensors to estimate the state of a system that would be better than using each sensor individually. The beginning was by the development of radar and ultrasonic sensor system that can measure the backscattered radar and ultrasonic signals. Land clutter models were studied to understand the dependence of the backscattered signals on the surface properties (roughness, material, etc.) and the system configuration (grazing angle, height of the antenna, beamwidth, etc.). This aided the optimization of the system configuration for the finest surface identification.

Radar and ultrasonic measurement results vary within a certain range which is distinctive for different sort of surfaces. A complete system comprising algorithms and classification methods was proposed to achieve the following:

1. Characterize and accumulate the collected data from different on-road and off-road surface into a database.
2. Extract the features and develop the algorithms to be used for classification.

3. Optimize the classification methods in terms of classification accuracy and complexity costs.

To obtain a reliable system, a study has been conducted on the influence of weather conditions and vehicle movement on measurements. The basic theoretical analysis was incorporated with outdoor experimental studies and computer modelling. We have analysed the performance of several common classifiers where both radar and ultrasonic backscattered signals features are inputted. All measurements have been done with real surfaces in different weather conditions.

1.5 Thesis Outline

Chapter two introduces the fundamentals of monostatic radars and ultrasonic sensors. It follows by a brief review of the use of radar and ultrasonic in the automotive industry. The second part explains the fundamentals of pattern recognition and the classification algorithms used in the road surface identification. Chapter three introduces the principles of cross section and land surface clutter. It shows the effect of surface roughness, material, and other factors on the reflectivity. Chapter four describes the outdoor experimental setup and the methods used for data collection. It also shows the effect of the weather conditions and the system configuration on the measurements. In chapter five, the classification algorithms are shown including the set up of a database and extracting the features out of the row data available from experiments. The performance of the classification algorithms is optimized as shown in that chapter. Chapter 6 shows the results obtained from radar, ultrasonic sensor, and fusion of both. An evaluation of the methods used to detect road and off-road surfaces is shown, in addition to the influence of exhaust air on the measurements. Finally, chapter seven draws the conclusions and summary of the work completed. Future research and proposals are also discussed in chapter seven.

Chapter 2

Literature Review

2.1   Introduction

A general introduction about radar and ultrasonic sensors is given in the first part of this chapter. It follows up with a literature survey of radars and ultrasonic sensors used for the surface discrimination in and out of the automotive industry. This includes the automotive environment and the surfaces of interest. The chapter also highlights the characteristics of surfaces affecting the reflectivity of the sensors. The second part introduces pattern recognition and highlights the classification methods used. Relevant research undertaken previously regarding surface classification is reviewed, and the factors affecting the performance of the classification system are explained.

2.2 Principles of Radar Technology

Radar systems transmit electromagnetic waves (EM) from a transmitting antenna, and wait for the reflected wave to be received by the receiving antenna. The reflected wave is analysed on the receiver side to gather information about the target hit by the EM wave. Any radar system consists of major subsystems: transmitter, receiver, and signal processor. There are three configurations of radar systems: monostatic, bistatic, and multistatic. Monostatic radars have the transmitter and receiver collocated and can use the same antenna for both

transmitting and receiving. Bistatic radars have their transmitter and receiver separated by a distance that can be compared with maximum range of the target. On the other hand, multistatic radars contain multiple monostatic or bistatic radar components with a shared area of coverage. Figure 2.1[10] shows the difference between a monostatic and a bistatic radar.

backscattered energy

reflected energy

Tx/Rx

Tx

baseline

Rx
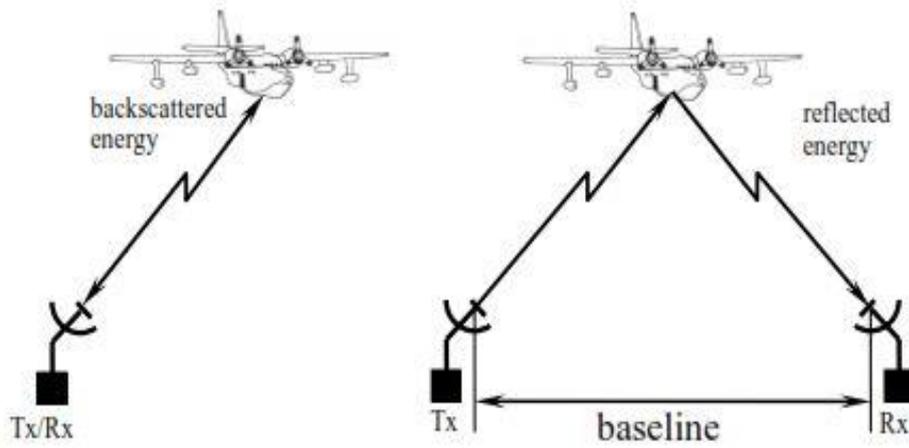
Figure 2.1 Monostatic and Bistatic radar[10].

In the monostatic case, the energy is backscattered to the receiver after it hits the target, while it is reflected to the receiver in the case of bistatic radar. As road surface identification sensors are to be mounted on the vehicle, both transmitter and receiver should be collocated to provide a system small enough to be integrated within the car. The radar system to be developed in this research is monostatic radar processing the backscattered signal from different surfaces. If the transmitter transmits power p_t through an antenna of gain G, the received power p_r is given by the monostatic radar equation:

$$p_r = \frac{p_t G^2 \lambda^2 \sigma}{(4\pi)^3 r^4} \tag{2.1}$$

where σ is the radar cross section of the target, λ is the wavelength of the signal, and r is the distance to the target. As the range of the target *r* increases, p_r decreases. This means that a high power transmitter and a low-noise receiver is needed for long range targets [11]. This equation and its relationship with the radar cross section will be discussed in details in the next chapter.

2.2.1 Microwave Signal Backscattering

There are many types of radars characterized by different major features discussed in [12]. The radar system developed for road surface identification is Pulse Compression radar that uses long pulse with internal frequency modulation. The developed radar system is designed to analyse the change in polarisation of electromagnetic wave which is reflected from the road surface. This change is referred to depolarisation, and is mainly governed by set of parameters affected by the dielectric constant of the surface material and the surface roughness. In general, the properties of a surface can be described by scattering matrix of the scattering target [13]:

$$S= \begin{bmatrix} S_{vv} & S_{vh} \\ S_{hv} & S_{hh} \end{bmatrix} \qquad (2.2)$$

where the first index of matrix elements refers to the polarisation of the transmitted signal, and the second index refers to the polarisation of the received signal. Each scattering amplitude S_{ij} is a complex quantity which comprises a magnitude $|S_{ij}|$ and a phase angle ϕ_{ij}:

$$S_{ij} = |S_{ij}|.e^{i\phi_{ij}} \qquad , i, j = v \text{ or } h. \qquad (2.3)$$

The backscattered polarised signal magnitudes ($|S_{vv}|$, $|S_{hh}|$, $|S_{vh}|$ and $|S_{hv}|$) are Rayleigh-distributed and related to the surface geometrical and dielectric properties. For most surfaces, the phase angles (ϕ_{vv}, ϕ_{hh}, ϕ_{vh} and ϕ_{hv}) and cross polarised phase angles ($\phi_{x1} = \phi_{hv} - \phi_{vv}$ and $\phi_{x2} = \phi_{vh} - \phi_{vv}$) are uniformly distributed over $[-\pi, \pi]$ and they are harder to extract surface information from. The co-polarised phase angle ($\phi_{c} = \phi_{hh} - \phi_{vv}$) is target-dependent and depends on the sensor parameters (incidence angle, wavelength) and on the surface parameters (roughness, dielectric constant). Microwave radiation is very sensitive to the presence of water in the medium through which it passes. Thus, the amplitude and phase of a wave reflected from a road contains information about water, snow and ice accumulation. The number of influential physical parameters on radar responses of surfaces such as surface dielectric constant, roughness, wetness, density, surface cover, etc. is rather large, which makes the effort of generating a comprehensive data set rather difficult.

2.2.2 Surface Recognition using radars

Since the invention of radars, it has been used for several applications including: military, remote sensing, air-traffic control, etc. The use of radars for surface roughness recognition and detection of low- friction spots caused by water, ice, or snow has been investigated in many studies. The most common method of surface identification is comparing backscattered signals at different polarizations at different incidence angles. The backscattering properties of dry, wet , and icy asphalt has been studied in [14] [15] at frequencies of 24 GHz and 77 GHz. Low friction spots were detected, in laboratory conditions, using backscattering ratios of different polarizations. Results showed that road covers like snow, ice, and water change the scattering properties of the surface. Detection of

snow on asphalt was shown by [16] using 95 GHz bistatic narrow-beam radar measuring different polarizations of reflected signal. The system uses stationary antennas located at 1.5 meters above the surface. In [17] a coherent polarimeter system operating at 24 GHz which can detect ice on the road surface is described. The system consists of two transceivers with orthogonal polarization and uses velocity gates to acquire road scattering only. Polarimetric monostatic 94 GHz radar system is reported in [18] for characterizing the backscattering of asphalt surface: dry, covered with ice or water. The backscattering coefficients and phase-difference statistics have been theoretically calculated and measured over a wide range of incidence angles. Authors of [19] showed a 10 GHz bistatic radar used to acquire the characteristics of the surface (dielectric constant and surface roughness) in laboratory conditions. Scattered power was measured at different incident angles to obtain the backscattering coefficient. Information was collected about the surface roughness and physical dielectric constant of the surface by studying the relationship between the backscattering coefficient and the incident angles. In addition, a 24 GHz automotive radar for the detection of road condition in laboratory and real road conditions was presented in [20]. The radar uses eight transceivers which illuminate the surface simultaneously at different incidence angles. Difference in cross polarization properties of dry, wet and icy asphalt was shown in the preliminary results. Multifrequency radar was used in [21] for the detection of snow and ice on roads in laboratory conditions. [22] Showed an excellent discrimination of four laboratory road surface classes: dry, wet, snowy, and icy using a 94 GHz dual-channel polarimetric radiometer data. [23] and [24] provided scattering measurements from different laboratory road surface at 76 and 24 GHz respectively. Mueller matrix was used to recognize the different road conditions[13]. Results showed that road's scattering properties is changed by the cover (water, ice, snow). A 61 GHz bistatic radar for road condition recognition was reported in [25].This system measures all four polarization

combinations and compares the amplitudes and phases of the responses. It was able to distinguish asphalt and cobblestone pavements. Another bistatic radar was used by [26] to measure road conditions. The system used an active bistatic radar and passive radiometer operating at 35 GHz and 90 GHz. In [27] the specular reflection from roads has been measured and the polarization properties were used to detect ice layers on road surfaces. The system used in [27] is based on 76 GHz bistatic radar. A 2.45 GHz road surveillance radar has been able to measure the thickness of the water layer and snow layer on the road [28]. Using 24 GHz Doppler radar, authors of [29] found out that different surface properties (dry road, frozen, wet and crusty snow) causes change in reflected signal amplitude, but there is no considerable change in the shape of the signal spectrum. Ground penetrating radars are used by in [30] [31] to monitor soil water content and measure the dielectric constant at frequencies of 250 MHz, 500MHz, and 1000 MHz In addition to surface recognition, radars are reported to be capable of detecting lanes [32] and advance paths [33]. Other automotive applications like collision avoidance are also provided by radar systems as reported in [34] . Some patents on the use of radars for surface recognition are presented in [35] [36] [37] [38].

Nevertheless, the survey shows that most studies were conducted in stationary laboratory conditions. In addition, only few surfaces have been investigated. The off-road surfaces were ignored as the focus of the studies was to differentiate wet, icy, dry and snowy asphalt. This leaves a space for further investigation of the reflected signals from different road and off-road surfaces under real conditions.

2.3 Principles of Ultrasonic Technology

Ultrasonic sensing or Sonar propagates acoustic energy of frequency higher than the normal hearing. Like the radars, ultrasonic sensors listens to the signal reflected back from a target. In this case, the reflected signal is called echo. Ultrasonic sensing is popular for many features like cheap cost, light weight, low power consumption, and low complexity [39]. It is used by Robots to measure range to objects. The object range is supposed to be proportional to the echo travel time as the speed of sound is known. In addition, ultrasonic sensors have different purposes: Obstacle avoidance, Sonar mapping, and object recognition. The interest of this research is to recognize surfaces through the ultrasonic echo. Many factors influence the waveform and amplitude of ultrasonic echo. The angle of incidence of ultrasonic wave, the bandwidth, the roughness, and the texture of the surfaces are important parameters that must be considered when dealing with scattering from uniform surfaces. The roughness of the surface is the relationship between the surface irregularities and the wavelength of the incident signal. The larger the electrical length of the irregularities, the rougher is the surface.

2.3.1 Theory of ultrasonic signal backscattering

The mean intensity of the echo signal is defined by two terms in the equation (2.4) [40]:

$$I_{bS} = \rho_S I_o \, + \, \langle I_d \rangle \qquad\qquad (2.4)$$

where I_o represents the coherent component, and $\langle I_d \rangle$ represnts the non-coherent component. The coherent component, also known as the specular component, is the intensity which

would have been reflected from an equivalent smooth surface. ρ_S is the equivalent reflection coefficient for scattering in the specular direction. The non-coherent component, also known as the diffuse component, is the mathematical expectation to account for the random nature of the reflections from rough surfaces. The coherent component dominates for smooth surfaces, while the non-coherent dominates for rough surfaces. The specular and diffuse reflection from smooth and rough surface will be discussed in depth in the next chapter. Both components of the echo signal depend on the surface properties, angle of incidence, and directivity pattern. Typically, ultrasonic signatures are noisy due to the interferences caused by air currents and other reflections [41]. Consequently, probing the surface at a large distance will bring uncertainty in the measurements due to the interference signal caused by the scattering of ultrasound by air fluctuations. Nevertheless, the analysis of the reflected echo can give reliable information about the surface but with restriction of the sonar range (to provide sufficient reflected power).

The amplitude of the ultrasonic signal received at an angle θ to the target is given by [42]:

$$A(r) = A_o \, G(r) \, \varrho(\theta) \, \rho_b \qquad (2.5)$$

where A_o is a constant, $G(r)$ is the transmission gain function, $\varrho(\theta)$ is the directivity of the receiver and transmitter pair at a grazing angle θ, and ρ_b is the scattering coefficient characterizing the texture and roughness of the surface. The ultrasonic emitter is modelled as a circular piston surface vibrating in an infinite planar baffle. The emitted pressure field forms a main lobe surrounded by side lobes when a > λ , where a is the radius of the circular piston surface [39].This beam is described by the directivity pattern in the far field (range greater than a^2>λ) and is given by [43]:

$$\varrho(\theta) = \left| \frac{2 \, J_1[k \, a \, .\sin(\theta - \theta_1)]}{k \, a \, .\sin(\theta - \theta_1)} \right| \tag{2.6}$$

where J_1 is the first order Bessel function of the first kind produced by the circular aperture of the piston surface. The wave number k is the ratio between the angular frequency of ultrasonic waves and the speed of ultrasonic waves in air. It is given by the following equation:

$$\text{k} = 2 \, \pi \, {F_0}/{c_{air}} \tag{2.7}$$

where F_0 is the transmitted signal frequency, and c_{air} is the speed of sound in air at sea level air density given by:

$$c_{air} = 20.0457 \sqrt{T_c \, + \, 273.16} \ \text{m/s} \tag{2.8}$$

where T_c is the Celsius temperature. A better estimate of the speed of sound can be done if humidity is known [39].

The backscattering coefficient, given by ρ_b in (2.5), can be derived by looking at the physics of propagation of the ultrasonic waves and the surface model. Road and off-road surface are considered rough surfaces that can be modelled using Gaussian distribution of surface's height. The incoherent component of echo dominates over all the incidence angles, and the backscattering coefficient is given by [42]:

$$\langle |\rho_b| \rangle = \sqrt{\frac{\pi}{S} \frac{\eta \, \rho_s}{2k \, cos^3 \, \theta}} \ e^{- \, tan^2 \theta \, . \, {\eta^2}/{2}} \tag{2.9}$$

where S is the area of reflecting surface, η is the roughness parameter, and ρ_s is the equivalent reflection coefficient for scattering in the specular direction. In chapter 4, a simulation was set based on equation (2.9) to show the dependence of backscattering coefficient of the ultrasonic sensor on the grazing angle for surfaces of different roughness.

The transmission gain function, $G(r)$ in equation (2.5), attenuates with distance to the surface r. $G(r)$ depends on two factors: the inverse of second power propagation loss and another exponential factor caused by the absorption in the propagation medium (air damping). The transmission gain function $G(r)$ is given by:

$$G(r) = G_0 \frac{W(r)}{r^2} e^{-2\alpha r} \tag{2.10}$$

where G_0 is a coefficient depending on the reflector properties, and α is the atmospheric ultrasound attenuation coefficient. Values of α are provided in the tables of [44], and their dependence on the signal frequency, temperature and humidity are given. *W(r)* is the increase of the insonified area with respect to distance from the transducer. It represents the length of a swathe sector at a distance r, and can be found by solving the equation of the intersection of the plane (surface) with the cone representing the main lobe. Figure 2.2 shows the insonified area (meters) with respect to distance (meters). The intersection of a cone and a plane surface can be elliptical, hyperbolical or parabolic shape depending on the grazing angle and beamwidth. Illuminated footprint shown in figure 2.2, simulated in Microwave Integrated System Laboratory (MISL), is based on beamwidth of $45°$ and height of ultrasonic

transducer above the surface of 0.4 meters. As the grazing angle decreases, the insonified area increases which leads to more beam spreading (i.e., distance damping).
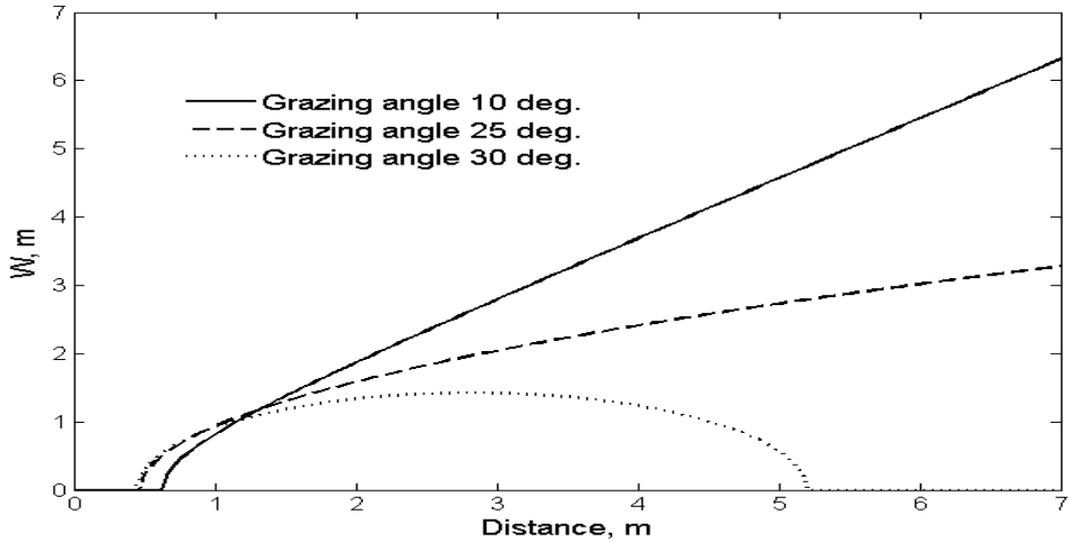
Figure 2.2 Intersection of a cone and a plane surface.

The analysis of expressions (2.5)-(2.10) allows drawing conclusions about the features of the signals reflected from various surfaces.

2.3.2 Surface Recognition using Ultrasonic sensors

Many studies investigated the use of ultrasonic signals for surface recognition according to its roughness and textures. Some of the investigations in the field of mobile robot applications are shown in [45], [42], [46] and [41]. An application of a surface classification method called ENDURA (Energy- Duration-Range) is explained in [47] and [48]. The method uses the echo intensity in terms of its energy content, duration, and range to classify surfaces according to its roughness. Detailed explanation will be followed in section 2.4.2. In [49], sonar performance in distinguishing between surfaces has been explored.   The authors used an extension of the Kirchhoff approximation method, describing the scattering

of acoustic wave, to model random rough surfaces. These models were used to derive a continuous transmission sonar signature which was compared to experimental measurements in order to differentiate between hard smooth floors, carpets, asphalts and surfaces made up of tiles. The results of correct recognition achieved were 92 % and 94 % using K-nearest neighbors (KNN) algorithm. A neural network has been trained in [50] to differentiate wood, carpet, curtain, ceiling and water covered surfaces. The system used Time Delay Spectrometry (TDS) to filter reflections from an acoustic sensor. TDS is a special technique used to separate and measure spectral response of an acoustic system. The spectral coefficients of each 64 frequency points were averaged so that the magnitude of the signal was separated into 64 equal bands. The energy in each of the 64 bands were normalized to create a feature vector with a dimension of 64. This feature vector was input to the neural network to achieve a classification accuracy of 96%-98%. In addition, authors of [45] developed a model combining a transducer model, acoustic reflection model, and a surface geometry model to predict the echo component. The quality of the 11 features extracted were experimentally measured and reduced to 5 features (energy distribution, range of angles, sum of amplitudes, end angle and amplitude of peak) to achieve 99.73% of correct recognition for 12 indoor surfaces with different profiles. Three sonar signature features were extracted in [41, 46] using attributes of a scattering model to distinguish between asphalt, grass, gravel, plastic, and carpet. This was achieved using different classification algorithms, such as K nearest neighbour algorithm, Radial Basis Functions (RBF) and Multilayer Perceptron (MLP) which will be discussed in the next section. In [46], an ultrasonic scattering mathematical model was derived, based on the Kirchhoff Approximation Method (KAM) [58], to calculate FFT images from the ultrasound continuous transmission frequency modulated sensor. The sonar images were estimated for rough and smooth surfaces and the final model consists of an equation of sonar image

comprising both rough and smooth component. Three features were particularly extracted: the rough and smooth components of the overall FFT of the scattered signal, and the range of the detected reflections in the FFT image. In order to extract the rough component, the rough surface was split into k elementary much smoother surfaces. The rough component is the sum of all reflections from the elementary surfaces, while the smooth component is the reflection from the rough surface as a whole. The use of KNN provided high probability of correct recognition when number of neighbours K didn't exceed 12. On the other hand, RBF network consisting of 9 hidden nodes was trained. EM (Expectation- Maximisation) algorithm was used to acquire the position of the centres of each radial basis function, while pseudo-inverse method was used to obtain the hidden to output weights. RBF showed better results than KNN and overcame the computer storage KNN disadvantages. In addition to RBF, a two layer perceptron network with four hidden nodes was trained using the scaled conjugate gradient. The network consisting of three inputs and five outputs (five surfaces) was trained using a twenty data point train set for each class. According to [46], even though MLP is computationally intensive, it provided the best results for rough surface classification. In [51] the properties of backscattered acoustic waves has been studied as a function of the transducer geometry and the roughness of the surfaces. A numerical evaluation was presented to evaluate the influence of surface roughness, grazing angle, range and transducer type on the reflected signal. A boom-mounted microphone was tapped on different floor materials (contact sensing) in [52]. A windowed power spectrum of the acoustic signature arising from the microphone contact is then compared to one of a family of prototypical signatures generated statistically from the same material (metal, wood, cement, plastic, glass). The technique is used to classify the floor type. It involves limited computational expense, and performs very well. The "acoustic transfer function" between an echo from a reference plane and that from object to be tested is calculated in [53] to

classify object surfaces. The acoustic transfer function of an object determines how the transmitted waveform is changed to give the echo waveform. Finally, several patents have described an ultrasonic ground speedometer utilizing Doppler effect[54]. The analysis of reflectivity from different surfaces at different signal wavelengths showed that the surface roughness can be defined by analysing reflectivity at different frequencies.

Most of the studies on ultrasonic signals were carried out in stationary laboratory conditions. Moreover, the influence of external factors, like weather conditions, was not highlighted. A lack of research on the fusion of sensors (radar and ultrasonic) was also reported. This leaves a space for further investigation of the reflected ultrasonic and radar signals from different road and off-road surfaces under real conditions.

2.3.3 Classification algorithms for surface Recognition

Surface classification has been investigated in many studies. Classification algorithms are used in remote sensing applications to classify land covers and surface types. Identification and classification of different types of road pavements was reported in [59, 60, 61] by analyzing the tyre noise profile. Two classifiers were used in this investigation, and correct recognition of 95% between 7 surfaces was achieved using Neural Network classifier [59,60]. Nevertheless, classification based on the tyre noise analysis can be difficult in real road conditions due to the dependence of the acoustic signal on vehicle speed and weather conditions. Different kinds of surfaces were discriminated in [32, 36, 37, 40, 41] using ultrasonic sensors. KNN method was used in [40] to classify eight surfaces. The surfaces were divided into two hyper classes with random and periodic textures. A probability of

recognition of 92 to 99 % was reached using this method. Surface's classification using radar reflections was investigated in [12, 22, 14 ,74]. Microwave signals backscattered from dry, wet, icy and snowy asphalt have been studied in [21] at two transmitter frequencies and under laboratory conditions. Although the performance was not provided, it was claimed that Maximum Likelihood Estimation provides reliable classification when using. Furthermore, the use of supervised learning algorithms for surface classification was reported in [23]. The analysis implemented the use of the neural networks, statistical classifiers and polynomial classifier without showing any numerical results. Differentiation between dry, wet, icy and snowy asphalt in laboratory conditions was established by a 94 GHz dual-channel polarimetric radiometer in [22,74]. The four road surface classes were classified, with a Classification accuracy reaching 94%-95%, using a Bayesian classification method based on principal component analysis. The first principal component was characterized by the air temperature, while the second was indicated by the mean of the brightness temperature value of vertical and horizontal polarization.

The survey shows that there is a lack of research on the efficiency of the statistical classification algorithms. Furthermore, most studies of the classification algorithms were conducted on the ultrasonic data. The performance of classification algorithms on radar signals reflected from different surfaces need to be further investigated.

2.4 Automotive environments

2.4.1 Road and off-road Surfaces

Paved roads provide a mean of transportation for vehicles and pedestrians. The structure of the roads is built to prevent vehicles from sinking into the surfaces. The road structure consists of two types: flexible and rigid. Flexible roads are more common in the road building industry as rigid roads are more complex and require special equipment [55]. Before building a road, the ground is dug-out down to a certain depth. The structure and materials of a flexible road is shown in Appendix A. Sub-grade is the ground dug out below the road, while capping is a layer added above for protection. The four layers above the capping are the layers of the road (Sub-Base, Base, Binder Course, and Surface course). These upper layers constitute of strong materials to lower the stress transmitted through the vehicles. Artificial compaction takes places during the construction of each layer to assure that gaps are filled and materials wouldn't compact down with time. Since the radar and ultrasonic sensors cannot penetrate the layers below the surface course (45 mm to 105 mm thickness), we have concentrated on the surface of the road. In this research investigation has been carried out on two types of road surfaces: rough and smooth. The terms smooth asphalt and rough asphalt used in the next chapters refer to a smooth and rough road surfaces that constitutes of the above mentioned layers. Apart from the road surfaces, few off-road surfaces are considered to be of interest as any vehicle would encounter these surfaces on an off-road trip: sand, gravel, and grass.

1.   **Asphalt:** This liquid or semi-solid black material is a form of petroleum. Asphalt, also known as bitumen, can be found in natural deposits. It is mainly used to construct the roads made of asphalt concrete[56].

2. Sand: This natural surface consists of rocks and mineral particles. Sand is a textural class of soil as the latter contains more than 85% of sand [57]. Sand types vary according to location and source of rocks, but most commonly it consists of silica in the form of quartz and calcium carbonate. The diameter of sand particles ranges from 0.625 mm to 2 mm. This range is finer than gravel but coarser than silt.

3. Gravel: This natural material is commonly used in surfacing roadways especially in rural areas. It consists of rocks fragments of different sizes. Gravel is categorized into granular gravel and pebble gravel according to its fragment's size. Granular gravel has particles of diameter of 2-4mm, while pebble gravel has particles of diameter 4-64 mm. There are many types of gravel with larger particle diameters, but the gravel investigated in this research is pebble gravel of average 50 mm particle's diameter.

4. Grass: or graminoids consists of plants with narrow leaves growing on the surface base. There are three families of grass: true grass (Poaceae), sedges (Cyperaceae), and rushes (Juncaceae). Each of the families has different heights and different ecology. The grass of interest in this research is the true grass that can be found on the roadsides and vegetation areas. True grass or Poaceae has hollow stem and leaves with parallel veins. The average height of the dense mown grass investigated is about 5 cm.

2.4.2 Characterization of surfaces

Analyzing the different types of automotive environments allows the integration of the surfaces into one analytical framework. Microwave and ultrasonic reflections differ for different materials and different roughness surface. It is essential to establish a relationship

between the features of the reflected signals and the properties of the surfaces, enabling the differentiation, classification and identification of the surfaces.

As electromagnetic wave reflections are imposed to polarization changes that characterize the properties of the surfaces (mainly material) demonstrated as a scattering matrix, ultrasonic backscattered signals are parameterized in terms of energy which is a function of surface roughness. Previous sonar system only relied on the extraction of the time of flight (TOF) information from the ultrasonic echo. ENDURA(Energy and DURation) method reported in [47] and [48] showed a significant improvement over the previous systems by applying the following methodology:

1.   Energy and duration templates were derived for smooth, moderately rough, and rough surfaces. The templates are based on the forward model which is based on the Kirchhoff Approximation method (KAM) explained in [58]

The echo energy describes the energy content of the reflected signal and is defined as:

$$\text{P} = \int S^2(t) \; dt \qquad (2.11)$$

where $S^2(t)$ is the echo from a particular structure. The echo duration is derived from the second moment of the echo and it is given by:

$$\text{D} = \left(\frac{\int (t-\mu)^2 S^2(t) \; dt}{\int S^2(t) \; dt} \right)^{1/2} \qquad (2.12)$$

where

$$\mu = \frac{\int t \, S^2(t) \, dt}{\int S^2(t) \, dt} \qquad (2.13)$$

Using the forward model, equation (2.11) and (2.12), energy and duration template are derived for each surface type. Figure 2.3 [48] shows the energy and duration templates used in ENDURA. In the smooth surface template, energy hits maximum at the normal incidence. It decreases as the incident angle θ increases. The smooth surface cannot be detected beyond $\pm\theta_o$ from the normal incidence, where θ_o is the half-beamwidth angle. The echo duration is approximately constant within θ_o since the echo's shape is similar to the transmitted pulse. In the moderately rough template, the coherent component of the echo dominates in echo energy within $\pm\theta_o$ from the normal incidence. The incoherent component dominates at larger incidence angles. Since both coherent and incoherent components are significant, no simple analytical description can be obtained for the echo duration. Nevertheless, it can be seen that the echo duration curve has two distinct regions. In the rough surface template, echo energy hits the maximum at normal incidence then decreases slowly. A slow increase is also shown in the echo duration curve as the angular deviation increases.

2.   Set of experiments is conducted on smooth, moderately rough and rough surfaces. Ensemble averages of the experimental echo energy and duration are compared to those estimated from the templates. A measure of mismatch given by Rosenfeld-Kak [59] is used to compare the experimental and theoretical results. The measure of mismatch is given by:

$$M = \int (f - g)^2 \, d\theta \qquad (2.14)$$

Where f is the template and g is the observed signal.

3. The normalized version of the mismatch measure represents a normalized matched filter.

Six filters are used, in total, to represent the mismatch between echo energy and echo duration for the 3 surface types. M^P and M^D are the filters of the energy and duration, respectively. $\Pi_i(\theta)$ and $\Omega_i(\theta)$ are the outputs of the energy and duration filters (i=1, 2, 3). The combination of the outputs is given by $\Gamma(\theta)$:

$$\Gamma(\theta) = \left[a\ \Omega^2(\theta) + (1-a)\ \Pi^2(\theta)\right]^{1/2} \tag{2.15}$$

where a is a weight factor that can be chosen as a function of range.
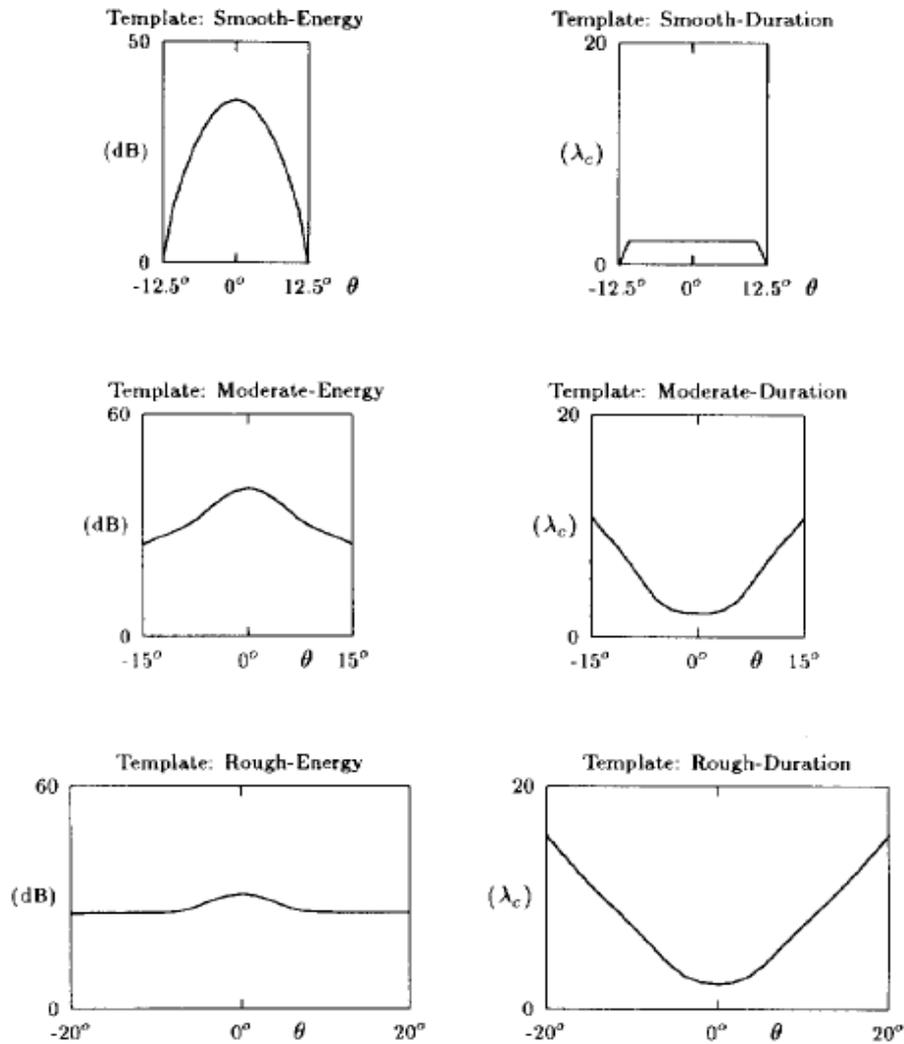
Figure 2.3 Templates of energy and duration used in ENDURA for smooth, moderately rough, and rough surfaces[48].

4. A decision of the type of the surface depends on the filter that results with smallest output (highest match). Furthermore, the centroid of the signal is calculated and the distance to the surface is estimated by the TOF measurement of the centroid.

The application of ENDURA provides reliable detection of the types and the location of the surfaces. It gives the insight of the parameters used in this research to acquire new features that characterize roughness of the surface (further explanation in chapter 5).

2.5 Pattern Recognition

The goal of this research is to provide surface classification using multisource remote sensing data. The challenge is to extract useful information from the data in order to characterize the signals and differentiate these signals according to the surface they were reflected at. Pattern recognition is the study of algorithms for recognizing the patterns of data and classifying it according to specific properties. However, this research doesn't intend to cover all aspects of pattern recognition. This section will cover the principles of machine learning, feature extraction, and a number of classification algorithms that we used for the surface identification.

The data set consists of examples or instances. Each instance is described by variables or features that can be numerical or discrete. Instances with similar or common features are classified by the machine learning algorithms into a class. In order to determine the unknown class of new examples, machine learning algorithms will be trained to approximate a representative model per class of data. There are two methods or categories for undertaking this training. The first method is called supervised learning and it is used when the data set

is labelled, i.e., the class of each instance in the data set is known to the machine. The other method is unsupervised learning, and it is applied when the class of the instances is not known. In the latter case, the algorithms try to find hidden structure in the unlabelled data in order to allocate it into a cluster. As we know the nature of the surfaces we are experimenting, all the instances in the data set are labelled. Therefore, the focus will be on classification algorithms. Figure 2.4 shows the whole pattern recognition system used for classification. After acquiring data, instances which represent a set of available recorded signal features are collected in a database. This database is used to train the model where each instance is included in a particular class. Classification later probes the features of the new unlabelled instances against
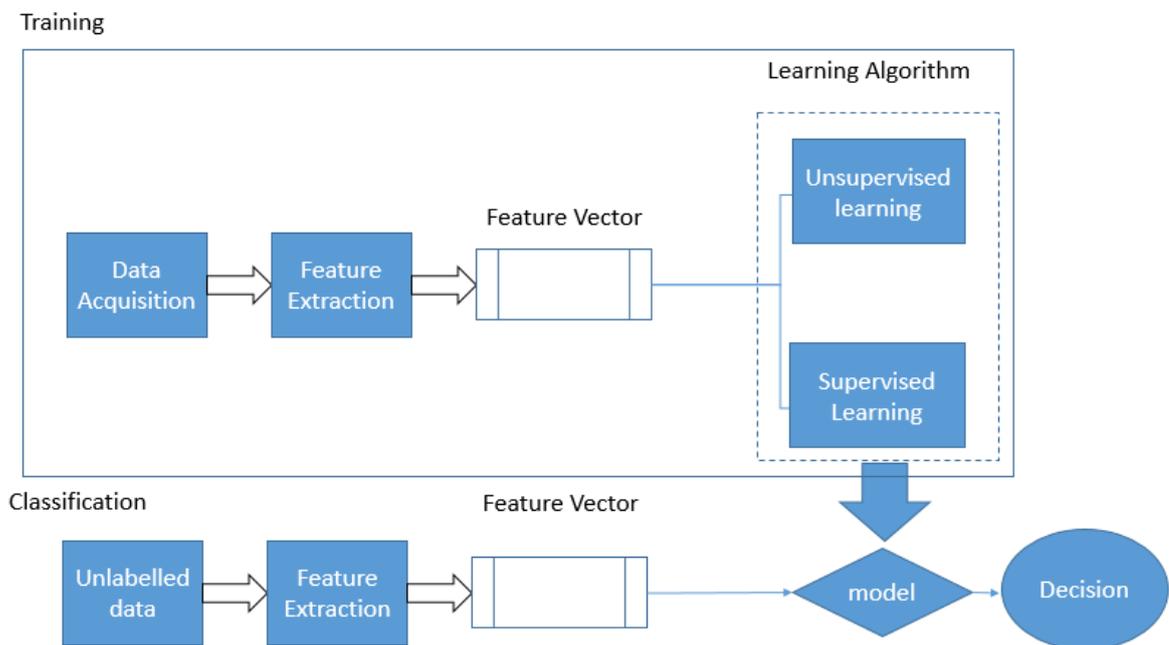
Figure 2.4 Training and classification phases of a pattern recognition system.

the database. The decision on affiliation of each measurement to a particular class is made. The Feature selection phase shown in figure 2.4 is expected to produce a reliable signature from each of the data samples available for training.

2.5.1 Feature Selection

Given a data distribution, a pool of features can be estimated. Within this exhaustive pool of features it is expected [61,119] that only a subset of features will be considered as reliable candidates for integration within a classification framework. Variety of methods from literature can be applied to select the features like Information Gain, ReliefF, Gain Ratio, etc. in a backward feature selection or forward feature selection [61,119,120,121] . Both methods were applied, in this proposed work, on the observations before it was provided to the system to train the models. The methods and the features selected will be discussed in details in Chapter 5.

After feature selection, instances are described by their feature vector, x = $[x_1, x_2, \ldots, x_N]$ where N is the number of features extracted. Figure 2.5 shows a 2 dimensional feature space for a supervised learning example in which feature vectors were fed into the learning algorithm. The similarity between the blue points (or red) refers to the proximity of their features in the feature space. It is very rare to have a feature space that can separate the classes correctly without errors especially when it is dealing with more than two classes.
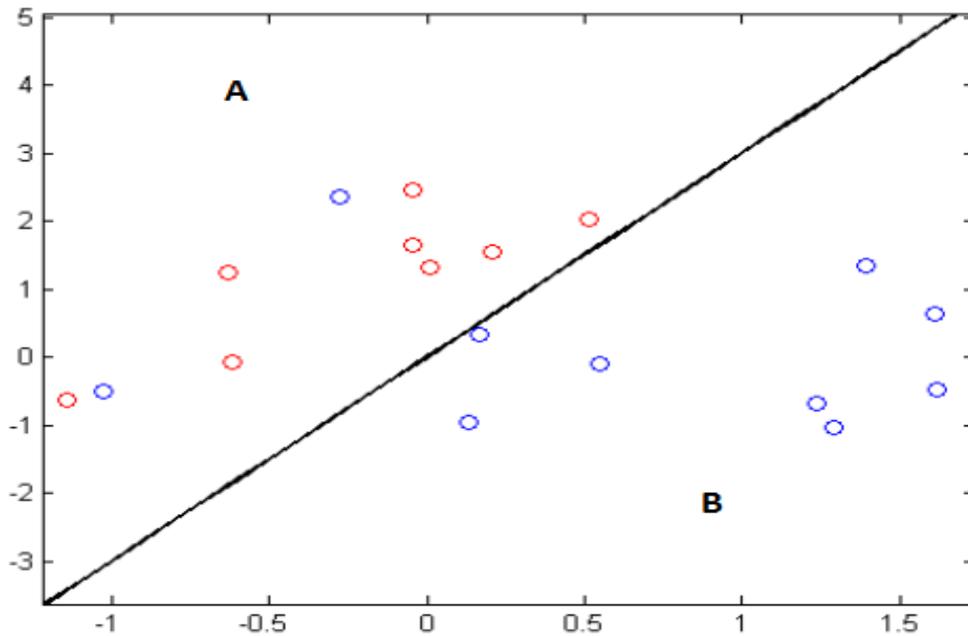
Figure 2.5 Two dimensional feature space

2.5.2 Algorithms of Classification

The work proposed in this thesis is valuated against some of the state of arts method which is based on classification algorithms that can be categorised as parametric and nonparametric. In parametric methods we assume that the sample is drawn from some distribution that obeys a known model, for example, Gaussian. The advantage of the parametric approach is that the model is defined up to a fixed number of parameters. In nonparametric methods all we assume is that similar inputs have similar outputs. The difference is that parametric methods make more assumptions of the data probability distribution. The number of parameters in non-parametric method grows with the training data, while that of parametric methods is fixed. The most widely used non-parametric method is the k-Nearest Neighbors (KNN) Further discussion about parametric and

nonparametric methods is provided in [63]. At the initial stage of this research, a study was conducted on many classifiers using our data set (Support Vector Machines SVM, Naïve Bayes, Decision Tree learning, Minimum Distance classifier and K nearest Neighbour, etc.). The four classifiers which offered an effective way of classification in terms of accuracy, time complexity and space complexity are described in the next section.

K Nearest Neighbours

KNN is considered one of the simplest machine learning algorithms. It is an instance- based nonparametric classifier which classifies data based on the distance between vectors or neighbours [64]. Any new unlabelled test sample belongs to the class which has the closest k neighbors where k is a positive integer. With too small k single instances have a large effect, which leads to the shortage of smoothing over the noise. With a big value of k, many instances contribute to the estimate of a point. Another factor affecting the performance of KNN is the distance function. The distance function (metric) is a function that describes the distance between the test point feature vector and the data set features vector. It represents the closeness of the test point to the classes. The most common distance function used is the Euclidean function. Nevertheless, many other distance functions were used and compared in this work (Euclidean, Manhattan, Chebyshev, and Mahalanobis distance). The distance functions and number of neighbors (k) chosen for KNN are discussed in details in chapter 5. Although KNN is simple and easy to implement, it has few disadvantages such as [65]:

a) Computing distance of each test point to all training sample yields a high computation cost.

b) A large memory is required.

c) Determining the value of parameter *k*.

Nevertheless, the classification performance by any other sophisticated classification method can only be better by a factor of two as compared to KNN [66]. The KNN algorithm is implemented in Appendix B.

Minimum Distance Classification

Minimum distance classification (MDC) is the second explored classifier in this research. The parametric method is an adaptation of k-means method. The concept of k-means is based on the partition of the observations into spherical clusters in which each observation belongs to the cluster with the nearest mean. MDC applies the same concept as KNN but it looks for the closest class's mean to the test point instead of the closest k neighbors. The mean is a representative value for the class, defining usually the centre of all the sample vectors in that class. MDC is considered a simple method and it works well when the distance between the mean is larger than the spread of training data of each class around its mean. An important factor affecting the performance of MDC is the distance function used. The MDC algorithm is implemented in Appendix B.

Gaussian Distribution Model

The Gaussian or normal distribution is a common continuous probability distribution. The Maximum likelihood Estimation (MLE) method is a method used for estimating the values of parameters of a Gaussian Distribution model. It estimates the parameters that maximize

the agreement of the model with the observed data given a chosen probability distribution model. The MLE uses the samples to provide estimates of the parameters mean μ and variance σ^2. The MLE method, implemented in Appendix B, produces two MLE estimators (mean and variance). These two MLE estimators correspond to the value of a parameter vector θ that maximizes the likelihood of the selected model with the observed data. After estimating the MLEs, it calculates the probability density function (pdf) of the test point using the normal distribution with mean and standard deviation. The normal pdf for a test point x is given by:

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (2.16)$$

where μ is the mean of the class, and σ is the standard deviation. The new unlabeled test sample belongs to the class in which it achieves the highest probability providing the class's two MLE estimators.

Multilayer Perceptron

The second explored nonparametric method is Multilayer Perceptron (MLP). MLP is considered one type of Artificial Neural Network models. ANN uses a number of interconnected processing elements called neurons to learn and adapt in response to external inputs [68]. As shown in figure 2.6, each neuron generates a single output out of the received number of inputs. If the input vector to the neuron is given by $x = [x_1, x_2, \ldots, x_n]$, each input is associated with a weight $(w_1, \ldots, w_n)$. The total output of the neuron is given by calculating the combination of the inputs and the associated weights:

$$y = \sum_{k=1}^{n} x_k w_k \qquad (2.17)$$

The output is usually compared with a specified threshold θ. A transfer function is also associated with the neuron and it sets the output to a range of values, within a range inherent to the transfer function properties.
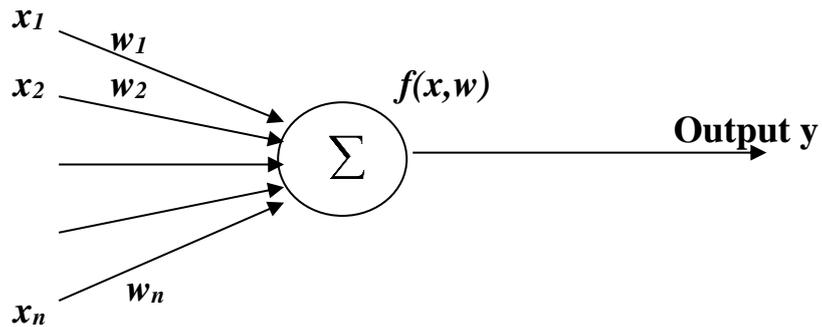
Figure 2.6: A simple neuron.
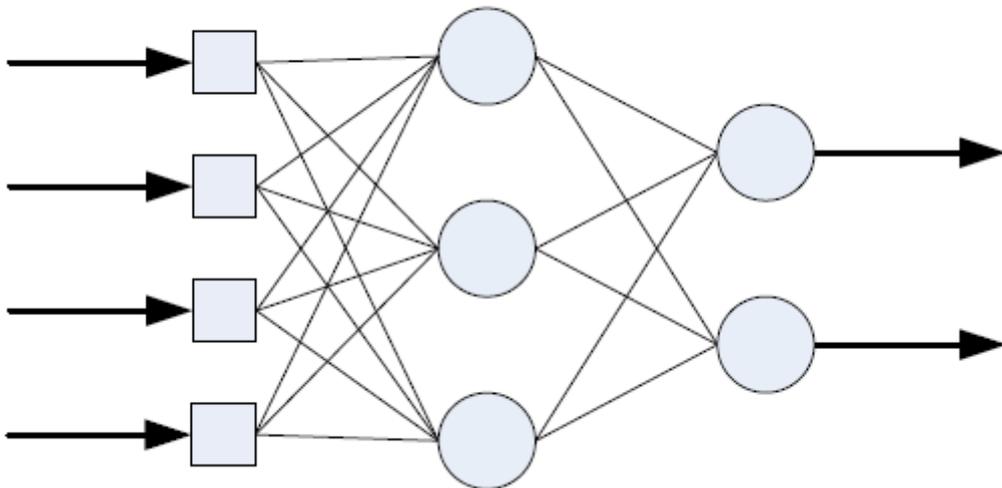
Figure 2.7: FNN architecture.

The architecture of FNNs is shown in figure 2.7. Each neuron is connected to all the neurons in the previous and next layers. Each neuron takes input from all the neurons in the preceding layer. The layer at the left side is the input layer while the layer at the right side is the desired

output. The number of neurons in the output layer depends on the number of classes to be classified. There exists at least one layer between the input and output layer and there is no connections between the neurons of the same layer [70]. Data flows in one direction in this network starting from the input to reach the output layer. The number of neurons in the hidden layer is determined experimentally.

The challenge in the multilayer neural networks is to set the weights based on the desired outputs. Backpropagation is one of the most general methods of multilayer neural network supervised training [62]. It changes the layer weights making the actual output more similar to the desired one. Since this is not as easy for the input-to-hidden weights as there is no proper known output for the hidden units, BP calculates an error for each hidden unit and derives a learning rule for the weights. The Backpropagation algorithm for MLP is implemented in the following way:

1. The network weights are initialised using the Nguyen-Widrow initialization method. This method assigns values to the weights randomly in order to evenly distribute the active region of each neuron in the layer.

2. Present an input training with the desired outputs to the network and determine the actual output. For a given input x in the training set with an actual output y, BP computes the training error:

$$E(w) = \frac{1}{2} \sum_{k=1}^{c} \left(y_k - z_k\right)^2 \tag{2.18}$$

where z_k is the desired output and w represents all the weights in the network.

2. Adjust the weights to reduce the measure of error. Initially random values are given to the weights and then changed to reduce the error. This learning rule is based on gradient descent:

$$\Delta w = -\eta \frac{\nabla E}{\nabla w} \qquad (2.19)$$

where η is the learning rate. It is used to specify if the neural network will make major or minor adjustments after each learning.

3. The weight vector at iteration m is updated by this iterative algorithm:

$$w\ (m+1) = w(m) + \Delta w(m) \qquad (2.20)$$

The learning repeats until the minimum mismatch between the desired and actual output is achieved.

2.6 Summary

The basic concepts of radar, ultrasonic, and pattern recognition were reviewed in this chapter. The surfaces of interest, the materials and the characterization of these surfaces were discussed. The change in polarisation of the reflected radar electromagnetic waves, and the echo energy and duration of backscattered ultrasonic signals characterize the properties of the surface (materials and roughness). Pattern recognition techniques proved to be suitable for remote sensing data classification. A survey of the literature review concerning radars,

ultrasonic sensors, and pattern recognition techniques for surface identification reveals the following:

1. Only few surfaces have been investigated by the analysis of backscattering properties. The focus was differentiating wet, dry, icy and snowy asphalt. Other challenging off-road surfaces like gravel, sand, grass, and mud were ignored in the previous studies.

2. Analysis of various classification algorithms was carried out in the case of ultrasonic signal. Nevertheless, attempts to analyse the performance of classification algorithms in the case of microwave signals haven't been done. The goal of most studies was to the find difference of the backscattered signals from wet and dry asphalt, and didn't aim to study the efficiency of statistical classification algorithms. Besides, choice of the classification algorithms was subjective as to the author's preferences.

3. The influence of external factors (vehicle movement and weather conditions) that might influence the accuracy of surface recognition was not highlighted in the reviewed papers.

4. Most of the studies were conducted in stationary laboratory conditions.

5. A lack of research on surface recognition using radar and ultrasonic sensor fusion including various combinations of frequency/polarimetric data was revealed. This approach can accumulate the advantages of different spatial information leading to the improvement of classification performance.

The goal of our study is to analyse the repeatability of reflected signal properties within the specified type of surface to trace the difference in individual measurement properties of backscattered signals. Since the real conditions highly affect the radar and ultrasonic reflections, it is necessary to implement a significant statistically representative number of measurements for each possible combination (sensor, surface, cover, frequency, etc.). This allows a reliable surface identification technique based on the statistical methods.

Chapter 3

Clutter

3.1 Introduction

It is necessary to understand the features of the backscattered signal and the factors affecting these features during the surface identification. The dependence of the backscattered signal on the surface properties (roughness, material, etc.) and the system configuration parameters (grazing angle, height of the antenna, beamwidth, etc.) are analyzed in order to assist with the extraction of the features and optimizing the system configuration for the finest surface identification. This chapter will study the clutter from land surfaces, in general, and the factors affecting it. A model of the reflected signal power and its dependence on surface properties and system configuration is developed in order to compare with the real time reflected signals.

Clutter is the reflections returned from an object or set of objects unwanted for a specific application. It describes the unwanted radar returns that might interfere with the radar operations [75]. Clutter can be classified into two categories: the clutter which enters through the main lobe of the antenna is called main lobe clutter; otherwise it is called sidelobe clutter. It can generally be classified into two types: surface clutter and volume clutter. The surface clutter is the unwanted reflections from trees, vegetations, terrains, and sea surface, whereas

the volume clutter describes the reflections from rain, dust, spray, birds, insects, etc. The latter usually has larger power but it is more predictable. Surface clutter is not predictable as it changes from one area to another [75]. An important concept about clutter is that its definition varies from one field to another. For someone trying to detect a pedestrian on the road a reflection from the asphalt would be considered as an unwanted reflection (clutter), while for someone trying to detect the nature of the surface on the road this same reflection would be the primary target [76]. Therefore, the same object can be the clutter for one application, and the main target for another. In this chapter, the surface clutter will be discussed as it is the target of interest for a road surface identification radar system.

3.2 Surface Clutter

Surface clutter is major concern for radars at low grazing angles and many measurements of land clutter have been performed before. Nevertheless, several studies have shown there is a lack of solid information about it. A summary of what has been performed and published regarding clutter is shown in [77][78]. The uncertainty in the value of land clutter was always a concern as , in addition to lack of published data, several investigators had different values of land clutter while looking at the same data[4]. The backscattering echo from Clutter is commonly given in terms of radar cross section density due to its distribution [4].The radar relies on the target radar cross section and the clutter radar cross section in order to distinguish target returns from the clutter returns. The cross section is a measure of the target as seen by the radar. The average clutter radar cross section is given by σ_c :

$$\sigma_c = \sigma^{\circ} A_c \tag{3.1}$$

where σ° is the cross section per unit area referred as clutter scattering coefficient, and A_c is the clutter area. The grazing angle, surface roughness and material, as well as the radar wavelength are the factors affecting the amount of clutter reflected back from the surface. Figure 3.1 shows the area where surface clutter is picked up for radar waveform [75] [4].

(a)

(b)
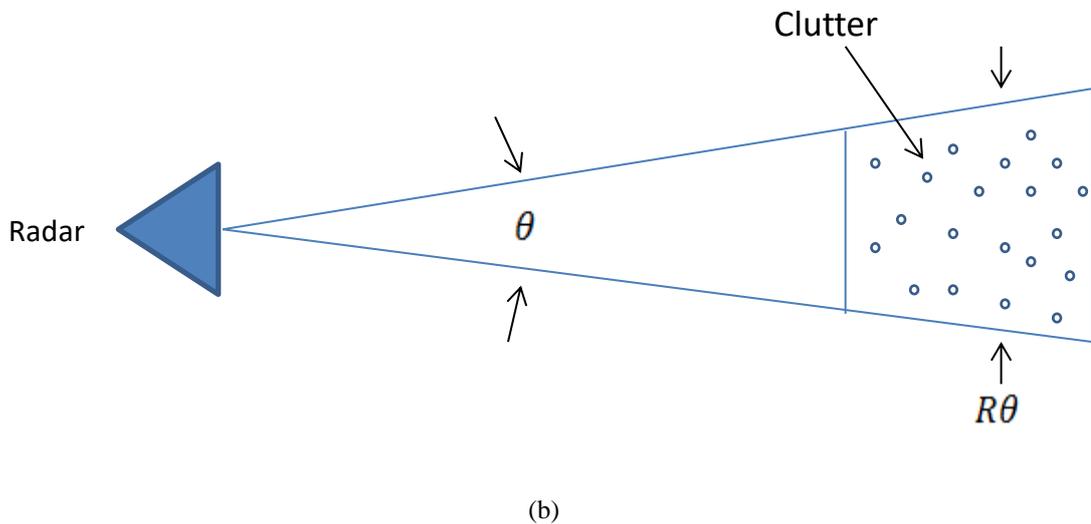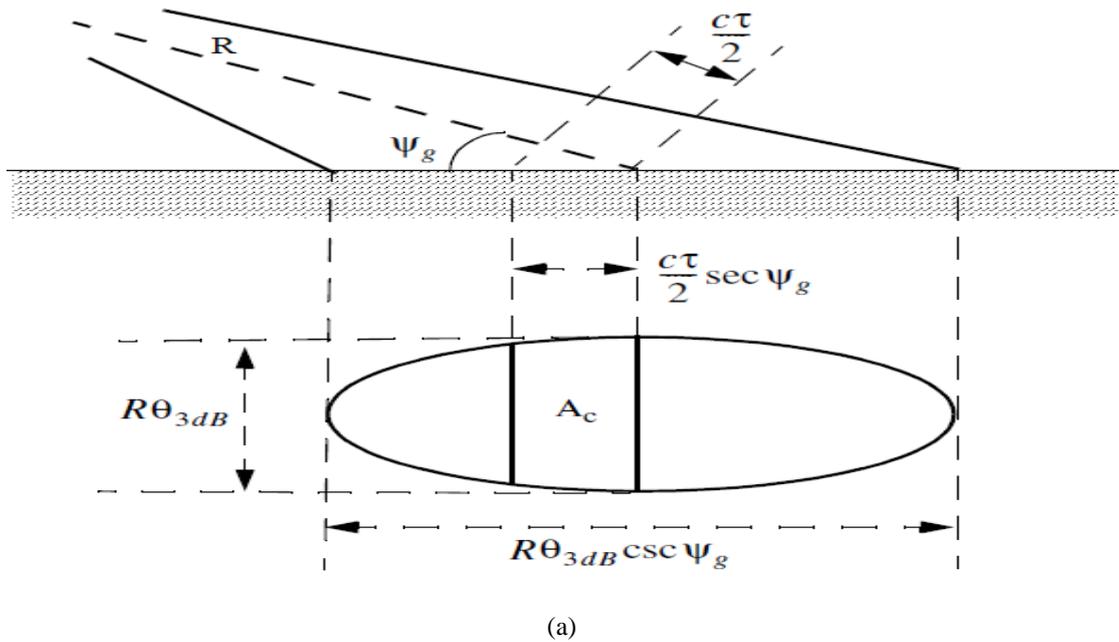
Figure 3.1: Geometry of radar clutter. (a) Elevation view, (b) View from above [4].

Figure 3.1 (a) shows a radar illuminating the surface at a grazing angle Ψ_g (rad). The elliptical footprint is defined by the intersection of antenna beam with the ground. The

elevation view shows that it is divided into range bins of size $\frac{c\tau}{2}$ sec Ψ_g , where τ is the transmitted pulsewidth and c is the speed of light (c $=3\mathrm{x}10^8$ m/s). Therefore, the range dimension of the clutter area A_c is determined by the radar pulse width τ [75]. Figure 3.1 (b) shows the horizontal view on the surface consisting of independent scatterers responsible for the clutter. The width of the clutter area A_c is determined by the azimuth beamwidth θ [77]. The cell of the linear frequency modulation "Chirp" used by our radar system depends on the time gating applied. The clutter area A_c (meters^2) is given by:

$$A_c = R\theta \frac{c\tau}{2} \sec \psi_g \tag{3.2}$$

where R is the slant range between the antennas and surface (meters) and θ is the 3 dB two way azimuth beamwidth (radians). Substituting Eq. (3.2) into (3.1), the cross section per unit area is equal to:

$$\sigma^\circ = \frac{\sigma_c}{R\theta \frac{c\tau}{2} \sec \Psi_g} \tag{3.3}$$

The power received by the radar from a target with cross section σ_t is given by:

$$S = P_r = \frac{P_t\, G^2\, \lambda^2\, \sigma_t}{(4\pi)^3\, R^4} \tag{3.4}$$

where P_t is the transmitted power, G is the transmitting and receiving antenna's gains (gains are equal since transmitting and receiving antennas are the same), and λ is the wavelength. Similarly the power received from the clutter is given by:

$$C = \frac{P_t\, G^2\, \lambda^2 \sigma_c}{(4\pi)^3 R^4} \tag{3.5}$$

Substituting equation (3.3) for σ_c in (3.5) gives:

$$C = \frac{P_t\, G^2\, \lambda^2 \sigma^\circ\ \theta\ \frac{c\tau}{2}\ \sec \Psi_g}{(4\pi)^3 R^3} \tag{3.6}$$

The signal to clutter ratio at low grazing angle is obtained by combining Eq. (3.4) and (3.6) [75]:

$$\frac{S}{C} = \frac{\sigma_t}{\sigma^\circ\, R\ \theta\, \frac{c\tau}{2}\ \sec \Psi_g} \tag{3.7}$$

For reliable detection of a target, the radar should have a high signal-to- clutter ratio. This doesn't apply in surface identification as the clutter is the target of interest.

3.2.1 Radar Cross Section of surface

In the design and development of radars, it is essential to describe the echo from target in terms of the characteristics of this target. Size, shape, and orientation are some of the target characteristics that are an interest for the operation of radar. Therefore, the target is recognized by an effective area called the radar cross section (RCS). RCS describes the intensity of backscattered wave and relates it to the amount of power transmitted to the target. The radar cross section definition is given by [79]:

$$\sigma = \lim_{R \to \infty} 4\pi R^2 \frac{|E_s|^2}{|E_0|^2} \tag{3.8}$$

48

where E_0 is the strength of the incident electric field on the target, E_S is the strength of the scattered electric field at the radar, and R is the range from the radar to the target. Although most target don't scatter the power equally in all directions, but equation (3.8) assumes that the power from the incident wave is radiated uniformly in all directions. The Friis transmission equation is given by the ratio of the received power at the receiving antenna to the output power of the transmitting antenna:

$$\frac{P_r}{P_t} = \frac{G^2 \lambda^2 \sigma_t}{(4\pi)^3 R^4} \tag{3.9}$$

The same ratio is given by the square of the reflected voltage S12 with describes the relationship between the reflected and incident wave [80]:

$$|S_{12}|^2 \sim \frac{P_r}{P_t} = \frac{G^2 \lambda^2 \sigma_t}{(4\pi)^3 R^4} \tag{3.10}$$

The radar cross section of the surface is given by the clutter radar cross section σ_c . Using equation (3.5):

$$\sigma_c = \frac{(4\pi)^3 R^4 C}{P_t G^2 \lambda^2} \tag{3.11}$$

where C is the power received for clutter (P_r). Equation (3.11) can be written as:

$$\sigma_c = |S_{12}|^2 . \frac{(4\pi)^3 R^4}{G^2 \lambda^2} \tag{3.12}$$

The scattering coefficient or scattering cross section per unit area σ° is generally used to describe the ground returns rather than the total scattering cross section σ_c. In contrary to σ_c, σ° is independent of the geometric radar parameters (such as pulse width and beamwidth) and doesn't vary with the illuminated area [79]. Since the road and off-road surfaces of our interest are considered rough surfaces, the backscattered signal is an accumulation of several weak signals due to the irregularities of the surfaces. This means that the backscattered signal changes randomly according to these randomly distributed irregularities [80]. These kinds of targets are characterized by the scattering coefficient given by the average radar cross section over the illuminated area (equation 3.3). Substituting (3.12) into (3.3) gives the scattering coefficient of the surface:

$$\sigma^\circ = |S_{12}|^2 \cdot \frac{(4\pi)^3 \, R^3}{\theta \, \frac{c\,\tau}{2} \, sec \, \Psi_g \, G^2 \, \lambda^2} \tag{3.13}$$

As can be seen from equation (3.13), the scattering coefficient σ° depends on the reflected voltage $|S_{12}|$. $|S_{12}|^2$ is calculated through real time experimentation and measuring S parameters of each surface and setup configuration. All the other parameters of dependence are provided for different frequencies, grazing angles, and antennas. Using equation (3.13), a model of land clutter was developed in Matlab, for demonstration purposes, to show the effect of the roughness of several surfaces and grazing angles on the reflected voltage. Figure 3.2 (a) shows the dependence of S12 on the scattering coefficient of three different surface types (flat land, farm land, and mountains) at 5.8 GHz (one of the ISM bands used in our measurements) with respect to distance. These surfaces can be eventually rescaled into our surfaces of interest. It is expected that as distance increases, S12 decreases rapidly (from -40 dB to -72 dB for flat land). The highest
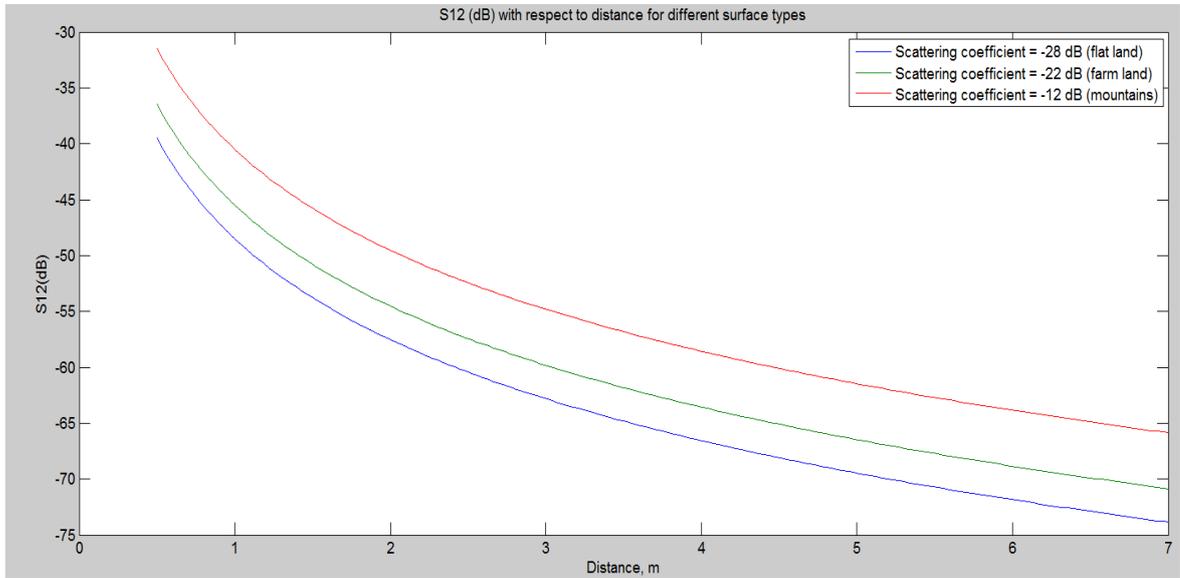
(a)
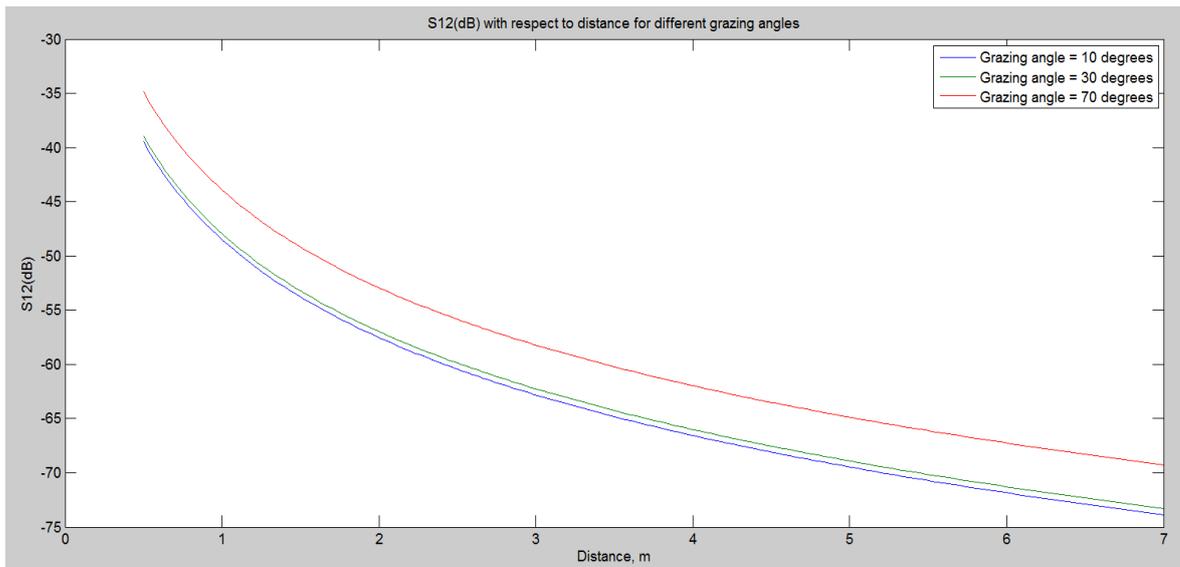
(b)

Figure 3.2: Model of land clutter. (a) Dependence of S12 on scattering coefficients, (b) Dependence of S12 on grazing angle.

the scattering coefficient, the higher is the reflected voltage S12. At any distance S12 of

mountains is higher than farm land and higher than the flat land. Figure 3.2 (b) shows the

dependence of S12 on grazing angle. The graph shows the reflection voltage of a flat land

with respect to distance at 5.8 GHz. There hasn't been a significant influence of the grazing angle on the slope of the reflected voltage, but higher grazing angles yields higher S12.

3.2.2 Dependence of Surface Clutter on Roughness

The predictability of a reflection out of a surface depends on the level of the surface (horizontal, angled, curved, etc.). An incident wave on a smooth surface leads to a scattering pattern that is concentrated in the specular direction upon leaving the surface. The reflection called Specular reflection is described by the Fresnel reflection laws [81]. As the signal rays meet a rough surface which has a different orientation, it will reflect partly in the specular direction (referred as coherent component) and partly in all other directions (referred as diffuse component). Reflection from very rough surfaces is dominated by the diffuse component where the coherent component is negligible. Fortunately, the road and off road surfaces of our interest are considered rough or slightly rough in the worst case scenario. This is due to the fact that a monostatic receiver wouldn't be able to receive power from smooth surfaces except for normal incidence [78]. Figure 3.3 shows the relationship between the surface roughness and the scattering patterns. Diffuse reflection from rough surfaces is shown in figure 3.3 (a). The scattering pattern shows scattered component in all directions, and a reflected component in the specular direction. On the other hand, specular reflection from smooth surfaces is shown in figure 3.3 (b). This obeys the law of Fresnel reflection. Since surface's roughness is a very important factor affecting the scattering pattern, the clutter received from the surface highly depends on it. Roughness is described by two parameters: the root mean square

(a)   (b)
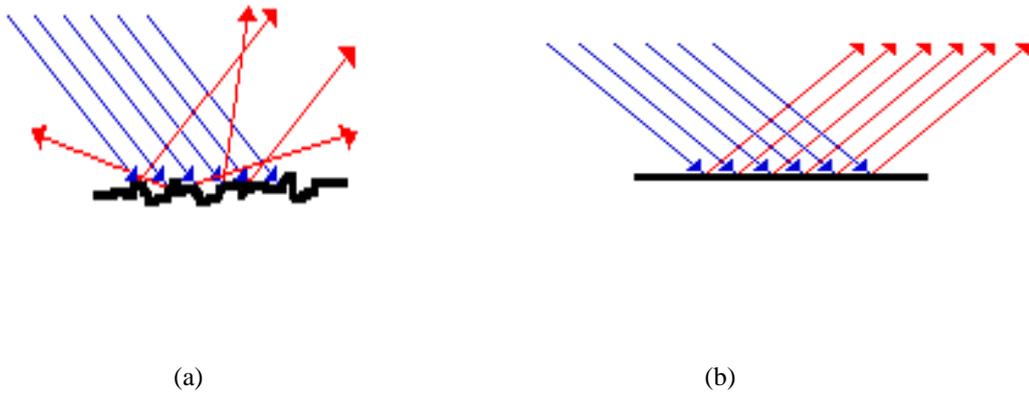
Figure 3.3: Relationship between surface roughness and scattering patterns. (a) Diffuse reflection, (b) Specular reflection.

height (rms) and the correlation length. Both parameters describe the surface height as compared to a reference surface that can be mean surface or periodic one. Figure 3.4 shows an example of a mean level surface reference and a random rough surface component [78].
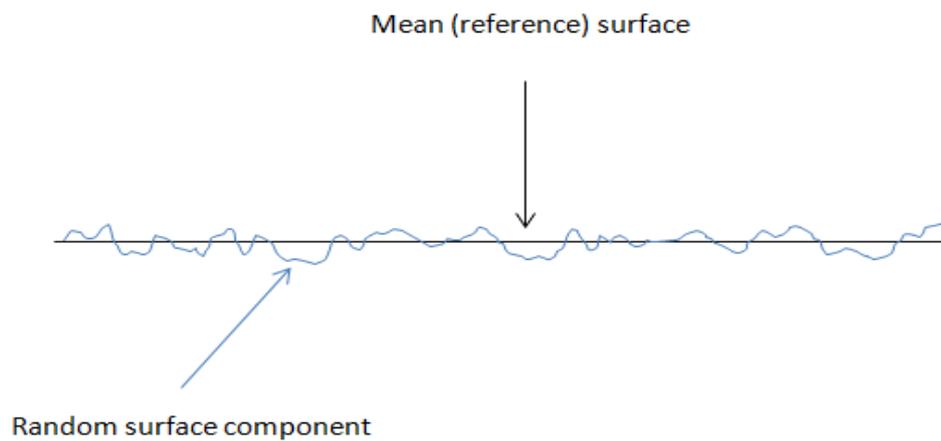
Figure 3.4: Random height variations on a flat surface [78].

The rms referred as the standard deviation of the surface height considers that for each point (x, y) of the surface there is a height z (x, y) in the x-y plane. The mean for a segment of the surface of dimensions L_x and L_y is given by:

$$\bar{z} = \frac{1}{L_x L_y} \int_{-L_x/2}^{L_x/2} \int_{-L_y/2}^{L_y/2} z(x,y) \ dx \ dy \qquad (3.14)$$

The second moment is given by:

$$\overline{z^2} = \frac{1}{L_x L_y} \int_{-L_x/2}^{L_x/2} \int_{-L_y/2}^{L_y/2} z^2(x,y) \ dx \ dy \qquad (3.15)$$

Consequently, the rms is given by [82]:

$$\sigma = (\overline{z^2} - \bar{z})^{1/2} \qquad (3.16)$$

In addition to the standard deviation of the surface height (σ), the surface correlation length also describes the variation of the surface height relative to a reference surface. The normalized autocorrelation function of a one dimension surface profile z(x) is given by measuring the likeness between the height z at a point x and at a point x':

$$\rho(x') = \frac{\int_{-L_x/2}^{L_x/2} z(x)z(x+x') \ dx}{\int_{-L_x/2}^{L_x/2} z^2(x) \ dx} \qquad (3.17)$$

54

The correlation length l is defined as the displacement distance x' (in cm) where $\rho(x') = \rho(l) = 1/e$ [82]. This estimates the dependence between any two points x and x' on the surface. If the distance of separation between the two points is greater than l, then their height are considered independent of each other. In case of perfectly smooth surface, $l = \infty$ so that every point on the surface is correlated and dependant on every other point. Thus, the autocorrelation coefficient is equal to 1. Figure 3.5 shows how surfaces with different roughness can be represented by the two parameters [48]. In figure 3.5 (a), the rms is almost zero for a mirror-like smooth surface, while figure 3.5 (b) is considered rougher as it has σ1> 0 and T1 as the correlation length. 3.5(c) is considered moderately rough with σ1> 0 and T2 >T1. It can be shown that figure 3.5 (c) has lower frequency of oscillations than that of figure 3.5 (b). The higher the correlation length, the bigger is the distance of separation between the points. Rough surfaces are shown in figure 3.5(d) where σ2 > σ1 > λ.

Smoother surface (σ ≈ 0)

(a)

T1 and σ1

(b)

T2 >T1 and σ1

(c)

T1 and σ2 > σ1

(d)
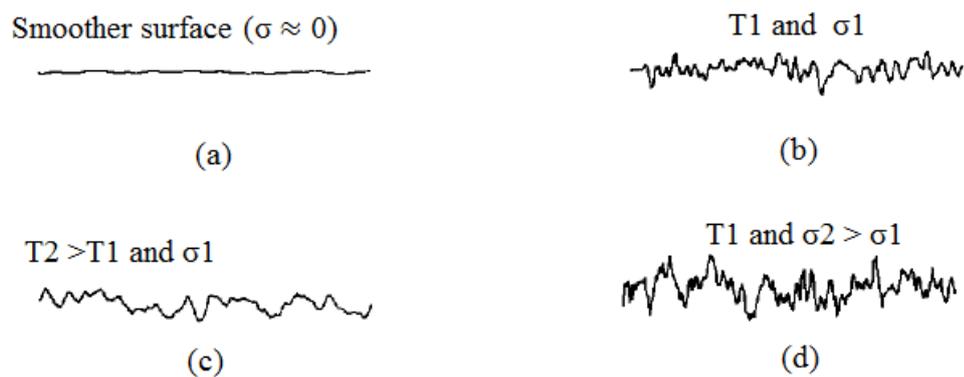
Figure 3.5 Surface description of different rms height and correlation length. (a) smooth , (b) and (c) moderately rough, (d) rough surface[48].

Although the aforementioned parameters describe the statistical variation of the heights of the surface, but some surfaces may appear smooth to microwave and rough to another technology (optics). This is due to the relativity of roughness to the wavelength. Two criteria defining scattering from objects were considered: Rayleigh criterion and Fraunhofer criterion. Rayleigh criterion specifies the separation needed between two diffractions, for them to be resolved as different objects. According to the Rayleigh criterion, the surface is considered smooth when the phase difference between two reflected rays is less than $\frac{\pi}{2}$ radians. The phase difference is geometrically given by:

$$\Delta\emptyset = 2kh \, \cos\theta \tag{3.18}$$

where k $=\frac{2\pi}{\lambda}$, h is the difference in height between the two points the rays were reflected at, and θ is the incidence angle. For a smooth surface, $\Delta\emptyset$ is less than $\frac{\pi}{2}$ which gives [81]:

$$h < \frac{\lambda}{8 \, \cos\theta} \tag{3.19}$$

Equation (3.19) can be replaced by $\sigma < \frac{\lambda}{8 \, \cos\theta}$ for a random surface with standard deviation of surface height (σ). Another criterion which is used for modeling the scattering behaviors of surfaces in the microwave region is called Fraunhofer criterion. This criterion states that the rays from the centre and edge of the antenna should have a maximum phase difference of $\frac{\pi}{8}$. Accordingly, a smooth surface should have a standard deviation (σ):

$$\sigma < \frac{\lambda}{32 \, \cos\theta} \tag{3.20}$$

Roughness is not defined with respect to wavelength only. Incidence angle plays a role on how rough the surface is considered. For asphalt roads, the surface is considered smooth for radar frequencies but it can be moderately rough for wideband radars. At small grazing angles (large incidence angles), surfaces appear smoother [82]. Figure 3.6 [83] shows the dependance of surface clutter on roughness of the surfaces. Figure 3.6 (a) shows the scattering coefficient $\sigma°$ for five soil surfaces with different roughness and similar moistures (similar dielectric constants) at 4.25 GHz and HH polarization with respect to angle of incidence (degrees). At low incidence angles (close to normal incidence), the Scattering coefficient $\sigma°$ varies significantly between the 5 surfaces. The smooth surfaces have higher energy reflected in the specular direction. As the incidence angle increase, $\sigma°$ of smooth surfaces decreases rapidly inversely proportional to the incidence angle (between 18dB at normal incidence and – 9 db at $\theta = 30°$). On the other hand , rough surfaces (rms= 4.1 cm) decreases slowly and doesn't show dependance of the clutter on the incidence angle. The scattering coefficient $\sigma°$ is strongly dependant on the surface RMS especially near to low incidence angles, but this dependence is dynamic due to the effect of frequency and incidence angle [83]. Figure 3.6 (b) shows the angular response of the same surface at a frequency of 7.25 GHz. Same conclusion can be drawn from Figure 3.6 (b) , but it is worth highlighting that the same surface with the same roughness has higher $\sigma°$ at lower frequencies. Therefore, the surface has more incident energy specularly reflected and is ,therefore, considered smoother at lower frequencies. In contrary to 4.25 GHz, the rough surface is somehow dependant on the angle of incidence at higher frequency (7.25 GHz). No significant difference was observed between different polarizations.
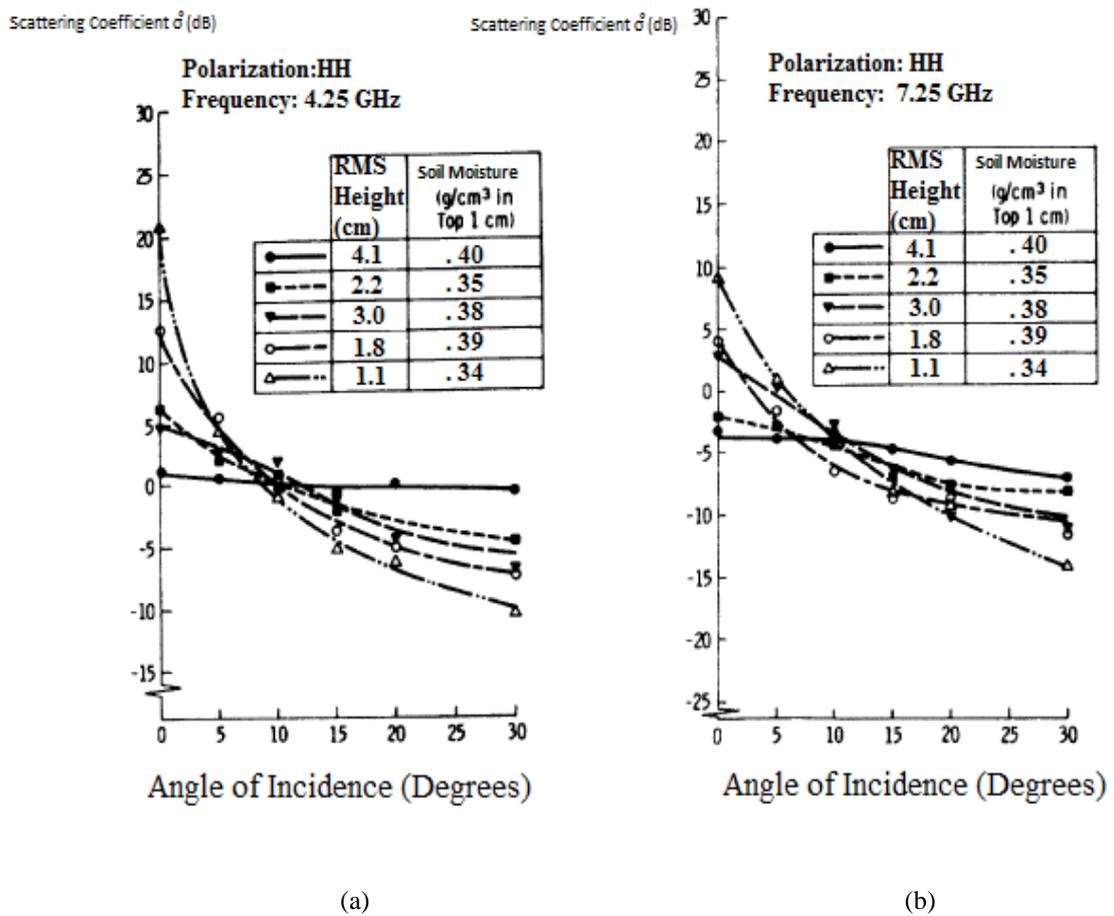
Figure 3.6 Angular response of scattering coefficient with respect to Angle of incidence (Degrees) for five fields with different RMS height and soil Moisture at C-band [83].

3.2.3 Dependence of Surface Clutter on Dielectric Constant

The scattering coefficient $\sigma°$ is dependent on the material of the surface which is characterized by the dielectric constant ε. Even if the surface roughness did not change, $\sigma°$ may vary up to 10 dB between dry and very wet surfaces or any different surface materials [84]. Therefore, asphalt, gravel, sand, grass, and snow have different scattering characteristics. There has been an attempt to create physical models of dielectric properties of different surfaces but the practical application of those models is complicated due to the difficulty of calculation of some parameters. Such parameters (like the permittivity) must be

retrieved experimentally for each type of surface [85]. The backscattering model from a statistically rough surfaces like road and off-road surfaces can be written as:

$$\sigma^\circ = \left| R_{pp} \right|^2 g(s, l, \theta, \lambda) \tag{3.21}$$

where $g(s, l, \theta, \lambda)$ is a function of incidence angle, wavelength, and surface roughness parameters (independent of dielectric constant). R_{pp} is the Fresnel reflection coefficient expressed for different polarizations as [81][86] :

$$R_{vv} = \frac{\varepsilon_r \cos \theta - \sqrt{\varepsilon_r - sin^2\theta}}{\varepsilon_r \cos \theta + \sqrt{\varepsilon_r - sin^2\theta}} \tag{3.22}$$

$$R_{hh} = \frac{\cos \theta - \sqrt{\varepsilon_r - sin^2\theta}}{\cos \theta + \sqrt{\varepsilon_r - sin^2\theta}} \tag{3.23}$$

where $\boldsymbol{\varepsilon_r}$ is the relative dielectric constant of the surface. This dependence leads to a big variation in the backscattering coefficients of different surfaces. At an incidence angle of 45°, at L band, a dry sand of $\varepsilon_r = 3 - i0$ and a water saturated sand of $\varepsilon_r = 79 - i1.5$ demonstrates a difference in backscattering coefficient of 13.9 dB for VV polarization and 7 dB for HH polarization [82]. Figure 3.7 shows an example of clutter σ° from different surfaces at vertical polarization and X band during dry weather conditions. The wide range of data for each terrain is represented by the wide boundaries of the regions in the graph [4]. At low grazing angles, the difference in scattering coefficient is significant between the four terrains. The highest difference of 40 dB approximately is measured between asphalt and city surfaces at 10° grazing angle.
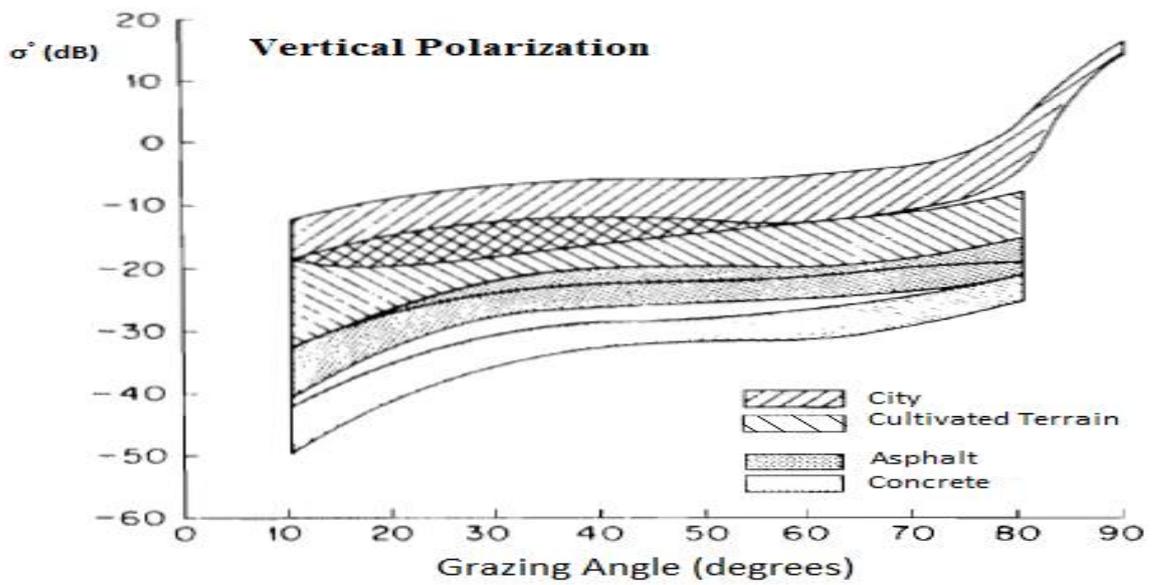
Figure 3.7 Measured clutter σ° for different terrains with respect to Grazing angle (Degrees) at X band [4].

Figure 3.8 shows the ratio σ_{vv}/σ_{hh} as a function of incidence angle for different dielectric permittivity. At high incidence angles, the ratio increases by 20 dB as the dielectric permittivity ε_r increases from 3 to 33 [80].
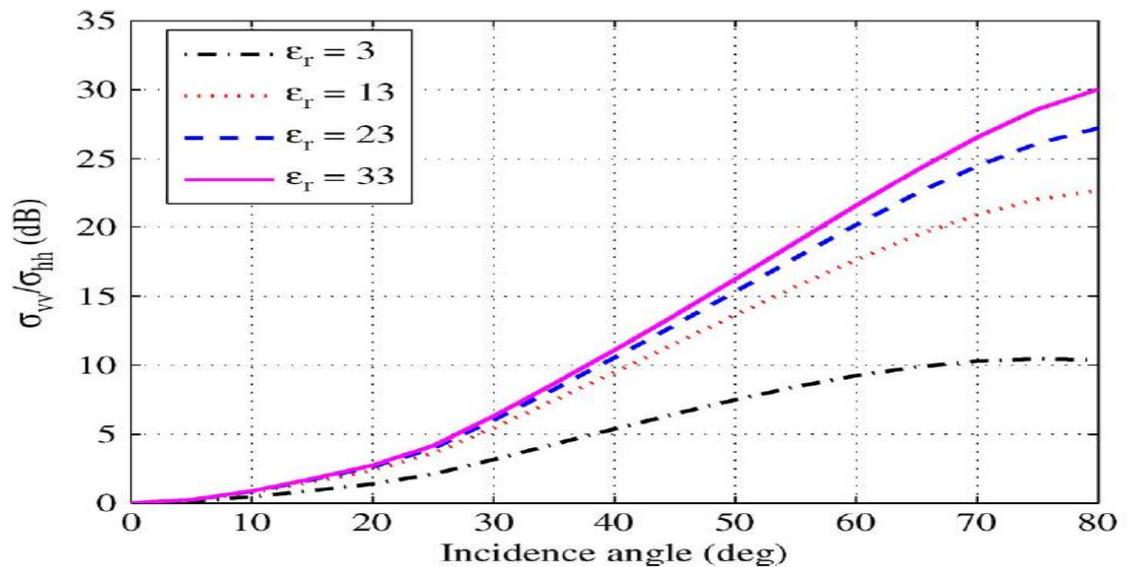
Figure 3.8 Ratio of backscattering at vv and hh polarizations for different permittivity [80].

All the studies and experiments performed showed that clutter is dependent on the dielectric constant of the surface regardless of the roughness or the grazing angle. Some studies showed that relative permittivity is the most influencing parameter amongst all the other values [80].This showed that looking at the scattering coefficients of different surfaces would be very useful for any classification algorithm used to differentiate between different surfaces with different conditions.

3.2.4 Dependence of Surface Clutter on Polarization

As stated in the previous chapters, polarization plays a big role in analyzing backscattered signals. The change in the polarisation of electromagnetic waves reflected from the road surface is governed by the dielectric constant of the surface material and the surface roughness. Moreover, radar cross section depends on the polarization of the incident and reflected wave. The components of the scattering matrix are given as a function of radar cross section [87]:

$$S_{ij} = \frac{\sqrt{\sigma_{ij}}}{4\pi R^2} \tag{3.24}$$

where $\sqrt{\sigma_{ij}}$ is a complex number with amplitude and phase. This agrees with equation (3.8) and is shown as:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = 4\pi R^2 \begin{bmatrix} |S_{11}|^2 & |S_{12}|^2 \\ |S_{21}|^2 & |S_{22}|^2 \end{bmatrix} \tag{3.25}$$

Once a scattering matrix is found, the radar cross section of the target can be deduced for all the combinations of vertical and horizontal polarizations [75]. Observations showed that the vertically polarized scattering coefficient $\overset{\circ}{\sigma}_{vv}$ is bigger than the horizontally polarized $\overset{\circ}{\sigma}_{hh}$ [82]. Figure 3.9 shows measured backscattered coefficient (dB) with HH and VV polarizations as a function of grazing angle at X-band [88]. As shown, there is a significant difference between the two polarizations for low roughness surfaces. Figure 3.10 shows the measured backscattered coefficient (dB) with different polarization ratios for dry and wet surface asphalt at 24 GHz. As discussed before, polarization ratios are used to reduce the effects of measurement distance, weather conditions, etc. Single measurements in figure 3.10 are represented by markers, and averaged values are represented by the solid lines [80]. For the vertical/horizontal ratio σ_{vv}/σ_{hh} there is a 5 dB difference between dry and wet asphalt. Nevertheless, there are more than 10 dB differences between σ_{vv}/σ_{hh} and σ_{vh}/σ_{hh} for the same dry asphalt. Although different surfaces weren't studied, this shows that using different polarizations
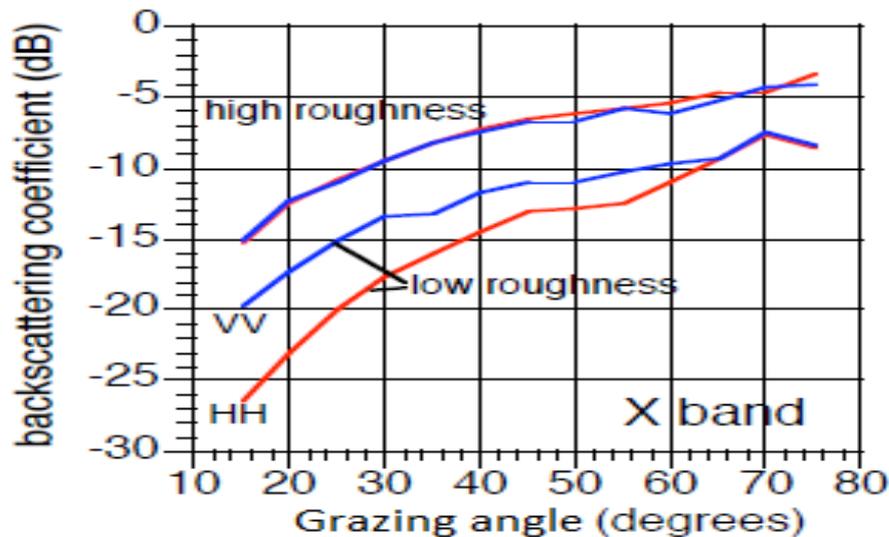
Figure 3.9 Measured clutter σ° for several terrains at different polarizations with respect to Grazing angle

(Degrees) at X band [88].
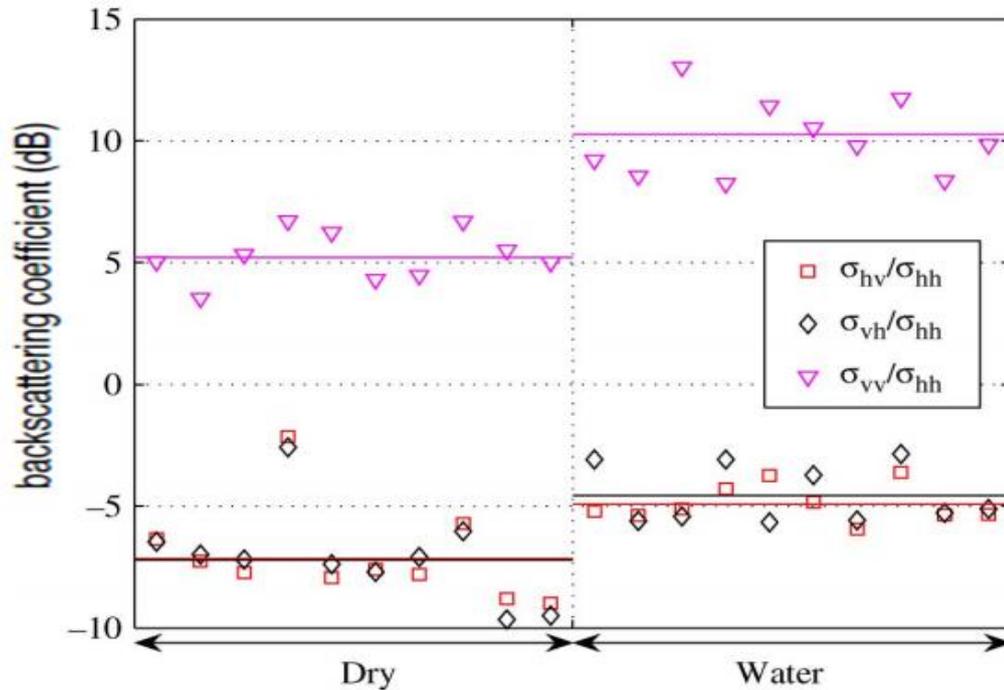
Figure 3.10   Backscattering from dry and wet asphalt at different polarizations at 24 GHz[80] .

would enhance the identification reliability for different surfaces conditions.

3.3 Ground Clutter Models

Formulating a theoretical model that can represent the backscattering from different land surfaces is considered difficult and very complex task due to the variety of the ground terrain. Furthermore, the weather, day time, season of the year, and the discontinuities of the surfaces are major factors that affect the backscattering and make it harder to a reliable land clutter model. Nevertheless, there were many attempts to model land clutter. It has been modeled as a Lambert surface with scattering coefficient σ° varying according to the grazing angle θ [89]. It has also been modeled as assemblies of spheres, hemispheres, and hemicylinders as

in [4]. In addition, it was modeled as the specular reflection from small segments called facets. Energy is backscattered when the facets are perpendicularly oriented against the radar line of sight [90]. Land clutter from rough surfaces was modeled as small-perturbation model which is suitable for asphalt but not off-road surfaces since it assumes that target is homogenous with low roughness compared to the wavelength [80]. Moreover, it was modeled as an application of the Kirchhoff-Huygens principle. This principle assumes that the current that flows at each point in the curved surfaces is the same that would do if the surface was flat and tangent oriented towards the actual surface [91]. Such an assumption helps to construct scattered fields when assuming that the current over the rough surface is the same magnitude if the surface was smooth but with different phase perturbations. Although these models provide some understanding of the land clutter, and relate some of the surface parameters to the backscattering properties but none of them have been considered a very successful general land clutter model [4]. The reason is that the clutter is predominantly a statistical process, and the variation in the radar backscattering occurs to be dependent on much more parameters than a single current model can offer. Furthermore, subtle differences in radar returns due to different deterministic characteristics, such as dielectric permittivity and roughness, can be obscured by the combination of parameters and unaccounted external interferences. Thus, though theory indicates dependence of clutter on those parameters the only way to differentiate practical surfaces is by using statistical classification algorithms.

3.4 Conclusion

This chapter showed the dependence of the power of the backscattered signal on the roughness of the surface and dielectric permittivity. It also showed that different surfaces

have different backscattered properties at different polarization ratios. A model of clutter was developed to show the dependence of the reflected voltage on the grazing angle and scattering coefficient, i.e., the surface type. The clutter models which were used by the previous studies to compare to the experimental results were highlighted and it showed that the measurement of polarization ratios for different surfaces is the best method to differentiate between these surfaces. The conclusion made was that the average power extracted from the reflected voltage S21, at a certain time gate for different polarizations ratio (VV/HH, VH/HH, and HV/HH), would be the most distinctive features to be used in the classification algorithms and optimization of the system configuration for the finest surface identification. In the next chapters, the use of the proposed methods will be analyzed for surface recognition. Practical results showed that more features can be extracted if the bandwidth was wide enough. These parameters are identified as "duration and power above the threshold" and will be explained in the next chapters.

Chapter 4

Surface Identification Data Collection

4.1   Introduction

In order to establish a reliable automatic surface identification technique, it is crucial to implement a sufficient statistically representative number of measurements for each of the surfaces. while accounting for all the possible combinations (incidence angle, frequency, bandwidth, etc.).

Collecting data and analyzing different on-road and off-road surfaces were carried out as outdoor experiments at different locations near Birmingham and Gaydon, UK. All the experiments took place in a real case scenario with different weather conditions which influenced the cover of the surfaces (wet, dry, icy, etc.). The collected data delivered by radar and ultrasonic sensor were used to build a database of different surfaces with multiple features. This chapter describes the outdoor experimental setup, the hardware used. It also highlights the influence of weather conditions and other factors affecting the measurement's accuracy.

4.2 Outdoor Experimental Setup

A developed radar system was used to analyze the polarization change of the

electromagnetic waves reflected from different surfaces. The system developed at the University of Birmingham   consists of forward-looking monostatic radar based on an Agilent FieldFox N9918A network analyzer, two WBH1-18 S horn antennas, and Ultrasonic sensor based on SRF08 pulsed ultrasonic range finder [92].On a later stage, four narrowband 24 GHz antennas were developed to be used instead of the horn antennas. The antennas functioning as transmitter and receiver with orthogonal polarization were coherently interconnected.   Measuring the phase shifts between transmitted and received signals involves the use of a radio frequency network able to supply transmitting antennas with the same reference signal. Fig 4.1 shows the block diagram of the experimental setup.
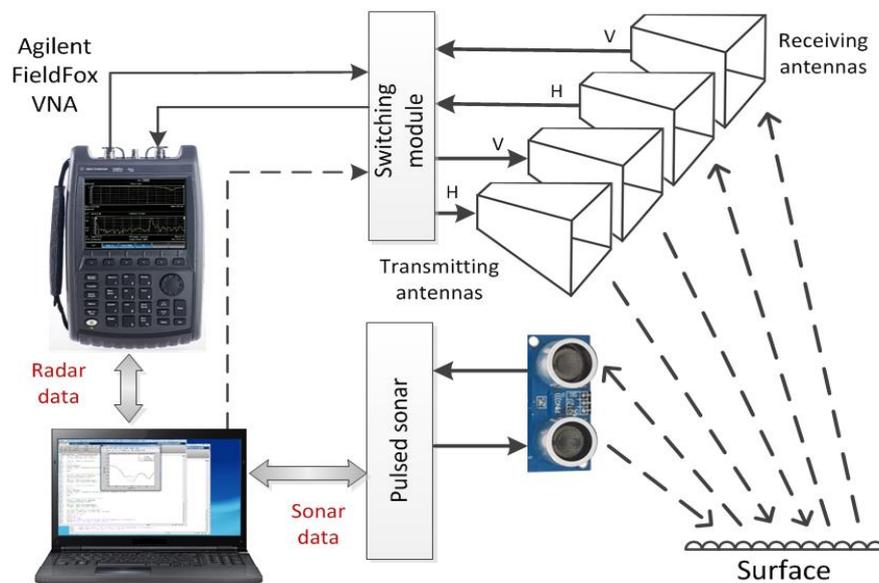
Figure 4.1: Block diagram of the measuring system.

The antennas were mounted, along with the ultrasonic sensor, on a rotating support to ease backscattering measurements at different incidence angles and different polarizations (Fig. 4.2). By mounting the antennas at 0.4 meters above the surface and rotating the support to an incidence angle of 10°, the slant distance to the surface can be measured (2.3 meters) at

the centre of the antenna beam (Fig 4.3). The signal received by the antenna is analyzed, in the frequency domain, by the VNA in the form of S_{21} where a time gate is applied to get rid of sidelobes, air fluctuations, and other obstacles. The echo signal received by the ultrasonic sensor passes through an Analogue-to-Digital converter (ADC) followed by a laptop where the data files are saved for signal processing. The latter includes recalculation of the received signals to dB scale by the receiver calibration curve, path loss compensation and signal averaging in time in 2ms (by sliding window algorithm).

Figure 4.2: Radar and ultrasonic front-ends
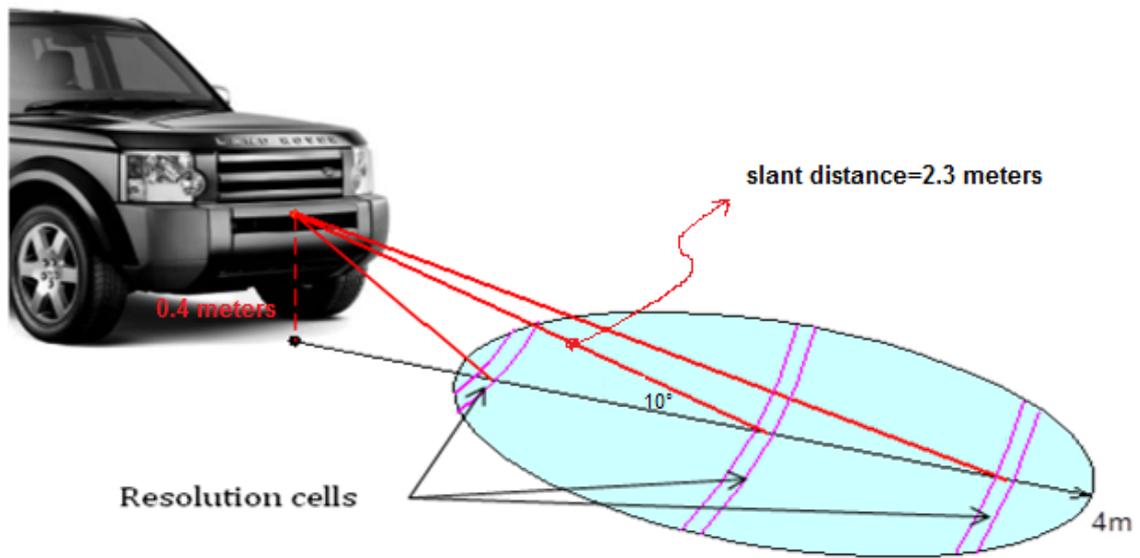
Figure 4.3:Surface identification infront of a vehicle.

4.3 Hardware Description

Some of the components used for the experimentation are available off the shelf, while other components were developed and created from scratch in the MISL laboratory. Below is a general description of each component.

WBH1-18 S horn antennas

The two antennas operate at the frequency band 1-18 GHz. The WBH1-18S, shown in figure 4.2, has an aperture size just over a quarter wavelength at the lowest frequency. They are considered as one of the lightest wide band 1 GHz horn in the market. Figure 4.4 shows the measured beamwidth of the WBH1-18S. More details about gain, VSWR and radiation pattern can be found in the data sheet attached in appendix D. According to figure 4.4, the E plane beamwidth at 3 GHz is 75°, at 5.8 GHz is 60°, and at 18 is 35°.
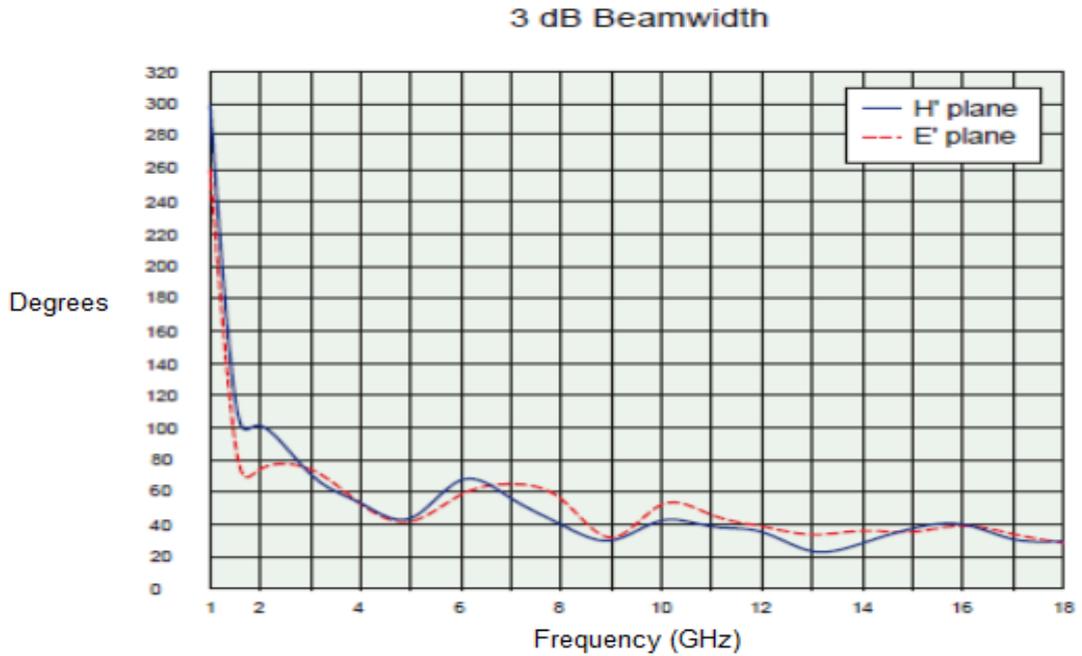
Figure. 4.4: WBH1-18S antenna's beamwidth in terms of frequency at the H and E plane.

Figure 4.5 shows the 24 GHz narrowband antennas, developed by Dr Edward Hoare. The antennas have a beamwidth of 60°, and bandwidth of 200 MHz.

Ultrasonic Sensor

The ultrasonic sensor developed at MISL by Dr Vladimir Sizov is shown in figure 4.2 (Blue Box). The sensor is based on SRF08 pulsed ultrasonic range finder. SRF08 is a set of 36 registers where only the first three can be written at. These three registers are used as:

1. Command register at Location 0 is used to start the ranging session. The default time for completion of ranging is 65 mS.

2. Max Gain Register at Location 1 is used to set the maximum gain. This is used to allow firing the sonar more rapidly than 65mS.
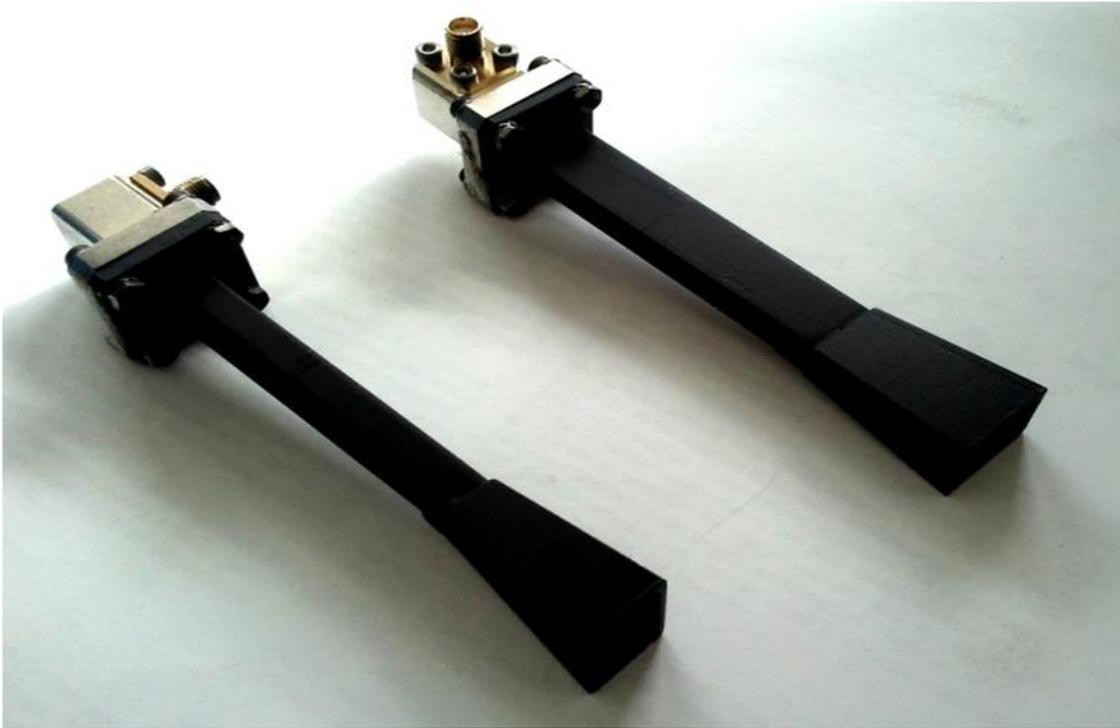
Figure. 4.5: 24 GHz antennas.

3. Range Register at Location 2 is set to 11 meters by default. SRF08 is actually capable of sensing 6 meter. Reducing the range is possible by writing into location

Locations 2 to Location 35 indicate echo's from distant objects. A zero indicates that no objects were detected. The 16 bit register gives the range in inches, cm, or the flight time in uS.

Figure 4.6 (a) shows the SRF08 board. The SDA and SCL should have a pull-up resistor to + 5v.Figure 4.6 (b) shows the beam pattern of the SRF08. The pattern is conical with a fixed beamwidth 45°. The echo is measured at $f_0 = 40$ kHz (wavelength 8.6 mm), while the signal

bandwidth and range resolution are 490 Hz and 0.35 m respectively. The power output of the ultrasonic burst is 100-150 mW.
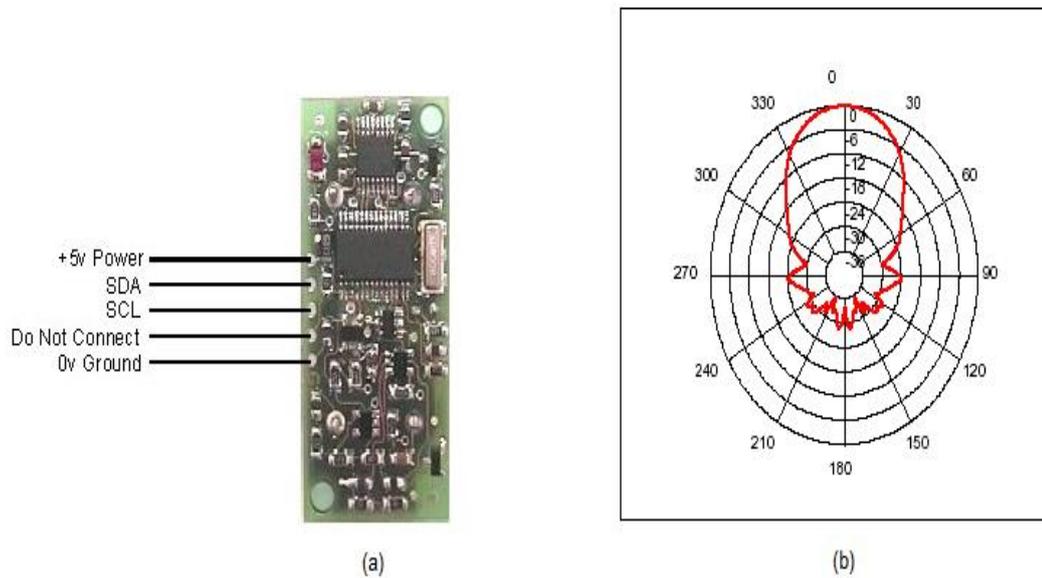
Figure 4.6:   (a) Range finder SRF08 board,   (b) SRF08 Beam patteern [92].

Detailed specification of SRF08 pulsed ultrasonic range finder is given in [92].

Network Analyzer Agilent Fieldfox N9918A

N9918A Analyzer integrates grand microwave capabilities in a single and compact instrument (figure 4.7). It also operates at wide temperature range ($-10°$ to $+55°$C). This eased the portability of the experiments on different surface and under all the circumstances and weather conditions. The Analyzer which can work as Spectrum analyzer, Cable and antenna analyzer, and Vector network analyzer (VNA) is very straight-forward to fully calibrate (full 2-port). The device was mainly used as a VNA for measuring transmission and reflection measurement (T/R), or S11 and S21, with magnitude and phase.

The time domain option allows computing the inverse Fourier Transform of the frequency response into time domain signal. Cable discontinuities, connecter mismatches, and any unwanted responses can be removed by time domain gating.

Figure 4.7: Network Analyzer Agilent Fieldfox N9918A

The wooden platform shown in fig 4.2 was designed at the MISL group to allow mounting the antennas and the sensor at the desired height and angle of incidence. It also allows changing the antenna's polarization manually.

4.4 Experimental Data Collection

4.4.1 Data Collection Method

AS mentioned before, the outdoor experiments took place in a Jaguar Land Rover site in Gaydon, and in the car park and some other locations around Pritchatts road at the University of Birmingham. Figure 4.8 shows the area of the car park and JLR site, respectively. Figure 4.9 shows the surfaces tested on these sites.

(a) Car park and sites around Pritchatts road.
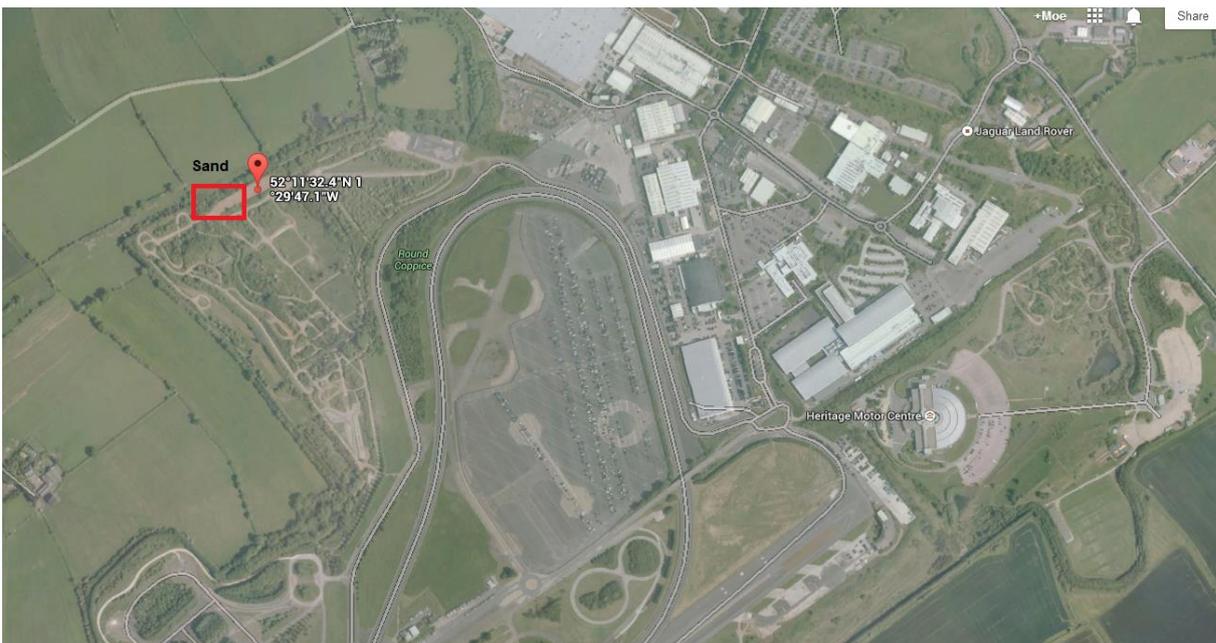
(b) Sand site in Gaydon,

Figure 4.8: Google maps of the testing sites.

Tests have been completed on smooth and rough asphalt, grass, gravel, and sand under the

same conditions (dry weather). Few experiments were conducted at the same location during

different weather conditions to study the effect of the cover (snow, ice, etc.) on the reflection

from surfaces (Figure 4.9). Since the goal of this project was dedicated to the feasibility of surface classification , the effect of weather conditions and vehicle movement was proposed for future work. Neither the platform nor the hardware provided at this stage was feasible to collect more data with different conditions. Following the timetable of Jaguar Land Rover, it was agreed that this information is crucial in order to appreciate experimental but will be postponed till JLR provides a demonstrator to analyse the external conditions.

Asphalt   Grass   Gravel

Snow   Ice   Sand

Figure 4.9: On-road and off-road surfaces tested.

4.4.2 Acquiring Data

In order to acquire data through antennas and the ultrasonic sensors, a Matlab program has been developed. The program consists of two parts: configuring ultrasonic sensor, and configuring antennas. Data files from the ultrasonic were saved on the disk of the laptop, while data files from the antennas were saved on the mini card of the VNA. The Matlab program developed for data acquiring was modified to account for different frequencies. Below are the steps of the data acquisition procedure completed after connecting the ADC

and the VNA to the laptop via USB. Once "Data_acq_Ac_Radar_15.m", which can be found in Appendix E, is run, it carries out the following:

1) The software prompts the user to enter the name of the measurement file.

2) The values of sampling frequency, acquisition time, number of samples, and sampling rate are set for the ADC channels. According to Nyquist, the minimum sampling rate required to avoid aliasing is equal to twice the highest frequency. According to the specification of the ADC, the sampling frequency is set as 10 KHz which agree with the Nyquist sampling theory. The number of frames acquired is set to 10. The number of samples per frame is given by:

$$\text{Number of samples} = T.f_s \tag{4.1}$$

where T is the ADC acquisition time for 1 frame given by 40 ms, and f_s is the sampling frequency. Thus, the number of samples in ADC window is given by 400 samples per frame. The maximum acoustic range is set to 6.8 meters.

3) The ultrasonic sensor is run, and the echo is collected over 10 realizations (number of frames). This data is averaged over 10 frames and transferred to dB then written to the hard disk as a file.dat .

4) Next step is establishing a TCP/IP connection between the antennas and the VNA. A TCP/IP object is created with an address and port number which are used to request a VNA ID.

76

5) The centre frequencies and the bandwidths are set for the antennas (5.8 GHz with 150 MHz bandwidth, and 18 GHZ/ 24 GHz with 200 MHz bandwidth). For each polarization, measurements at the mentioned frequencies (which can be set to more than two frequencies) are launched and the reflected signal is save to s2p.file on the mini card of the VNA.

The procedure takes around 30 seconds for one polarization. This procedure was repeated for each single measurement on each location for the three polarizations (vertical-vertical, horizontal- horizontal and vertical-horizontal).

4.4.3 Visualizing Data

After acquiring data, the Matlab program "Data_Import_NewVNA_5_Graphs" was developed to visualize it. This program, shown in appendix E, performs the following steps:

1) It prompts the user to enter the file name.

2) It reads the frequency, amplitude power (in dB), and phase from the VNA file.

3) It sets up the frequency range, time range, number of data points.

4) It applies a Gaussian window with standard deviation of 0.5 on the linear amplitude power to reduce the sidelobes.

5) It restores the signal into time domain using the amplitude, phase and time range.

6) It applies Hilbert Transform to create an analytical signal which forms the signal envelope (mathematically, the envelope of the signal is defined as the magnitude of the analytic signal), and then downsample the signal by smoothing.

7) It plots the amplitude power in dB versus distance (figure 4.10)
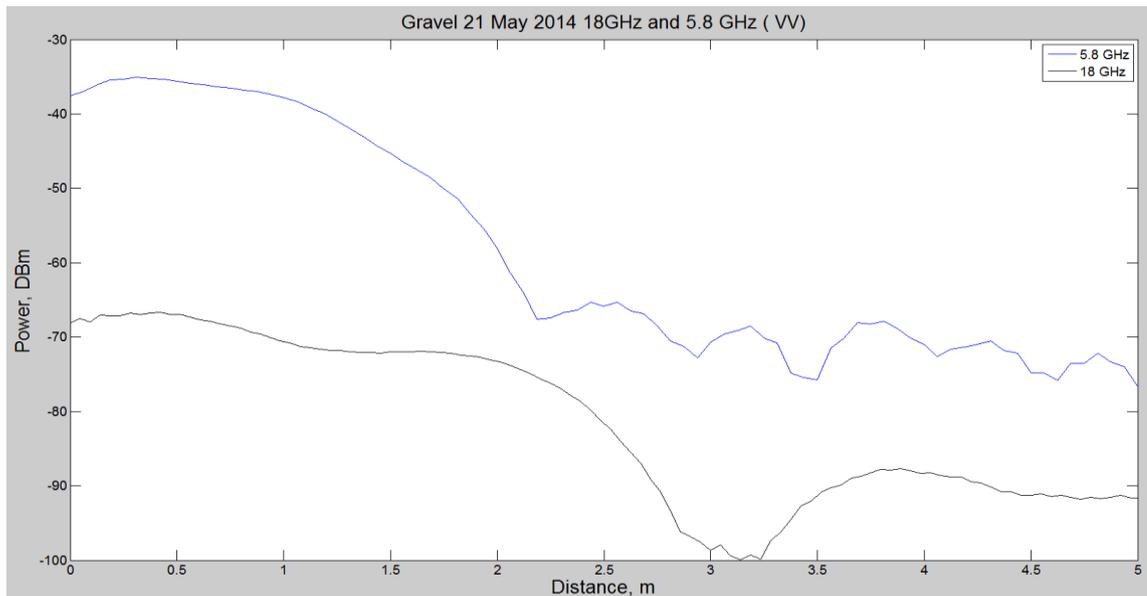
Figure 4.10: Power dB versus distance at a18 GHz and 5.8 GHz for dry gravel.

Two Matlab programs were mainly developed to extract the amplitude and the standard deviation of the reflected signal for different polarizations and store them in an Excel file to be later added to the database. The two Matlab programs used are also shown in appendix E:

a) Data_Import_Radar_200.m

b) Data_Import_Acoustic_200.m

"Data_Import_Radar_200.m"   performs   the   same   first   6   steps   of "Data_Import_NewVNA_5_Graphs" and, on top; it calculates the average or mean of the radar signal amplitude and the standard deviation over a time gate (between 1.5 meters and 4 meters). This is done in several iterations to compute the parameters for different files with different polarizations. Table 4.1 shows the output file when computing the parameters for

78

dry gravel with frequencies 5.8 and 18 GHz out of 40 different spots on the 2 mentioned locations. The columns in the table represent the frequency, average amplitude for VV polarization, average amplitude for VH polarization, average amplitude for HH polarization, relative standard deviation for VV polarization, and the file number.

"Data_Import_Acoustic_200.m" on the other hand, performs the same steps on the ultrasonic files. It calculates the average or mean of the ultrasonic signal amplitude and the standard deviation over a time gate (between 1.5 meters and 4 meters). Figure 4.11 (a) is a plot of the amplitude power in dB versus the sensor range for one of the measurements performed in this research. It demonstrates an ultrasonic signal reflected from dry gravel surface. A normalised round trip path loss compensation (explained in chapter 5) was applied to the signal. The graph showed that the reflection has an average amplitude of - 10 DBm over the range 0 to 7 meters. Figure 4.11 (b) shows the output file when computing the parameters for dry gravel at 4 different locations. The columns in the figure represents the average power in mv, average power in dB, power above threshold, duration above threshold, relative standard deviation, and the file number. Further explanation about each parameter will follow in the next chapters.

	A	B	C	D	E	F
1	FREQ_BW	VV	VH	HH	STD_VV	FN
2	5.8	1.8328	0.091474	0.5437	1.367	125
3	18.2	0.16317	0.073347	0.10473	0.91844	125
4	5.8	1.9734	0.077999	0.55522	1.3651	126
5	18.2	0.19789	0.082804	0.087176	0.82919	126
6	5.8	1.68	0.36852	0.74973	1.3518	127
7	18.2	0.25535	0.13289	0.21344	0.60443	127
8	5.8	1.6322	0.38699	0.7678	1.3605	128
9	18.2	0.2568	0.12034	0.22474	0.60276	128
10	5.8	1.6535	0.25927	0.2957	1.4292	129
11	18.2	0.20973	0.084544	0.18147	0.99809	129
12	5.8	1.7532	0.2605	0.26172	1.4127	130
13	18.2	0.23307	0.080516	0.19018	1.0031	130
14	5.8	1.7736	0.17935	0.63943	1.4546	131
15	18.2	0.13076	0.14915	0.098841	0.41008	131
16	5.8	1.7667	0.23584	0.71344	1.4282	132
17	18.2	0.11819	0.15939	0.10599	0.36511	132
18	5.8	1.4863	0.35202	0.59291	1.4252	133
19	18.2	0.18838	0.038379	0.12546	0.78528	133
20	5.8	1.6167	0.34	0.6155	1.4178	134
21	18.2	0.1951	0.032584	0.12816	0.74211	134
22	5.8	2.1652	0.3381	0.6011	1.3549	135
23	18.2	0.26967	0.047365	0.14997	0.79625	135
24	5.8	2.0215	0.35408	0.51625	1.3717	136
25	18.2	0.25593	0.049362	0.16526	0.78065	136
26	5.8	2.0324	0.16019	0.5482	1.3385	137
27	18.2	0.21307	0.11647	0.10137	0.66071	137
28	5.8	1.9886	0.30894	0.60964	1.3369	138
29	18.2	0.22641	0.12968	0.10929	0.75911	138
30	5.8	1.4646	0.25729	0.71325	1.3567	139
31	18.2	0.11712	0.088711	0.086204	0.43256	139
32	5.8	1.4963	0.37494	0.72293	1.3376	140
33	18.2	0.10003	0.088722	0.078939	0.51706	140
34	5.8	2.086	0.30715	0.23498	1.376	141
35	18.2	0.23177	0.15992	0.22432	0.61186	141
36	5.8	2.1657	0.25182	0.24363	1.3602	142
37	18.2	0.23587	0.16298	0.2223	0.6286	142
38	5.8	1.3897	0.41094	0.56639	1.3758	143
39	18.2	0.16402	0.13648	0.061083	0.66978	143
40	5.8	1.3107	0.34579	0.55901	1.3719	144

Table 4.1: Output file of "Data_Import_Radar_200.m
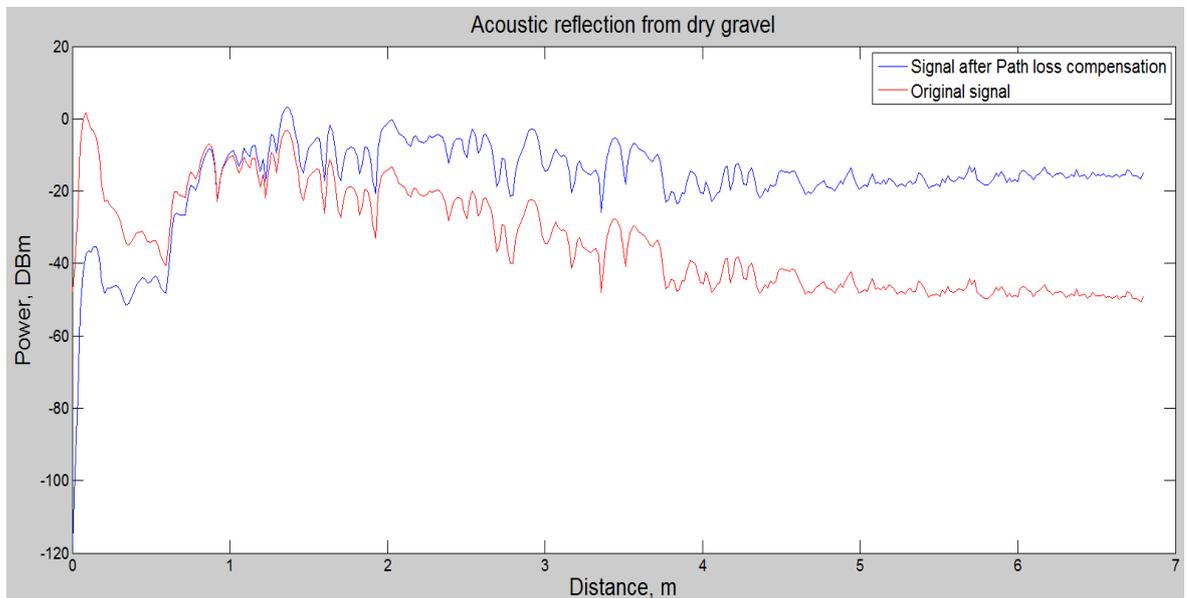
(a)

A	B	C	D	E	F
Power mv	power dB	Duration above thres.	Power above thresh.	St deviation	File_No
1.621	-28.121	1.802	0.32561	0.13845	125
1.6228	-28.05	1.819	0.32656	0.13243	126
1.5953	-29.148	1.802	0.23817	0.12872	127
1.6083	-28.628	1.972	0.28915	0.1253	128

(b)

Figure 4.11: Output file of "Data_Import_Acoustic_200.m"

4.5 Preliminary Study

In this section, the frequencies used for both of the radar and ultrasonic sensor will be discussed. Moreover, investigation of the effect of grazing angle, exhaust gases, vehicle movement, and weather conditions on the measurements will take place.

4.5.1 Signals and Frequencies

2002 FCC Report and Order authorized the unlicensed use of UWB in the frequency range from 3.1 to 10.6 GHz. The frequency band allocated for short-range automated radars falls between 21.65 and 26.65 GHz [93]. Our goal is to use the best of these frequency ranges intended for automotive and unlicensed use.

	Radar						Ultrasonic
Frequency	5.8 GHz	9 GHz	9 GHz	18 GHz	18 GHz	24 GHz	40 kHz
Bandwidth	150 MHz	200 MHz	3 GHz	200 MHz	3 GHz	200 MHz	Pulsed
Wave length	5.7 cm	3.3 cm	3.3 cm	1.7 cm	1.7 cm	1.25 cm	0.85 cm
Range Resolution	1.0 m	0.75 m	0.05 m	0.75 m	0.05 m	0.7m	0.034 m
Beamwidth $2\Theta_b$*	60°	35°	35°	30°	30°	30°	55°

Table 4.2: Frequency ranges and characteristics of the signals.

We measured backscattering at 4 frequencies: 5.8, 9 and 18 GHz. These frequencies were used with the previously mentioned horn antennas as its frequency range is up to 18 GHz. 24 GHz was later used with the set of dual- polarized antennas, automatically switched. Table 4.2 shows the frequencies, bandwidths, wavelengths, range resolution, and the beamwidths used by the radar and ultrasonic sensors.

Range resolution is the ability of the radar to distinguish between different targets at different ranges. It defines the minimum range separation of two targets in order for the sensor to resolve them as separate targets. Range resolution depends on the width of the transmitted pulse or what we call Bandwidth. The signal bandwidth used was 150 MHz at 5.8 GHz central frequency, 200 MHz and 3 GHz bandwidths at 9/ 18 GHz frequencies, and 200 MHz at 24 GHz. The wider the bandwidth, the higher the range resolution; although 3 GHz bandwidth is not allowed for unlicensed use but, for research purposes, we used it at some of the (9 and 18 GHz) measurements.

The linear frequency modulation "Chirp" is used by our radar system. Chirp or "Compressed High Intensity Radar Pulse" is a signal in which frequency increases or decreases with time. The magnitude of the frequency response is constant with a value of 1, while the phase is a parabola [94].

The chirp slope represents the rate of change of frequency. It is constant in the linear case and can be calculated as:

$$\Delta f = \frac{f_b \cdot T_b}{T_p} \tag{4.2}$$

where f_b is the beat frequency, T_b is the chirp time or signal duration, and T_p is the round-trip time to the target and back. T_p can be written as :

$$T_p = \frac{2R}{c} \tag{4.3}$$

Where R is the target range, and c is the speed of light in vacuum. Substituting (4.3) into (4.2) gives:

$$\Delta f = \frac{f_b \cdot T_b \cdot C}{2R} \tag{4.4}$$

Using (4.3), the range resolution δR can be calculated for a frequency resolution δf [95]:

$$\delta R = \frac{T_b \cdot C}{2\,\Delta f}\, \delta f_b \tag{4.5}$$

Since the frequency resolution δf_b defines the resolution bandwidth of a signal between its 3dB (half power) points, it falls within the $1/T_b$ region centred at f_b . Substituting $1/T_b$ in (4.4) gives the equation of range resolution in terms of the total transmitted bandwidth:

$$\delta R \;=\; \frac{c}{2\,\Delta f} \tag{4.6}$$

This equation is equivalent to the classical pulsed radar range resolution equation where $\tau = 1/\Delta f$, and it was used to calculate the range resolution in table 4.1. For 3 GHz it is equal to 5 cm, for 200 MHz - 75 cm, and for 150 MHz - 100 cm.

For the Sonar sensor, range resolution was calculated using the following equation:

$$R = \frac{l \,.\, C_{sound}}{2} \tag{4.7}$$

Where l is the pulse length, and C_{sound} is the speed of sound in air approximately 346m/S at 24 degrees Celsius. Since the SRF08 sends out an 8 cycle burst of ultrasound at 40 KHz, pulse length is given by :

$$l= \frac{1}{frequency} \,.\, 8 \tag{4.8}$$

with l equals to 0.25 ms, the range resolution R of the sonar is equal to 3.4 cm.

4.5.2 Influence of Vehicle Movement

The standard Range Rover is larger in height than the average car. Its standard ride height is 1,835 meters. The bumper of the car is located at a height of 0.4 to 0.65 meters (Figure 4.12). Therefore, the sensors must be located at the most convenient height that belongs to the range (0.4 - 1.835 meters). For practical implementation, the antennas were mounted on the bumper at a height of 0.45 meters.

Figure 4.12: Range rover top and bumper height.

All the measurements were obtained in a stationary condition using the "Stop-and-go" technique. Standing still, Sensors fire signal towards the surface and captures the backscattered signal. The sensors are then manually moved to face another part of the surface for another measurement. For practical realization of surface recognition, it is essential to analyse the influence of vehicle movement on the performance of the antennas and the sonar sensor. In real road conditions, several factors influence the measurement accuracy:

1.   Vehicle's pitch and roll angle.

2.   Vibrations of the transceiver.

3.   Reflections from objects on the roadside.

4.   Reflections from air fluctuations.

5.   Uneven absorption of ultrasonic signals by air currents.
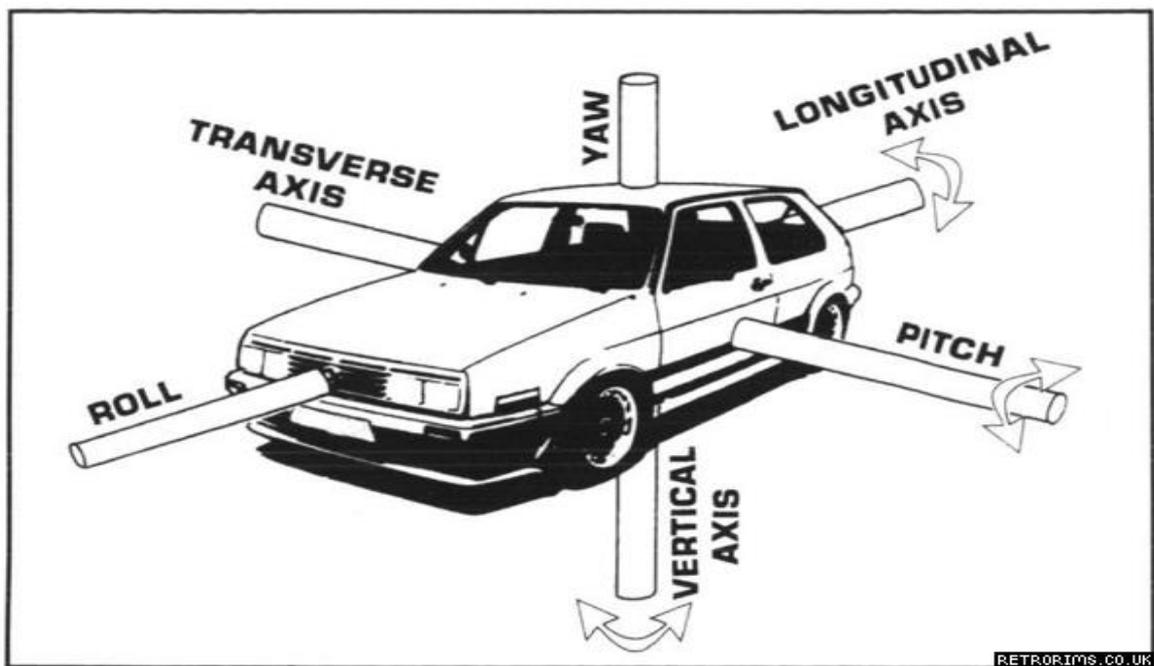
Figure 4.13: The Pitch, Roll and YAW axis **[96]**

These factors disproportionately affect the measurement's accuracy. Motion around the transverse axis is called pitch and it's a measure of how far is the car's nose tilted up or down (figure 4.13) . The pitch angle is an important factor since any change it is subjected to will lead to a change in the grazing angle of the insonified area. Figure 4.14 shows the influence of the vehicle pitch on the echo signal power at 2 meters distance for different beamwidths as obtained in our labs. The analysis shows that the change of grazing angle caused by the change of the pitch $\theta_p = \pm 2.5°$ leads to a change of backscattering coefficient of less than 0.5 dB. This change can be neglected and, therefore, influence of pitch angle can be neglected in most of the practical cases. Nevertheless, including the vehicle's position data,

received from the internal sensors, in the processing algorithm can be used to reduce the pitch error further.
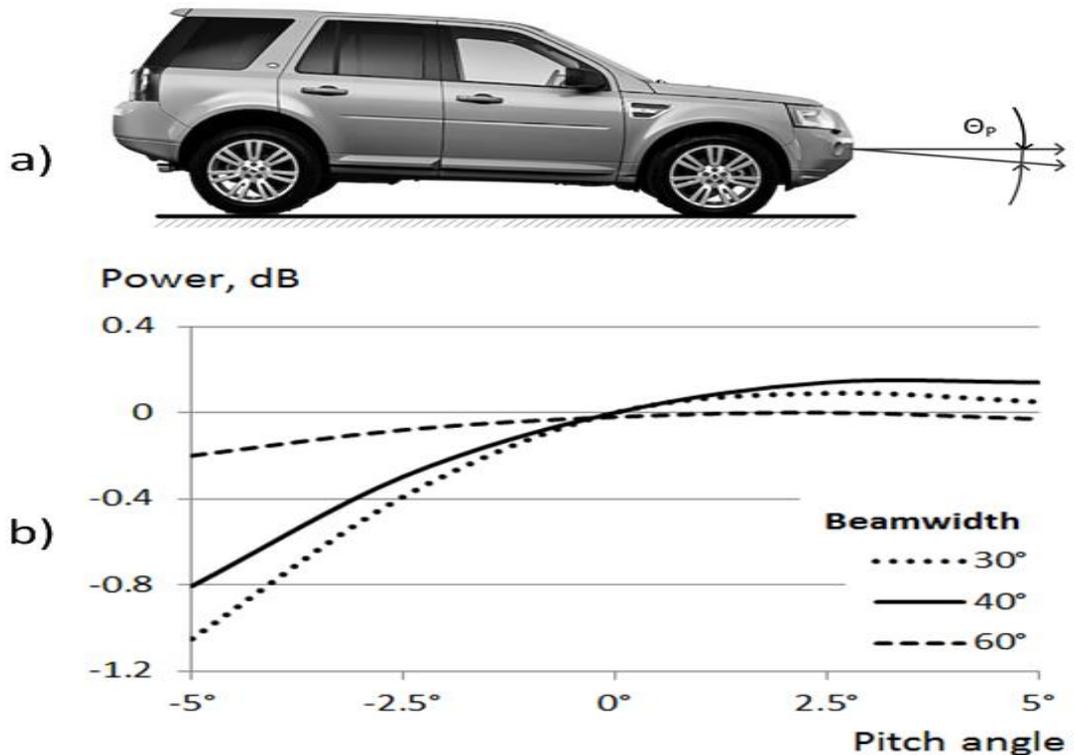
Figure 4.14: (a) pitch angle θ_P
(b) Influence of the vehicle pitch angle θ_P on the echo signal power at 2 m distance for different beamwidths.

Motion around the longitudinal axis is called roll. Roll angle is controlled by the anti-roll bar of the car. On the other hand, Yaw is the motion around the vertical axis and it is controlled by the suspension components. The roll and yaw factors do not have a significant effect on the measurement accuracy in the widespread case of axial symmetry of the beam as they don't impose any change in the grazing angle of the insonified area.

4.5.3 Effect of Exhaust Gases

Exhaust gases are the gases emitted through the exhaust pipe as a result of the combustion of fuels. These gases might affect the accuracy of measurements when it interferes with the echo signal.

At the outlet of the exhaust pipe the temperature of the gases reaches 150 to 200° C, but it cools down rapidly afterwards. After repeating fifty measurements on the same surface with or without running engine, the results averaged over these measurements showed that the power of the reflected signal will decrease 4 dB at a distance of 2 meters if the exhaust gases are directed towards the measurement object (surface). Exhaust gases has significant influence on sonar signature which leads to increasing variance of measured parameters, i.e., increasing the data spread from one measurement to another. The error caused by exhaust gases may significantly deteriorate the reliability of surface recognition. Results and discussion will be shown in the 6th chapter.

4.5.4 Effect of Grazing Angle

Incidence angle is the measure of deviation from the straight. In optics, incidence angle is considered to be the angle between the ray incident on a surface and the line perpendicular to the surface called normal. In some cases it is useful to use grazing angle instead, which is the angle between the ray incident and the surface (Figure 4.15). Grazing angles are typically used when the beam is nearly parallel to the surface.
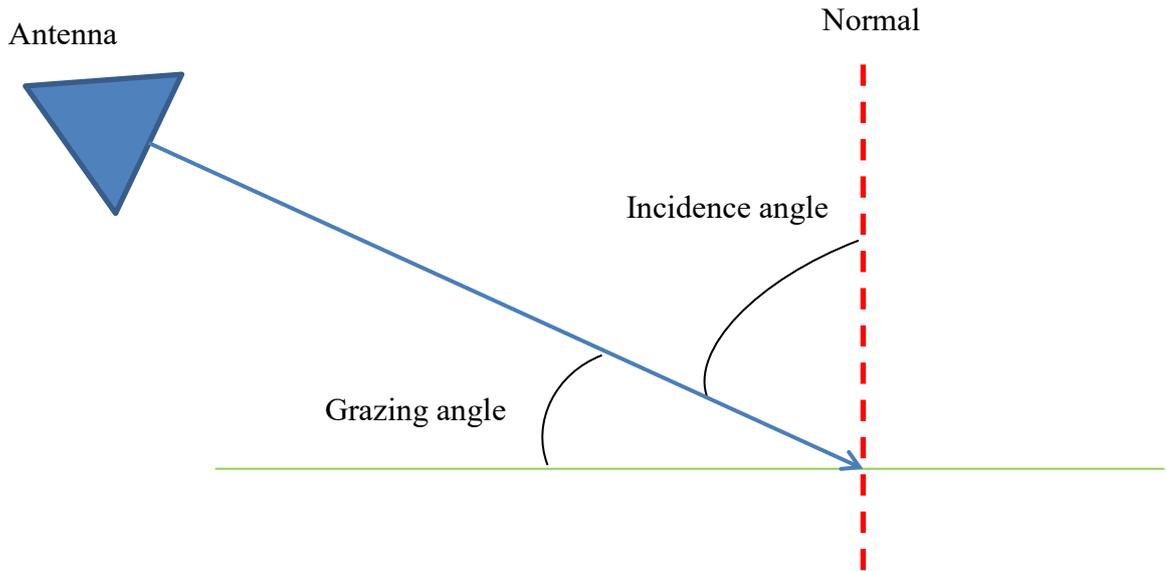
Figure 4.15: Incidence and grazing angles.

As mentioned before (previous chapters), the amplitude of the sonar signal and radar signal depend on the grazing angle. Grazing angle affects the directivity of the transmitter and receiver pair and the backscattering coefficient.

Figure 4.16 illustrates the dependence of backscattering coefficient of the ultrasonic sensor on the grazing angle for surfaces of different roughness. The simulation was developed based on the equation of backscattering coefficient:

$$\langle|\rho_b|\rangle = \sqrt{\frac{\pi}{S} \frac{\eta \, \rho_s}{2k \, cos^3 \, \theta}} \; e^{- \, tan^2\theta \, . \, \eta^2/2} \tag{4.9}$$

Where S is the area of reflecting surface and $\eta = {}^{T_s}/_{\sigma_s}$ is the roughness parameter. ρ_s is the equivalent reflection coefficient for scattering in the specular direction. As expected, smooth surfaces(red dots) has a big specular component around low incident angles, while rough

surfaces (blue stars) shows higher reflections at bigger incident angles. The coherent component of the echo signal depends on the surface roughness around the specular direction since it only occurs in the direction normal to the surface. The diffuse component depends on the surface roughness and the angle of incidence [41]. In the majority of terrestrial communications systems, reflection occurs at very small grazing angles. In our system, we are limited by the height of the antennas as the bumper of the car varies between 40 and 65 cm in height (figure 4.12).
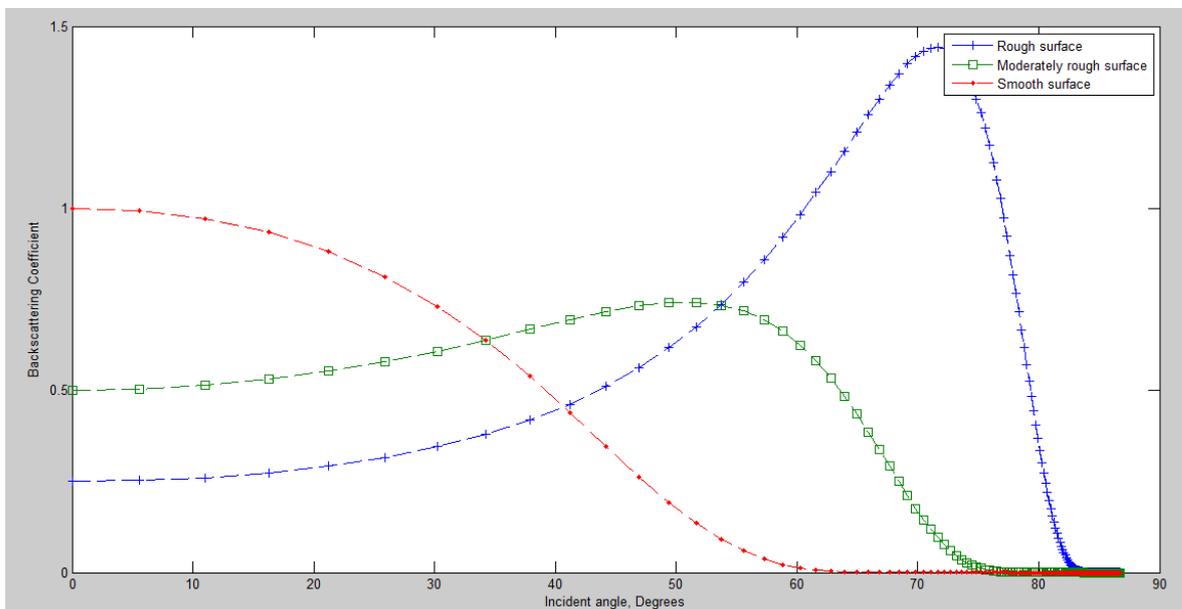
Figure 4.16: Simulation of Power of received pulses (dB) vs. incident angle (degrees).

The requirements are to cover the area from 5 to 10 meters in range in front of the vehicle, but for preliminary studies and due to the equipment limitations, we are scanning the range between 0.7 and 6 meters ahead of the vehicle. At the initial set of experiments, we have changed the grazing angles between $10°, 15°, 45°$ and $60°$. As the height of the antennas and beamwidth is fixed, changing the grazing angle changes the range of the scanned area. Lower grazing angles lead to probing the surface at bigger range from the antennas. Higher ones lead to probing a short-range distance from the antennas. The aim is to choose the

90

optimum grazing angle which satisfies the practical implementation of surface identification. The surfaces we are probing are considered rough surface so we don't need low inclination angles (high grazing angle) which is usually used in the presence of a major specular component. In contrary, we need a low grazing angle to probe the surface at the maximum possible distance as it is necessary to alert the driver as soon as possible. Nevertheless, air fluctuation scatters the ultrasound and causes an interference signal which increases by increasing the distance. Therefore, the reflected power must be sufficient for reliable surface recognition by both sensors, which leads to the restriction of using low grazing angles. Due to the scope of the research, a grazing angle of $10°$ was chosen as it is the lowest grazing angle we can use to get the sufficient power needed for a reliable surface identification. An experimental evaluation is suggested for the future work in order to optimise the value of the grazing angle employed.

4.6 Effects of Weather Conditions

Weather conditions are the main challenge to be considered in outdoor remote sensing. Attenuation due to weather conditions is a vital factor to assess the performance of the sensors. As the Electromagnetic wave propagates through the atmosphere, it is absorbed by atmospheric gases. The gases that forms the atmosphere are 78% Nitrogen (N_2), and 21% Oxygen (O_2), and 1% Argon (A_r). Other gases exist with smaller amounts like H_2O and CO_2. Water also exists in the atmosphere and its percentage ranges between 0 and 7 where the higher the percentage, the higher is the density of air. Haze particles, fog, cloud, and rain constitute of Aerosols. Aerosol is defined as a

gaseous suspension of fine solid or liquid particles. Atmospheric aerosols scatter and absorb electromagnetic waves. These actions cause a wavelength dependant attenuation of the waves leading to degradation of the performance of the sensors. The presence of water in the atmosphere can be seen as raindrops and snow. It also exists in the form of clouds or fog at the lower part of the atmosphere [97]. The attenuation of the radar and ultrasonic waves caused by clouds and raindrops will be discussed in the upcoming sections.

Attenuation by Haze

Haze is an atmospheric phenomenon in which small particles of tiny salt crystals, fine dust, and products of combustion are dispersed throughout the aerosol causing obscurity of the sky. The radii of the particles are known to be very small (around 0.5 µm). Haze attenuates visible and IR sensors, but its attenuation decreases with increasing wavelength. Thus, haze is transparent to RF radiation because of the tininess of the haze particles as compared to the RF radiation wavelength.

Attenuation by Clouds

Clouds consist of water droplets which cause attenuation of EM radiation due to absorption and scattering. The radii of cloud water droplets range between 1 µm and 30 µm. This radius is much smaller than the RF radiation wavelength. Therefore, Clouds are partially transparent to RF radiations and are expected to attenuate visible and IR radiation much more than they will for RF radiation [98]. Moreover, clouds will not be considered as a severe condition as the short range used for road surface identification doesn't encounter clouds.

Attenuation by Fog

Moisture can condense on the small particles of salt crystals and dust in regions of high humidity. It then forms large drops acting as condensation nucleus. Fog is formed when these nucleus grow into water droplets or ice crystals with radii of 1 µm [99]. There are two kinds of fog: coastal and inland. Fog changes it characteristics with time but it is very similar to clouds except that it contacts the ground.

Attenuation coefficient by small water droplets, in the λ= 0.5 cm to λ= 10 cm wavelength region, can be written as the following:

$$\gamma \ = \frac{0.438 \ \ M}{\lambda^2} \quad dB/Km \qquad\qquad (4.10)$$

where λ is the wavelength in cm, M is the liquid water content which ranges in clouds between 1 to 2.5 gm/ m^3. The maximum liquid water content of fog is M= 1 gm/m^3 [100]. This equation is valid for both clouds and fog with small water droplets (diameters of 10 µm to 50 µm). Figure 4.17 shows the attenuation coefficient by water droplets as a function of wavelength at temperature of $18°$ C. It can be observed in figure 4.17 that the attenuation is really low for any sensor of wavelength greater than 0.5 cm. The sensors used in the experiments have wavelengths starting at 1.25 cm so the attenuation by fog is less than 1 dB. Since the index of refraction of water is temperature dependant, table 4.3 shows the correction factor of the attenuation coefficients for other temperatures where γ(t) = Ø(t).γ ($18°$C$)$ [101].
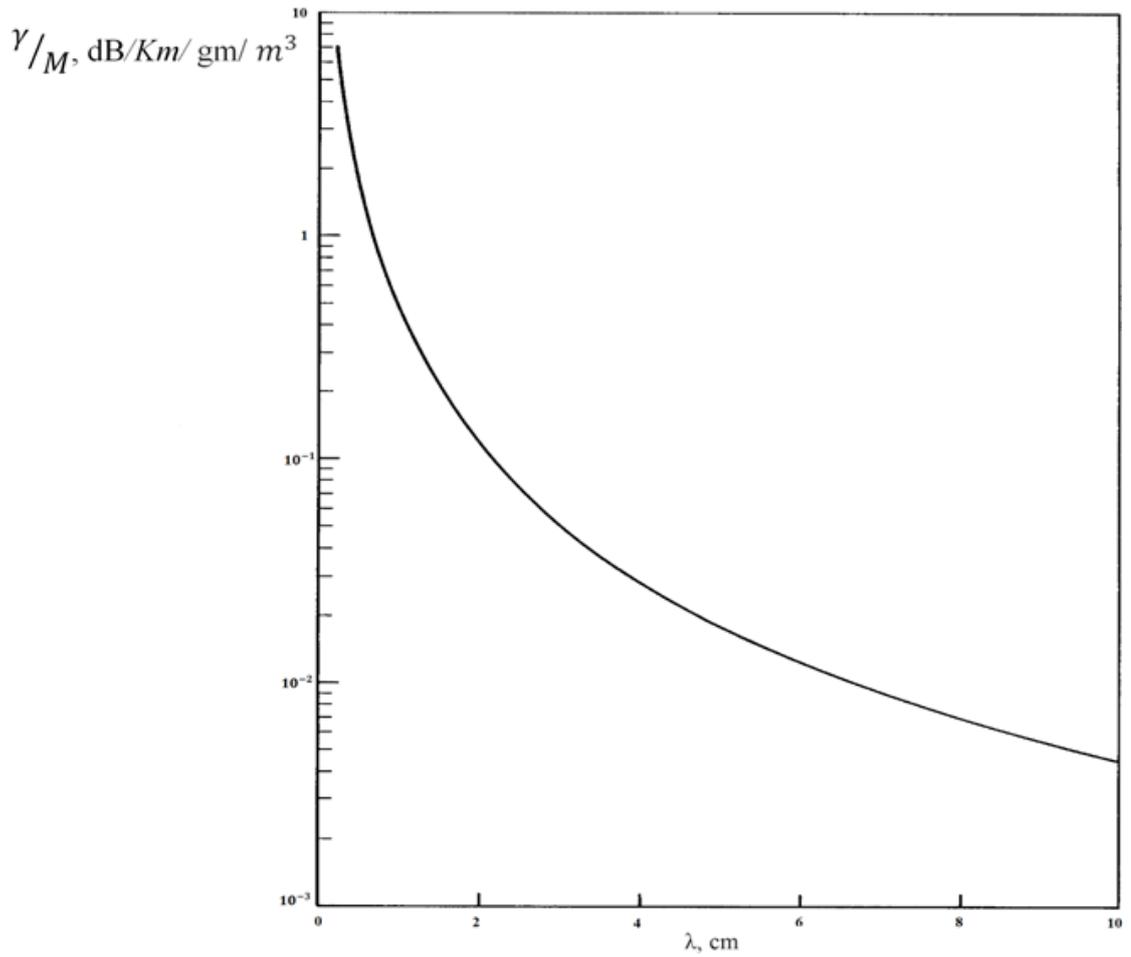
Figure 4.17: Attenuation coefficient by small droplets of water (clouds or fog) at $18^°$ C [101].

λ ,cm	Ø(t)					
	$0^°$C.	$10^°$C.	$18^°$C.	$20^°$C.	$30^°$C.	$40^°$C.
0.5	1.59	1.20	1.0	0.95	0.73	0.59
1.25	1.93	1.29	1.0	0.95	0.73	0.57
3.2	1.98	1.30	1.0	0.95	0.70	0.56
10	2.0	1.25	1.0	0.95	0.63	0.59

Table 4.3: Correction factor Ø(t) for attenuation coefficient at different temperatures [101].

Fog is usually characterized by its visibility. If the visibility given in meters is greater than 1 km, the atmosphere is considered hazy, otherwise it is foggy. The visibility of coastal fogs is given by the following empirical relationship [101]

$$V_2 = 59.4 \, M^{-0.7} \tag{4.11}$$

where V_2 is given in meters, and M is the liquid water content (less than 1 gm/ m^3). The one for inland fogs is given by equation (4.10) [102]:

$$V_2 = 24 \, M^{-0.65} \tag{4.12}$$

Figure 4.20 shows the relationship between the water content and visibility for both coastal and inland fog. The lower the liquid water contents the better the visibility.
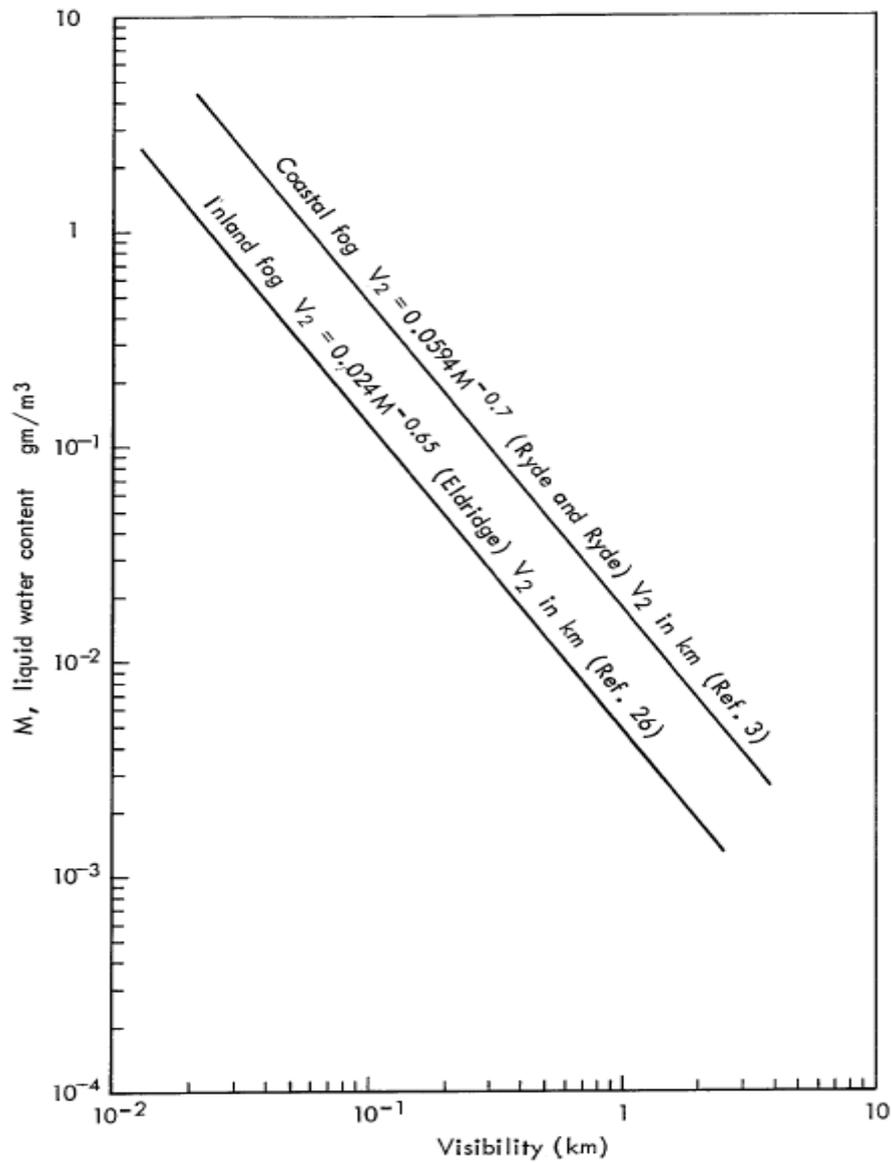
Figure 4.18: Relationship of visibilty and liquid water content in inland and coastal fog[99].

Using figure 4.18 and equations (4.11) and (4.12), the values of attenuation by inland and

coastal foggy atmospheres can be calculated as a function of frequency. Figure 4.19 shows

the attenuation by fog in dB/km for different visibilities in meters as a function of frequency.

The attenuation is calculated at $18°$ C. Attenuation decreases with increasing wavelength,

i.e., decreasing frequency. As shown in figure 4.19, frequencies lower than 24 GHz has very

low attenuation at different visibilities (less than 1 dB/km loss at low visibility of 30 meters).
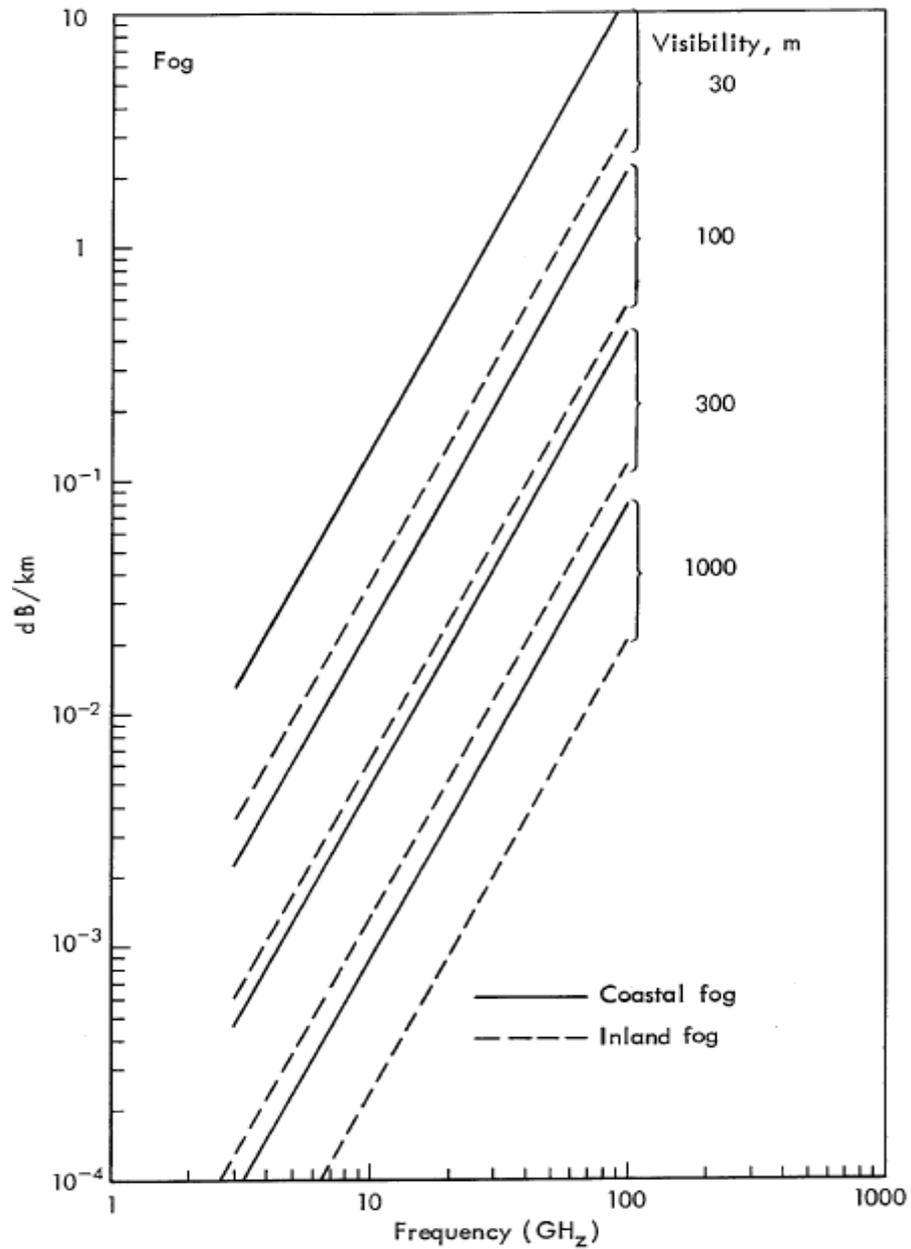
Figure 4.19: Attenuation by coastal and inland fog (dB/km) versus frequency (GHz) at different visibilty

(meters) [99].

Attenuation by Rain

Rain, as clouds, causes attenuation of EM radiation due to absorption and scattering. Nevertheless, rain has much bigger water droplet radius. Raindrop's radius is typically 0.05 cm, which is much larger than visible and IR wavelengths but comparable to the RF wavelengths. IR and visible sensors would be considered useless to operate under any heavy rain.

Attenuation by rain is considered the most severe among the other weather conditions. Figure 4.20 shows measured and calculated attenuation of RF signals (dB/km) by rain as a function of frequency at 20° C temperature [101]. At frequencies higher than 24 GHz, Attenuation is considered to be severe especially at high precipitation or rainfall rates (50 mm/hr and above).However, at frequencies less than 24 GHz attenuation decreases to an acceptable level. As shown in figure 4.20, attenuation at lower frequencies decreases to be negligible at 10 GHz. For the frequency ranges used in the experiments, only heavy rain with a very high rainfall would attenuate the signal more than 2 dB/km, otherwise any lighter rain would have an insignificant attenuation.

The attenuation is temperature dependant since it depends on the refractive index of water in addition to the radius of the water droplet. Therefore, a correction factor of the attenuation coefficients is needed for other temperatures. Table 4.4 shows the correction factor $\emptyset(t)$ for different temperature, different wavelengths, and different precipitation rate [101]. The correction factor doesn't exceed 0.1% of the overall value. The effect of temperature for small wavelengths (less than 1.25 cm) is small and can be neglected.
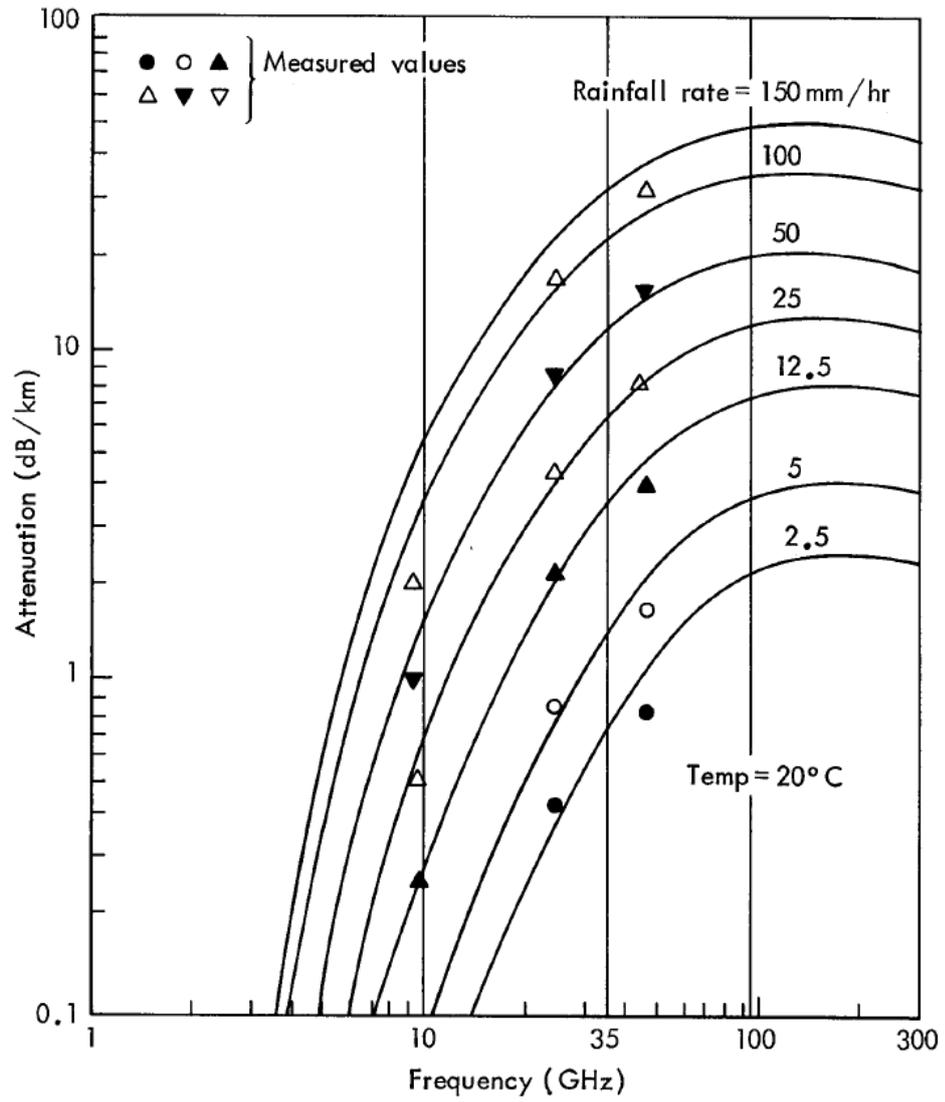
Figure 4.20: Attenuation by rain (dB/km) versus frequency (GHz) at different rainfall rates (mm/hr) and temperarute of 20° C [101].

Precipitation Rate (mm/hr)	λ ,cm	Ø(t)				
		0°C.	**10**°C.	**18**°C.	**30**°C.	**40**°C.
0.25	0.5	0.85	0.95	1.0	1.02	0.99
	1.25	0.95	1.0	1.0	0.90	0.81
	3.2	1.21	1.10	1.0	0.79	0.55
	10	2.01	1.40	1.0	0.70	0.59
2.5	0.5	0.87	0.95	1.0	1.03	1.01
	1.25	0.85	0.99	1.0	0.92	0.80
	3.2	0.82	1.01	1.0	0.82	0.64
	10	2.02	1.40	1.0	0.70	0.59
12.5	0.5	0.90	0.96	1.0	1.02	1.00
	1.25	0.83	0.96	1.0	0.93	0.81
	3.2	0.64	0.88	1.0	0.90	0.70
	10	2.03	1.40	1.0	0.70	0.59
50	0.5	0.94	0.98	1.0	1.01	1.00
	1.25	0.84	0.95	1.0	0.95	0.83
	3.2	0.62	0.87	1.0	0.99	0.81
	10	2.01	1.40	1.0	0.70	0.58
150	0.5	0.96	0.98	1.0	1.01	1.00
	1.25	0.86	0.96	1.0	0.97	0.87
	3.2	0.66	0.88	1.0	1.03	0.89
	10	2.00	1.40	1.0	0.70	0.58

Table 4.4. Correction factor Ø(t) for attenuation coefficient by rain at different temperatures and different precipitation rates (mm/hr) [101].

4.6.1 Atmospheric Absorption of the Ultrasonic Signals

The attenuation of sound by the atmosphere is a function of frequency, temperature, and humidity. There are two mechanisms in air that attenuates the sound energy: classical absorption and relaxation processes. Classical absorption represents the losses due to the friction with air molecules which results in heat generation. Relaxational processes are the absorption of sound energy in air molecules leading to the vibration and rotation of these molecules. This causes the molecules to re-radiate sound at a later time partially interfering with the incoming sound.

These mechanisms were empirically studied into an international standard for calculation [103], and showed that the absorption α at kHz frequencies is less than 1 dB/100 meters. Since SRF08 pulsed ultrasonic range finder measures echo at $f_0 = 40$ kHz, humidity and temperature are not expected to cause a severe absorption and attenuation of the sound signal. Figure 4.21 illustrates the influence of temperature and humidity on the ultrasonic echo signal at 2 m distance and pressure (101.325 kPa) of the air. The calculations are based on the calculation method for absorption of sound by the atmoshphere [105], and the power is normalized over a reference point (power at 20° C and 80% humidity).As shown in figure 4.21, increasing the temperature by 5 degrees reduces the power of the echo signal by an average of 1 dB. On the other hand, increasing humidity by 20 % leads to a signal attenuation of 1 dB.
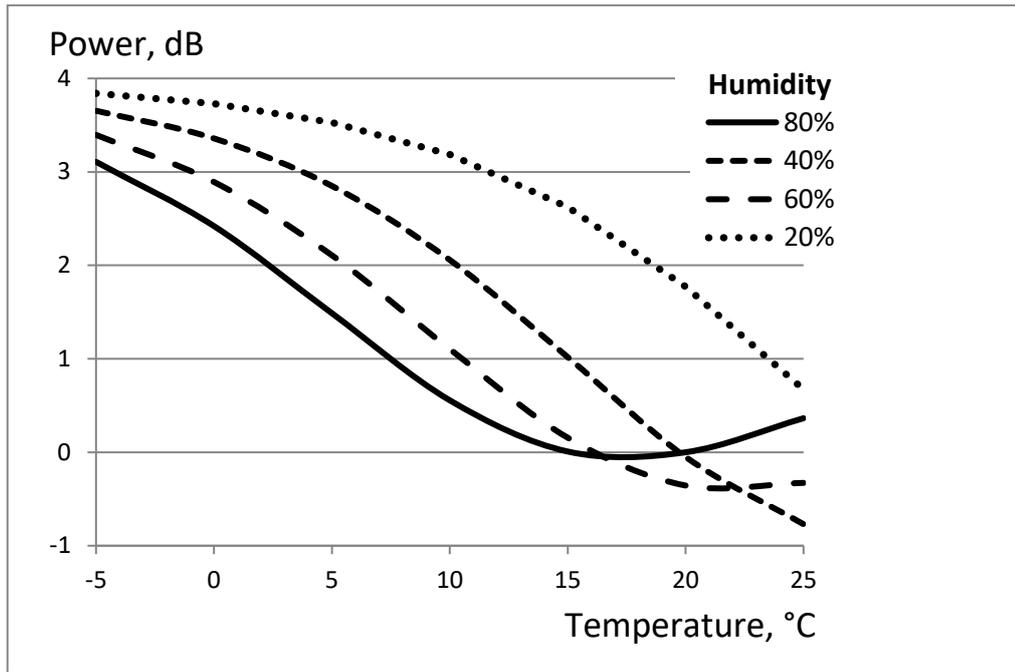
Figure 4.21: Predicted atmospheric absorption in db/100 meter at temperature of 20 C.

4.7 Conclusion

The Experimental set-up and the hardware used were described in this chapter. The data was acquired and visualized through signal processing. The data collection method used to accumulate the database shows that the proposed system is simple and capable of building up a large database that includes several features.

Section 4.5 showed the frequencies used and considered all the factors influencing the performance of the sensors or the accuracy of the measurements. Studies showed that the influence of these factors (vehicle movement, exhaust gases, etc.) can be reduced to be neglected. The effect of several weather conditions was discussed, and in most cases weather is not expected to attenuate the signal significantly. Weather conditions depend on many

factors like temperature, humidity, and size of water droplets. It is influence is lower than that of the vehicle's movement, and mostly it is transparent to RF radiations. The diversity of sensors used (RF and Ultrasonic) is essential to overcome such conditions.

Chapter 5

Classification Systems

5.1   Introduction

According to the methodology used for road surface identification and explained in chapter 2, the features extracted from both microwave and ultrasonic signals were used to define five classes and sort them out in a database according to the type and condition of the surface. Supervised classification algorithms were applied to take a decision of categorizing the newly obtained data into one of the classes (surfaces).

This chapter will explain the procedure of data training to accumulate a database. It will also describe the process known as "feature selection" used to select the most prominent features that can be integrated within a classification framework. Furthermore, it will investigate the classification algorithms applied to the different features and the evaluation methods used to compare the performance of the algorithms.

5.2 Database and Features

The variables used in the road surface identification set are considered continuous ratio. The class variable is a nominal variable with four or five categories (soft asphalt, rammed asphalt, grass, gravel, sand).

As mentioned in section *4.4.3*, features were extracted from each backscattered signal. These variables constitute the properties of the signal known as features. Each instance has the following raw features:

1. The average power of each polarization over 2.5 meters range.

2. The standard deviation for vertically polarized signal over 2.5 meters range.

3. The average power of the ultrasonic signal over 2.5 meters range normalized to the power of transmitted signal.

4. The standard deviation of the ultrasonic signal over 2.5 meters range normalized to the power of transmitted signal.

The standard deviation was calculated after applying the path loss compensation. It is related to the signal waveform and can be regarded as a characteristic of the surface roughness related to the wavelength.

In addition to the raw features, the following relative features were computed and added to each instance:

1. Ratio of vertical mean power over horizontal mean power P_{vv}/P_{hh}.

2. Ratios of cross polarized mean power over horizontal mean power P_{vh}/P_{hh}.

3. Ratio of vertical mean power at two frequencies. $P_{vv}(f_1)/P_{vv}(f_2)$

4. Ratio of horizontal mean power at two frequencies. $P_{hh}(f_1)/P_{hh}(f_2)$

5. Ratio of cross polarized mean power at two frequencies. $P_{vh}(f_1)/P_{vh}(f_2)$

6. Ratio of vertical mean power over ultrasonic power. P_{vv}/P_a

where v stands for vertical polarisation, h stands for horizontal , and P_a stands for ultrasonic power. In order to further understand the effect of roughness, two features that characterize roughness of the surfaces were added. The two parameters based on the ENDURA method theory used for ultrasonic signals helps separating between diffused reflections from rough surfaces and specular reflections from smooth surfaces. It is expected that adding those parameters (power above threshold & duration above threshold) would give better results, as for different roughness, we expect different duration and power. Figure 5.2 shows the signal processing behind the two parameters. Firstly, normalised round trip path loss compensation is applied to the signal. The green envelope in the figure shows the signal after compensation. The path loss given by:

$$\text{Path loss (dB)} = 40 * \log\left(\frac{4\pi d}{\lambda}\right) \tag{5.1}$$

where λ is the wavelength and d is the distance between the antennas and the surface. Second step is setting up a threshold based on the average amplitude of the signals. The threshold was taken the same for all the surfaces. Use of threshold reduces the dependence of the features on the transmitted power and other random factors. The optimization of the threshold is suggested for future investigation.
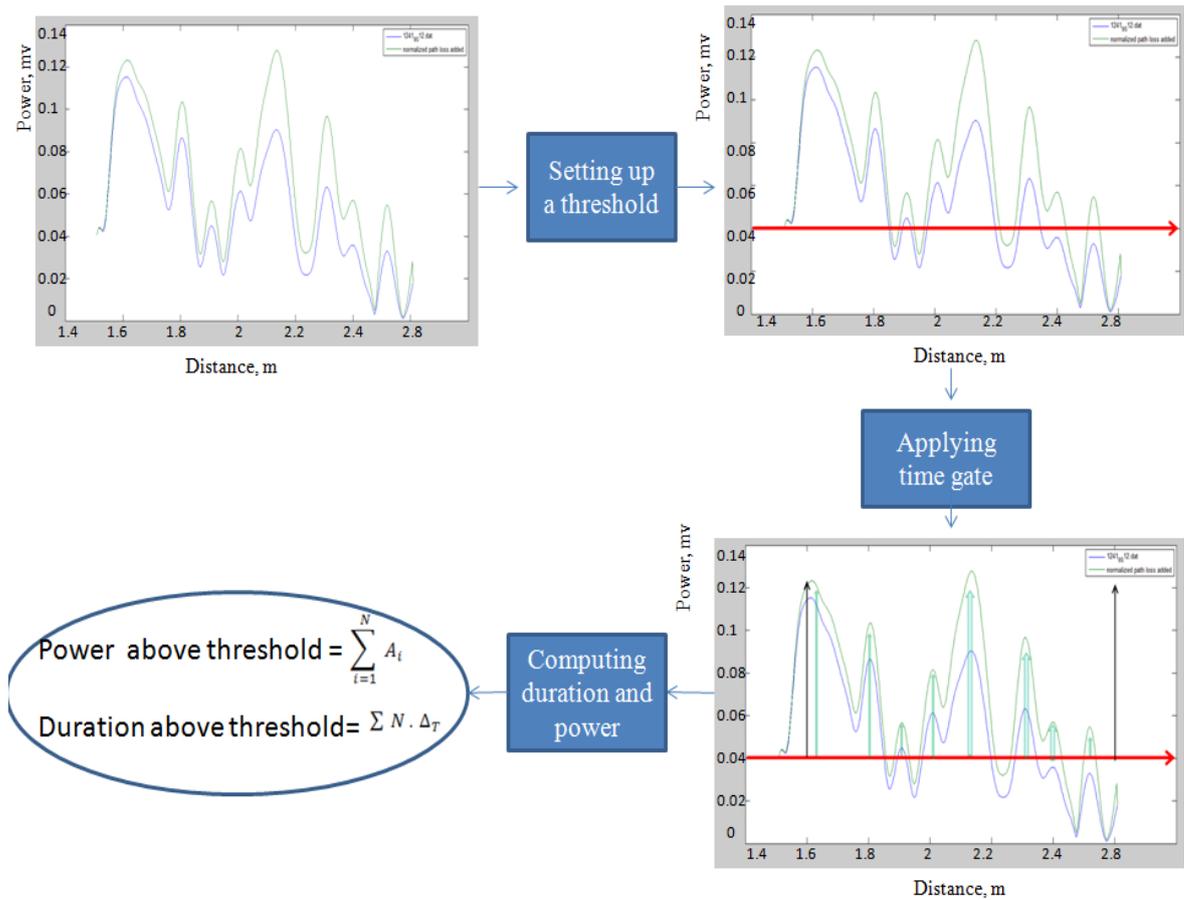
Figure 5.1 Power and duration above threshold.

Third step is applying the time gate (1.6 to 4 meters) then calculating the power and duration within the time gate. Power above threshold P is the amplitudes exceeding the threshold and it is given by:

$$P = \sum_{i=1}^{N} A_i \qquad (5.2)$$

where A_i is any amplitude within the time gate above the threshold. The duration above threshold is the time of backscattering signal with amplitudes exceeding threshold. This is given by:

$$D = \sum N . \quad \Delta_T \tag{5.3}$$

where N is the number of points exceeding the threshold, and Δ_T is the time step. The time step is the change in time between the steps of the time range. The total number of points is 1601 and the time step is computed as 0.4 seconds. Figure 5.3 shows the difference in duration and power above threshold for vertically polarized 9 GHz backscattering from three surfaces gravel, asphalt, and grass. The data in figure 5.3 is for research purposes only since the bandwidth used is 3 GHz (range resolution 0.05 meters) which is not allowed for unlicensed use. Extraction of similar features using a threshold is not possible for the radar signals at other frequencies and bandwidths due to the relatively low range resolution of microwave signals used (0.75 or 1 meter).

At low range resolution (1 meter), only 2 resolution cells can be found in the 2.5 meters studied portion of the signal. This is not enough to extract the power and duration above threshold.
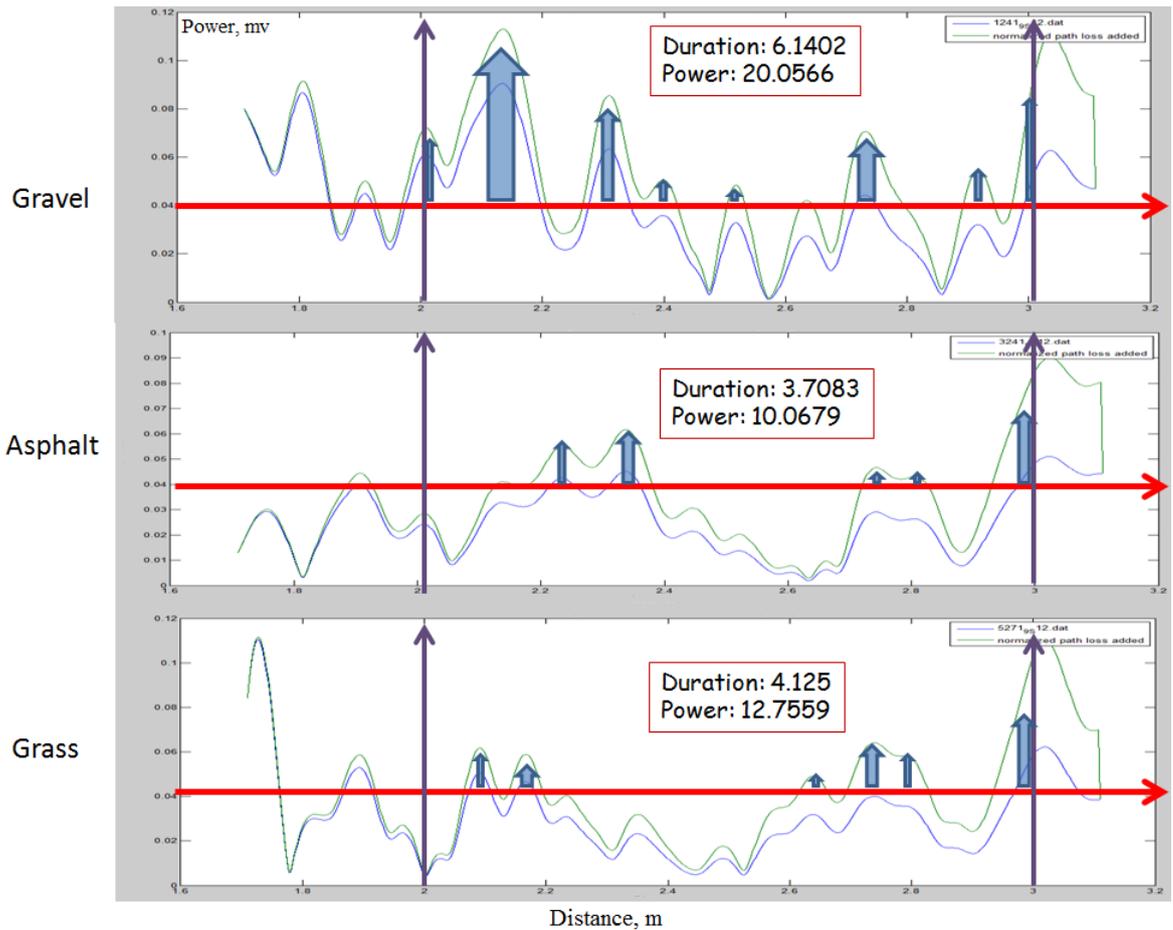
Figure 5.2 Extraction of Power and duration above threshold.

The initial set of features acquired out of backscattering from radar and ultrasonic signals are shown in table 5.1. Eleven features out of the 20 displayed features are applied for 5 frequency-bandwidth combinations which give 55 features. Adding it to the 14 possible combined features of radar and ultrasonic, gives 69 features in total. In addition to these features, the instances of the database also includes the file number, date of the experiment, time, temperature, humidity, weather (dry, wet), surface (gravel, asphalt, grass, sand), cover (snow, ice, frost), details (soft, rammed), bandwidth, frequency, grazing, angle, antenna's height. These detailed records help to identify any external factors affecting the measured row variables. The database consisted of more than 1100 instances conducted on different surfaces under the same conditions.

109

	Feature	Description
1	P_{vv}	S_{21} parameter vertically Tx, vertically Rx over 2.5m range.
2	P_{vh}	S_{21} parameter vertically Tx, horizontally Rx over 2.5m range.
3	P_{hv}	S_{21} parameter horizontally Tx, vertically Rx over 2.5m range.
4	P_{hh}	S_{21} parameter horizontally Tx, horizontally Rx over 2.5m range.
5	P_{vv}/P_{hh}	Normalized S_{21}
6	P_{vh}/P_{hh}	Normalized S_{21}
7	P_{hv}/P_{hh}	Normalized S_{21}
8	Std_{vv}	Standard deviation of S_{21} vertically Tx, vertically Rx
9	Std_{hh}	Standard deviation of S_{21} horizontally Tx, horizontally Rx
10	Dt_{vv}	Duration above threshold for S_{21}, vertically Tx, vertically Rx
11	Pt_{vv}	power above threshold for S_{21} ,vertically Tx, vertically Rx
12	$P_{vv(f_1)}/P_{vv(f_2)}$	Relative vertical power at two frequencies
13	$P_{hh(f_1)}/P_{hh(f_2)}$	Relative horizontal power at two frequencies
14	$P_{vh(f_1)}/P_{vh(f_2)}$	Relative cross polarized power at two frequencies
15	$P_{hv(f_1)}/P_{hv(f_2)}$	Relative cross polarized power at two frequencies
16	P_a	Average power of ultrasonic signal over 1 meter range (1.5 to 2.5m).
17	Std_a	Average power of ultrasonic signal over 1 meter range (1.5 to 2.5m).
18	Dt_a	Duration above threshold for ultrasonic signal
19	Pt_a	Power above threshold for ultrasonic signal
20	P_{vv}/P_a	Relative vertical radar power over ultrasonic power

Table 5.1 whole set of features.

The variables are reduced as some of them can be ignored like the duration and power above threshold for low range resolution signals (Dt_{vv}, Pt_{vv}). Although dimensionality reduction is somehow achieved by computing the new features out of the combinations of the row variables, the large amount of features makes the task of optimising the road surface classification algorithms a very complex one. Performance of the classification algorithms highly depend on the quality of features used Therefore, the number of features had to be reduced and the features giving the best classification should be chosen.

5.3 Feature selection

In classification, it is very challenging to address the redundant features and reduce its dimensionality. Feature selection reduces both the data and computational complexity saving time and providing more efficiency. At the initial stage of feature selection, redundant features had been eliminated. Normalized features are preferred since they depend less on the transmitter power and grazing angle. However, the power extracted from the reflected voltage (S21 values) we are using are normalized to the transmitted power. Moreover, experiments showed that average power of vertical-horizontal polarization is almost equal to that of horizontally transmitting – vertically receiving which agrees with the reciprocity theory. Therefore, only P_{vh} was used for classification.

To provide further understanding, a study was conducted on the correlation between the features. Table 5.2 shows the correlation table between a set of features. As shown in the table, green cells represent low correlation. Although some features showed high correlation (80 % correlation between vertical polarized average power and horizontal polarized average power), it wasn't enough at this stage, to eliminate features due to the lack of confidence and

certainty. Consequently, each feature was applied on its own to each classification algorithm comparing two surfaces at a time. Table 5.3 shows the classification accuracy for the different features and different combinations of surfaces at different frequencies and bandwidths. This provides an understanding of the most influencing features for each classifier and on each surface pair. For instance, the first row/first column cell shows the classification between asphalt and ice when applying S_{21} parameter vertically transmitted, vertically received over 2.5m range with frequency of 9 GHz and bandwidth of 3 GHz. This method shows the effect of each feature on the classification of surfaces. The table showed that some features like Duration above the threshold gives better classification for certain surfaces. Nevertheless, this method was not clear enough in terms of which optimal set of features to use for all combination of surfaces and classification algorithms.

There was a need to use filter and wrappers for feature's reduction. Filters were applied through the "Orange Canvas" software. Orange canvas is an open source, for data visualization and analysis, which takes the database as an input and filters the features according to ReliefF, Gini Gain, Gain Ratio, or Information Gain [110]. This ranks a score for each single feature used to separate the given data set. Below is a summary of the filters used:

	P_{vv}	P_{hh}	P_{vh}	P_{vv}/P_{hh}	P_{vh}/P_{hh}	Dt_{vv}	Pt_{vv}	Std_{vv}	Std_{hh}
P_{vv}	1	0.8	0.6			0.15	0.2		
P_{hh}	0.8	1							
P_{vh}	0.6		1						
P_{vv}/P_{hh}				1	0.2				
P_{vh}/P_{hh}				0.2	1				
Dt_{vv}	0.15					1	0.8		
Pt_{vv}	0.2					0.8	1		
Std_{vv}								1	0.75
Std_{hh}								0.75	1

Table 5.2: features correlation.

Feat. No	Freq-Bandwidth	Feature	Asphalt Ice	Asphalt Snow	Ice Snow	Asphalt Gravel	Asphalt Grass	Asphalt Sand	Gravel Grass	Gravel Sand	Grass Sand	Sand Smooth Sand Rough
1	Radar only	Pvv_9.0-3.00	0.75	0.64	0.81	0.87	0.76	0.79	0.54	0.61	0.55	0.84
2	9.0-3.00	Phh_9.0-3.00	0.75	0.83	0.81	0.84	0.91	0.82	0.66	0.66	0.66	0.98
3		Pvh_9.0-3.00	0.76	0.73	0.81	0.87	0.84	0.81	0.67	0.59	0.51	0.9
4		Pvv_9.0-3.00/Phh_9.0-3.00	0.76	0.84	0.81	0.7	0.9	0.84	0.79	0.78	0.59	0.9
5		Pvh_9.0-3.00/Phh_9.0-3.00	0.75	0.81	0.88	0.64	0.78	0.71	0.88	0.82	0.59	0.76
6		DTvv_9.0-3.00	0.99	0.92	0.92	0.88	0.92	0.83	0.88	0.71	0.81	0.94
7		PTvv_9.0-3.00	0.9	0.89	0.9	0.91	0.78	0.85	0.88	0.76	0.83	0.9
8		STDvv_9.0-3.00	1	1	1	0.87	0.85	0.95	0.85	1	0.83	0.63
9		STDhh_9.0-3.00										0.76
10	Radar only	Pvv_18.-3.00	0.75	0.61	0.84	0.76	0.65	0.67	0.61	0.66	0.64	0.85
11	18.-3.00	Phh_18.-3.00	0.75	0.83	0.85	0.82	0.89	0.81	0.68	0.57	0.58	0.96
12		Pvh_18.-3.00	0.75	0.73	0.83	0.82	0.79	0.75	0.52	0.62	0.58	0.85
13		Pvv_18.-3.00/Phh_18.-3.00	0.86	0.92	0.86	0.77	0.9	0.86	0.68	0.67	0.6	0.79
14		Pvh_18.-3.00/Phh_18.-3.00	0.81	0.76	0.82	0.69	0.77	0.74	0.64	0.73	0.6	0.77
15		DTvv_18.-3.00	0.99	0.94	0.98	0.88	0.89	0.83	0.87	0.73	0.83	0.83
16		PTvv_18.-3.00	0.8	0.94	0.83	0.93	0.95	0.8	0.94	0.6	0.81	0.92
17		STDvv_18.-3.00	1	1	1	0.83	0.88	1	0.87	0.92	0.85	0.62
18		STDhh_18.-3.00										0.63
19	Acoustic only	Pa	1	0.93	1	0.77	0.91	0.84	0.9	0.72	0.9	0.71
20		STDa	1	1	1	0.92	0.9	0.95	0.91	0.95	0.84	0.59
21	Acoustic and	Pa/Pvv_9.0-3.00	1	0.81	1	0.79	0.95	0.81	0.83	0.64	0.79	0.76
22	radar	Pa/Pvv_18.-3.00	1	0.85	1	0.8	0.94	0.78	0.89	0.64	0.9	0.71
23	Combined	Pvv_18.-3.00/Pvv_9.0.-3.00	0.75	0.57	0.81	0.57	0.55	0.68	0.57	0.64	0.66	0.69
24	radar	Phh_18.-3.00/Phh_9.0-3.00	0.75	0.59	0.81	0.55	0.51	0.57	0.63	0.7	0.58	0.63
25	9.0-3.00	Pvh_18.-3.00/Pvh_9.0.-3.00	0.82	0.56	0.81	0.67	0.55	0.62	0.53	0.65	0.59	0.75
26	18.-3.00	DTvv_18.-3.00/DTvv_9.0-3.00	0.75	0.91	0.97	0.86	0.82	0.78	0.84	0.68	0.7	0.71
27		PTvv_18.-3.00/PTvv_9.0-3.00	0.89	0.92	0.98	0.85	0.89	0.77	0.82	0.73	0.7	0.71
28		STDvv_18.-3.00/STDvv_9.0-3.	1	1	1	0.85	0.83	0.8	0.87	0.8	0.85	0.54

Table 5.3 Classification accuracy for different combinations of features and surfaces.

1. Information Gain: Information gain (IG) measures the reduction in the entropy from a random state to another state achieved by learning a new random variable A. Let S be a set of training examples where each example is of the form $(x, y) = [x_1, x_2, \dots, x_n, y_1]$. The overall entropy of S is defined as:

$$E (S) = - \sum_{i=1}^{c} p_i \log_2 p_i \qquad (5.4)$$

Where c is the total number of classes, and p_i is the portion of instances that belong to class i. For each feature A, the reduction of entropy is computed by:

$$IG (S, A) = E(S) - \sum_{v \in A} \frac{|S_{A,v}|}{|S|} E(S_{A,v}) \qquad (5.5)$$

where 'v' is a value of A, and $S_{A,v}$ is the set of instances in which A has a value 'v' [108]. IG will determine an ordering of features from most useful to least useful. The feature which maximizes the difference between the overall entropy $(E(S))$ and the average entropy $(\sum_{v \in A} \frac{|S_{A,v}|}{|S|} E(S_{A,v})$) will have higher IG. A drawback against Information gain is the possibility of adjusting the learner towards very specific random features which has a large number of distinct values but no relation to the target of classification. This phenomenon known as "Overfitting" usually occurs when the model is complex and has too many parameters.

2. _Gain Ratio: Gain ratio is the IG divided by the intrinsic information of each attribute. It is used instead of IG in some cases as it biases the algorithm against considering the attributes with a large number of distinct values. Nevertheless, using the Gain Ratio deals

unfairly with the attributes of very low information. The intrinsic information of an attribute A is given by :

$$\text{IntI} \ (S, A) \ = \ -\sum_{v \in A} \frac{|S_{A,v}|}{|S|} \ \log \left(\frac{|S_{A,v}|}{|S|} \right) \tag{5.6}$$

3. Gini Ratio: Gini ratio is the probability that two randomly chosen instances will have different classes. It measures the inequality among values and gives zero when there is a perfect equality. The equation of Gini is given by:

$$\text{Gini} \ (S) \ = 1 - \sum_i f_i^{\ 2} \tag{5.7}$$

where f_i is the impurity measure which is defined by the fraction of instances labelled with class i. The average Gini index (used instead of average entropy) is given by :

$$\text{Gini} \ (S, A) \ = \ \sum_{v \in A} \frac{|S_{A,v}|}{|S|} \ . \ Gini \ (S_{A,v}) \tag{5.8}$$

4. ReliefF: ReliefF is an updated form of the feature selection algorithm "Relief". The algorithm which can be used for both binary and continuous data is noise tolerant. Mainly, the algorithm searches for instances randomly then computes their nearest neighbours. It then updates the weight of the feature according to its contribution to the discrimination of the instances from its neighbors of different classes [109]. At each iteration, the feature vector of one random instance (x) is compared to the feature vector of the closest instances from each class (by Euclidean distance). The weight vector estimated by ReliefF for each feature is updated by:

$$V_i = \ V_i - (x_i \ - \ nearHit_i)^2 + (x_i \ - \ nearMiss_i)^2 \tag{5.9}$$

where nearHit is the closest same-class instance, and nearMiss is the closest different-class instance. Therefore, the weight of the feature will increase if it differs from the feature of nearby different-class instances more than the nearby same-class instances. This method is slow since it needs to find K neighbors for each of the instances and for each feature at a time.

Figure 5.4 shows an example of "Orange Canvas" ranking of features. The optimal set of 10 features with highest ranks were taken and applied to a wrapper for all the available classification algorithms. Wrappers were used to search the features space by the forward selection method. Starting with no features, and adding one feature at a time then measuring the reduction in classification error until any further addition doesn't decrease the error. This has been applied on the different algorithms and on features for each radar frequency. Table 5.4 shows an example of the wrapper forward selection method. For radar and ultrasonic features and a classification algorithm KNN with 2 nearest neighbours the classification accuracy was computed for each feature set starting with no features. Using wrappers, the optimal set of features was set for radar features, ultrasonic features, and combination of radar and ultrasonic. The same forward selection method was used in order to figure out the optimal number of features. Table 5.4 shows that adding more than 6 features lead to decreasing the classification accuracy. Figure 5.5 shows the probability of correct recognition with respect to number of features for the "K nearest neighbours" KNN and "MDC" algorithms.

116

Figure 5.3 Ranking of Features using Orange Canvas filters.

Feature Set	Classifier	Classification Accuracy
$\{.\}$	KNN	
$\{P_{vv}, P_a\}$	KNN	83 %
$\{P_{vv}, P_a, Dt_a\}$	KNN	86%
$\{P_{vv}, P_{hh}, P_a, Dt_a\}$	KNN	89%
$\{P_{vv}, P_{hh}, P_a, Dt_a, Std_a\}$	KNN	91%
$\{P_{vv}, P_{vh}, P_{hh}, P_a, Dt_a, Std_a\}$	KNN	89%
$\{P_{vv}, P_{vh}, P_{hh}, Std_{vv}, P_a, Dt_a, Std_a\}$	KNN	87%
$\{P_{vv}, P_{vh}, P_{hh}, Std_{vv}, P_a, Pt_a, Dt_a, Std_a\}$	KNN	85%

Table 5.4 Wrapper predictions.

The probability of correct recognition was compared for asphalt, gravel, sand, and grass. This was applied to all the other surfaces with covers and to each one of the 4 classification algorithms (K nearest neighbour KNN, Multilayer perceptron MLP, Maximum likelihood MLE, and Minimum Distance Classification MDC with 2 different metrics). Figure 5.6 compares the average probability of correct recognition for the 4 classification algorithms. As shown in figure 5.6, increasing the number of features to more than 6 features doesn't improve the probability of correct recognition.
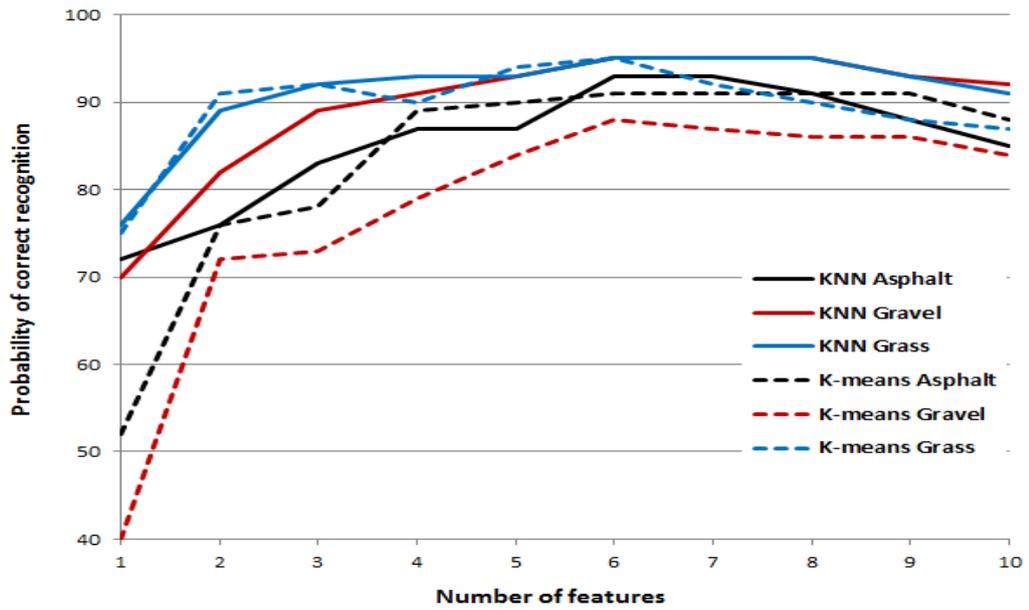
Figure 5.4 Number of features for KNN and MDC (based on kmeans).
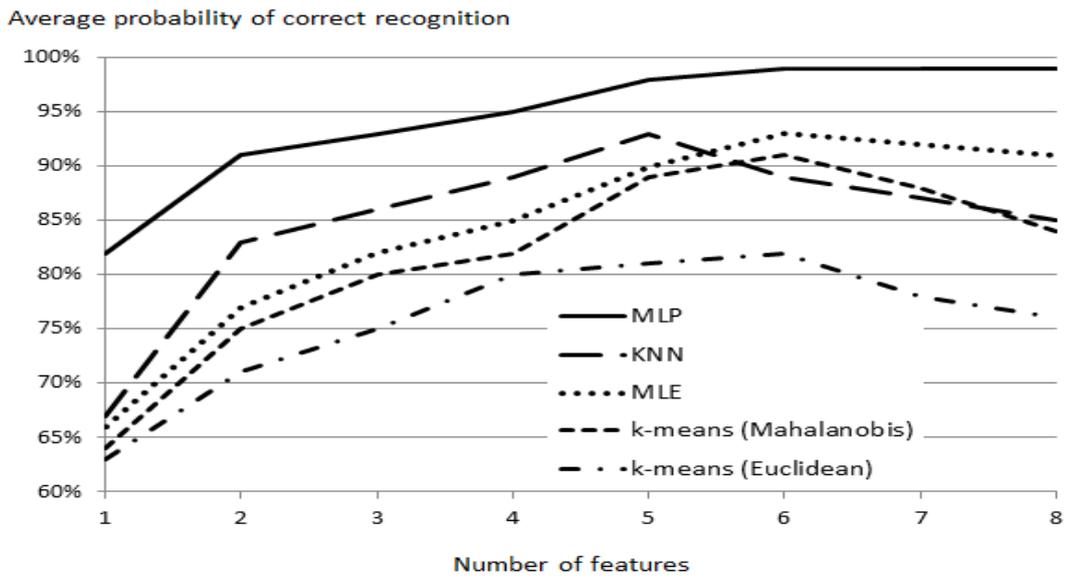
Figure 5.5 Average probability of correct recognition vs. number of features.

According to most algorithms, the highest probability was achieved by 5 or 6 features. The optimal 6 features for radar, ultrasonic, and combination of radar and ultrasonic features deduced for road surface identification is shown in Table 5.5. These features worked the best

for all surfaces and most of the classification algorithms. The data used for training and testing is explained in section 5.7.

	Radar Features	Ultrasonic Features	Radar & Ultrasonic
1	P_{vv}	Dt_a	P_a
2	P_{vh}	Pt_a	P_{vh}
3	P_{hh}	P_a	P_{vv}
4	Std_{vv}	Std_a	Dt_a
5			Std_a
6			P_{hh}

Table 5.5 Optimal features.

5.4 Orange Canvas

The data analytics features of Orange Canvas were used in this research to extract features and evaluate the classification algorithms. The scheme shown in figure 5.7 shows the data mining workflows of road surface identification. The components in the scheme are called widgets. Widgets exchange information amongst each other starting with the "File" widget which access the database through a text file and then communicate through channels with the other widgets. Data table widget present the data set as a spreadsheet format, and Scatterplot provides a 2 dimensional visualization of the features.

The data is displayed as a scatterplot of points with positions depending on the X-axis and Y-axis features set by the user. The most important widget is the "Test learners" as it takes the input data set after feature selection then applies it to the classification algorithms and output the results to the evaluation methods. The classification algorithms used in Orange canvas and Matlab will be discussed in details in the next section. In this section, the emphasis will be on the evaluation methods used to compare algorithms and detect errors. Test learners widget tests the algorithm on data. It uses different methods sampling schemes to train the data. (cross validation, random sampling, test on train data, and leave-one-out).
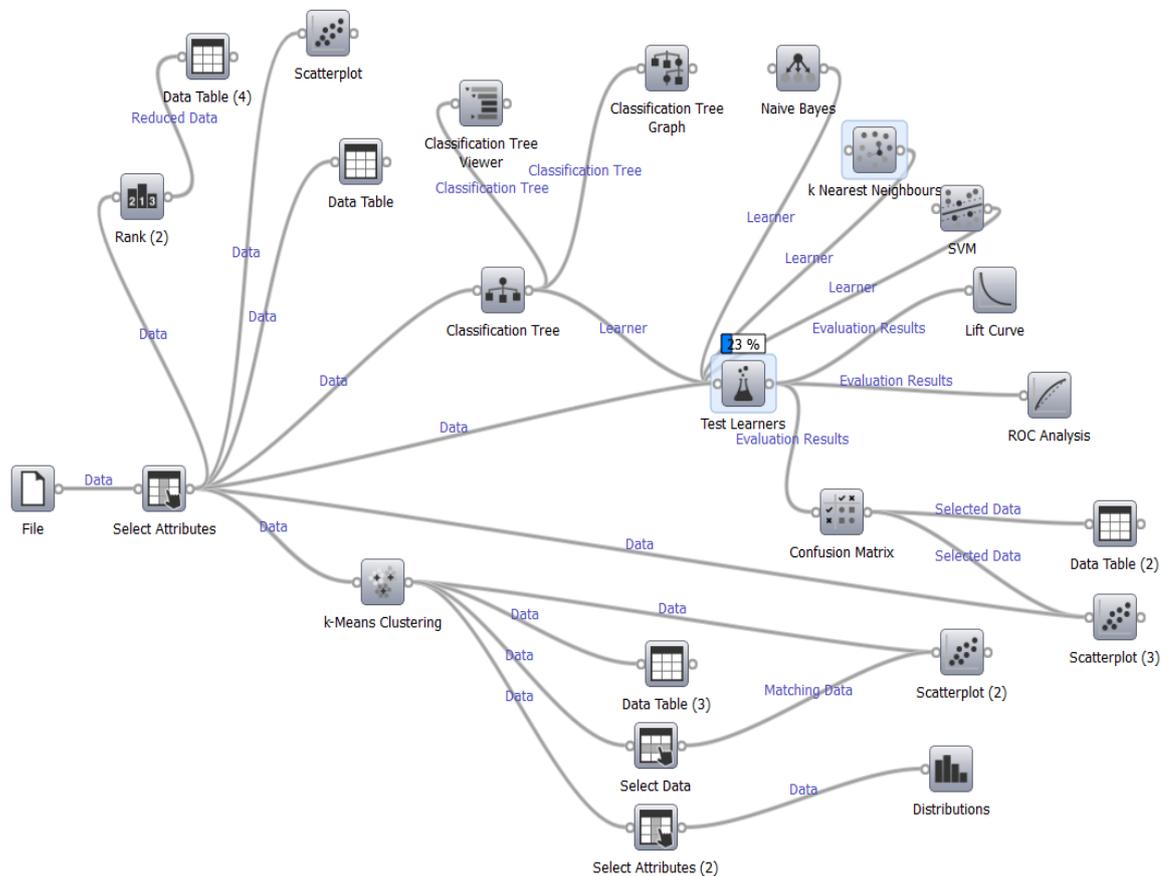
Figure 5.6 Scheme of road surface identification.

Cross validation splits the data into given number of folds set by the user. One fold is to be classified while the others are used as training data to induce the model. Test on train data

uses the whole set of data for training, and random sampling randomly splits the data into training and testing according to proportion set by the user (for example 80:20). The validation technique adopted for this investigation is "Leave-one-out". This is similar to Cross Validation but it holds out one fold at a time, induce the model from the other folds, then classify the held out fold. Later, it chooses another fold to be held out and repeat the process. This method is stable and reliable but slow in case of very large datasets. Additionally, test learners widget shows different performance measures of the classifiers (such as classification accuracy), and outputs data to the other widgets that evaluate the performance of the classifiers (like ROC Analysis and Confusion Matrix).

5.5 Measures of Performance

In order to distinguish mistakes, one of the classes is usually selected as the positive class or the target class. To ease understanding, the example of patients will be used where the positive class means having a disease. This yields 4 cases of correct and wrong predictions:

1. True positive (TP): represents the instances which were correctly classified to the positive class.

2. False positive (FP): represents the instances which were falsely classified to the positive class. For example, predicting the patient has the disease although he actually doesn't.

3. True negative (TN): represents the instances correctly classified to one of the negative classes.

4. False negative (FN): represents the instances wrongly classified as negative. For example, model predicted the patient is healthy although he has the disease.

These predictions were used to define the measures of performances. Indeed, each prediction can be more or less important according to the data set of the investigation. In medicine, the concern is to have FN as low as possible since its error is fatal. There are multiple measures of performance that can be used for comparing classifiers. Each has its advantages and drawbacks. For this study, the following measures were used:

1. Classification accuracy (CA): It measures the proportion of the correctly predicted instances the model gives. A classification accuracy of 90% means that 90 out of 100 cases were correctly classified. The drawback of this measure is that it considers all mistakes equal.

$$\text{CA} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5.10)$$

2. Sensitivity: also called True Positive Rate (TPR), the sensitivity of the model is the number of correctly classified positives over all the actual positives. It is given by:

$$\text{S} = \frac{TP}{TP+FN} \qquad (5.11)$$

3. Specificity: The specificity of the model is the number of correctly classified negatives over the entire actual negative.

$$\text{Sp} = \frac{TN}{TN+FP} \qquad (5.12)$$

4. Precision: also called Positive Predictive Value (PPV) is the number of True positives divided by the total number of instances predicted as positive. It is the probability that the instance predicted positive is indeed positive.

$$\text{P} = \frac{TP}{TP+FP} \qquad (5.13)$$

Many alternative measures can be used like Brier score, Recall, F-measure, and Area under ROC. All of these methods were initially used to measure performance, but the methods were reduced to four and on a later stage only classification accuracy (CA) was adopted. Reader is referred to [61][110] for detailed description of the measures. Table 5.7 shows an example of applying Test learner to KNN algorithms while changing the number of neighbours when gravel was considered as the target class.

The high sensitivity (above 0.9) shows that KNN classifies 90 % of gravel instances correctly. Nevertheless, when the target class was changed to the other surfaces it also gave high sensitivity. The challenge was to achieve a high CA since it represent the overall performance of the classifier regarding all the surfaces under investigation.

Number of Neighbours/ Measures	CA	Sensitivity	Specificity	Precision
3	0.7391	0.9048	0.84	0.8261
4	0.7717	0.9286	0.86	0.8478
5	0.7826	0.9048	0.88	0.8636
6	0.7826	0.9286	0.86	0.8478

Table 5.6 Measure of performance for KNN with respect to number of neighbours.

Two methods were used to test and compare the classifiers and visualize their performance: ROC Analysis and Confusion Matrix. The two methods provided by Orange Canvas widgets were developed in Matlab in order to cover two different sources of evaluation.

1.   ROC Analysis

A receiver operating characteristics (ROC) graph is a graphical plot that has been used in signal detection theory and then extended for analysing the behaviour of diagnostic systems[111]. The graphical plot shows the relationship between sensitivity and specificity. It is created by plotting the true positive rate (sensitivity) vs. the false positive rate (1-specificity).

At the initial stage of the research, ROC curve was used to compare the available supervised learning models in Orange Canvas. Figure 5.8 shows an example of ROC curve for 4 classifiers: Classifier 1 (dark blue), classifier 2 (cyan blue), classifier 3 (green line), and classifier 4 (red line) where the target class is grass. The diagonal line represents the behaviour of a random classifier. The more the plot is directed to the left, the more accurate the classifier is. Below the diagonal, the classifier is considered of low performance. Figure 5.8 shows that, for grass surface, classifier 1 and classifier 2 achieve high sensitivity (between 0.9 and 0.95) with a reasonable specificity (between 0.8 and 0.85). In contrary, classifier 3 displays lower sensitivity while classifier 4 falls below the diagonal. This method is helpful to compare the classifiers as it shows the relationship between two measures of performance. Hence, it was developed in Matlab as a method of evaluation beside the confusion matrix.
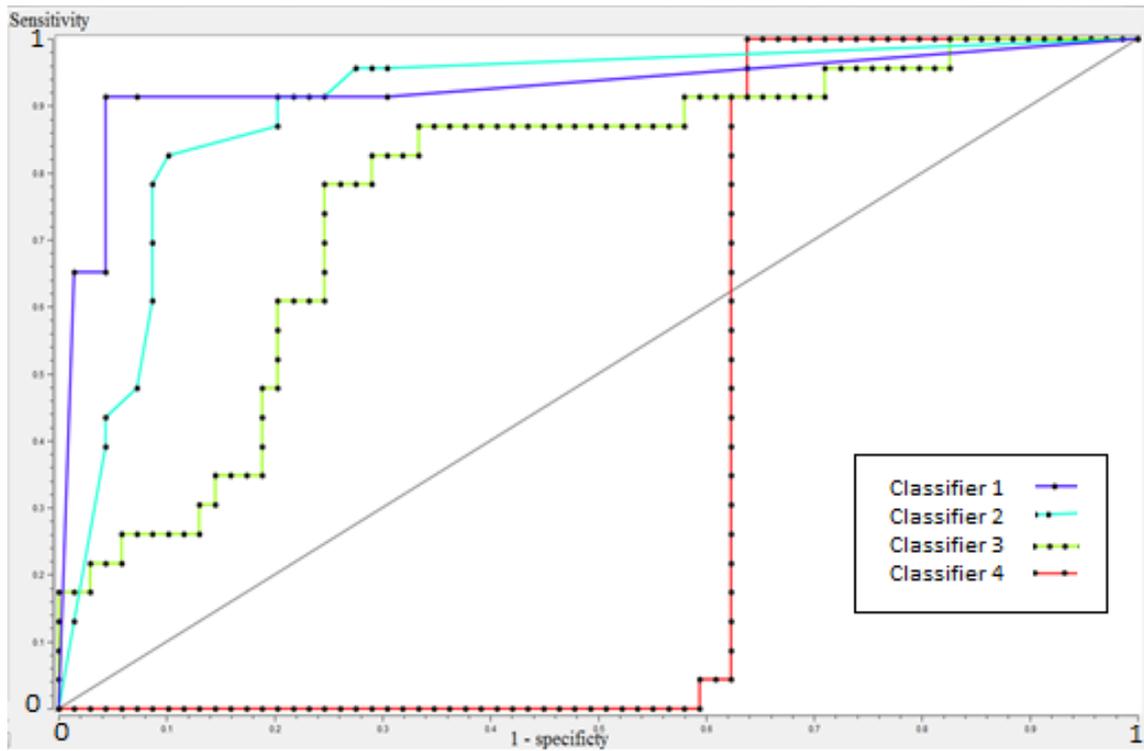
Figure 5.7 ROC curve.

2.   Confusion Matrix

Also known as contingency table or error matrix, confusion matrix is a matrix layout that allows visualization of the performance of an algorithm. It makes it clear to detect if the system is confusing any of the classes with the other. Confusion matrix determines the number or proportion of instances classified into another or same class. Each column in the matrix represents the predicted classes, while each row represents the correct class. Table 5.7 shows an example of a confusion matrix for three surfaces (asphalt, gravel, grass). The rightmost column gives the number of examples from each class (e.g. 35 is the total number of examples which belongs to Asphalt), and the next to the last row gives the number of examples classified into each class (e.g. 37 instances were classified into Asphalt).

	Asphalt	Gravel	Grass	Total
Asphalt	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	35
Gravel	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	25
Grass	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$	41
Total	37	25	39	101
Error	E1	E2	E3	Etotal

Table 5.7 Confusion Matrix

The last row was added using Matlab to show the error of each class, and consequently the total error. The confusion matrix for the classification algorithms is given as follows:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \ \text{x} \ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} \tag{5.14}$$

Where $a_{i,i}$ = number of examples correctly predicted as surface i, and $a_{i,j}$ $(i \neq j)$ = number of examples of surface i wrongly predicted as surface j. T_i is equal to the number of examples belonging to surface i. In table 5.7, where $i=1$ is Asphalt, $i=2$ is gravel, $i=3$ is grass). In order to check the accuracy, the error's row added to the confusion matrix where,

$$E_i = \frac{a_{i,j} + a_{i,k}}{T_i} \tag{5.15}$$

e.g. E_1 is the number of Asphalt's instances wrongly predicted as Gravel or Grass, divided over the total number of Asphalt's instances. E_{total} is the sum of the errors for the three surfaces and it can be used to compare the performance of the classification algorithms.

This matrix provided by the confusion matrix widget acquires its input from Test learner. Confusion matrix was also developed in all the Matlab codes as the main method to visualize the performance of the classifiers. This is due to the fact that the classes used are balanced and the number of instances for each class is almost equal to that of the other classes. The

126

average probability of correct recognition is calculated by averaging the diagonal of the matrix:

$$APCR = \frac{\sum a_{i,i}}{number\ of\ surfaces}$$ (5.16)

5.6 Classification Algorithms of Road Surface Identification

After deducting the optimal features, the whole dataset was used to obtain a model, for each surface, describing the distribution of the optimal features in the feature space. Figure 5.9 shows the distribution of a part of the dataset with respect to three optimal features. The data and features used here for demonstration purposes shows that some classes like snow and grass are harder to distinguish, while other classes have very good distinction.
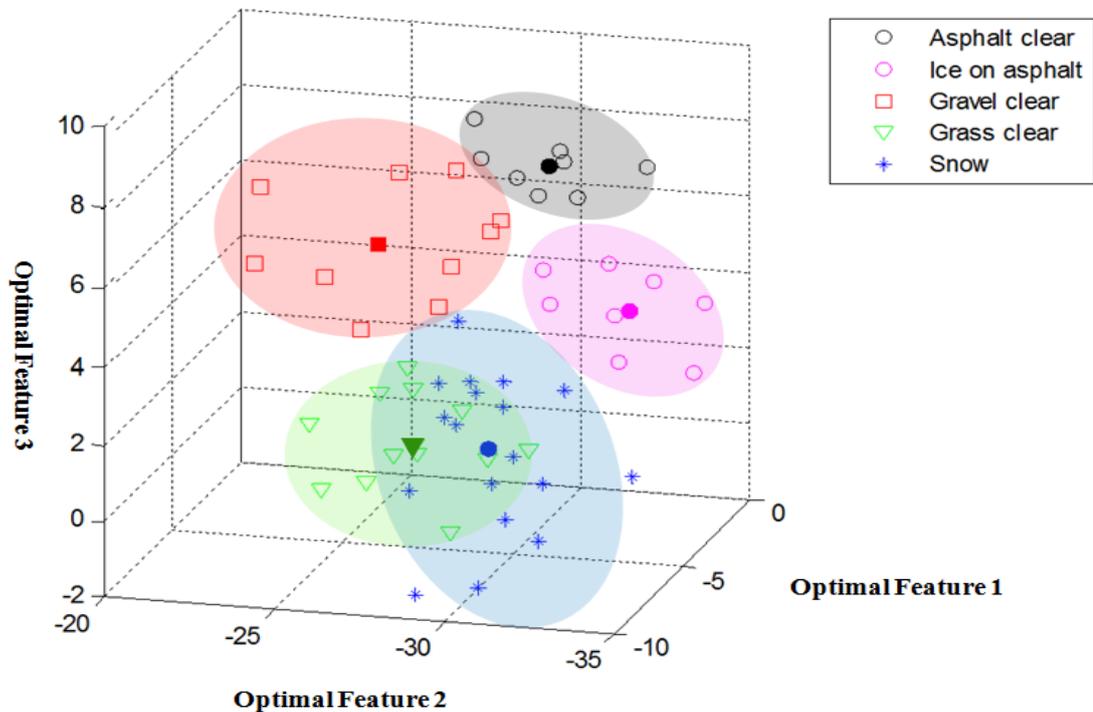
Figure 5.8 Distribution of classes on the feature space.

127

The 4 classification algorithms used were applied to the data set and compared through two methods: graphical user interface scheme in "Orange Canvas" and Matlab codes. For each algorithms, there was few parameters affecting its accuracy and defining its performance.

Figure 5.10 shows the block scheme of the Matlab codes developed for the classifiers KNN, MDC, and MLE. The Matlab codes which can be found in Appendix F have many versions all of which can be summarized in the block scheme. At the start of the program, the initial settings are set depending on the algorithm used. The program then extract the training database from an excel file. This database is the full database with the whole set of variables including weather conditions, surface cover, and all the above mentioned raw variables. It then applies feature reduction by choosing 3 to 6 features. In some versions, this step prompts the user to enter the optimal features he wishes to apply, while in other the optimal features were extracted automatically. Following, database is divided into smaller databases representing the training data for each class. The division is based on surfaces, different covers, frequencies, and roughness in some cases (for example rough asphalt and soft asphalt). Many test points are probed against the training database, and based on the decision of the algorithm the confusion matrix is automatically built. In addition, other versions of the code were developed to finalize by drawing the ROC curve and scatterplot of the classes showing the shortest distance between the test point and the nearest class. The next sections will explain how the 4 classifiers were used with the parameters affecting its accuracy and defining its performance.
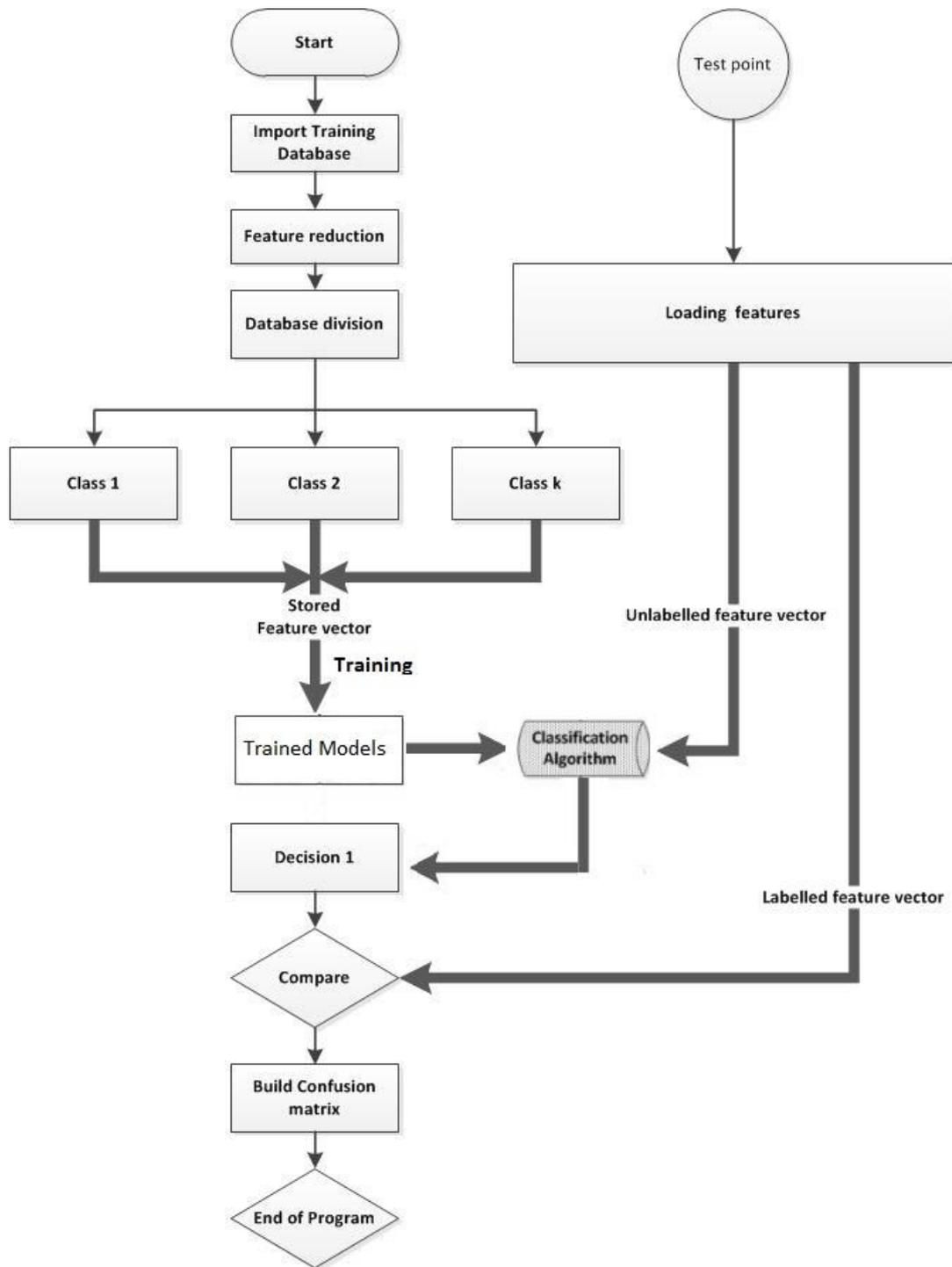
Figure 5.9 Block scheme of the Matlab codes.

5.6.1 K Nearest Neighbours (KNN)

There are two important factors the affects the performance of KNN:

1. The number of nearest neighbours K.

2. The measure of distance (metrics) between the points.

Our approach is to get the optimal number of neighbours and the best metrics so that the highest classification accuracy can be achieved through KNN. Figure 5.11 shows a scatterplot of KNN algorithm applied on our surface data with k=3 neighbours. D_1 ,D_2, and D_3 are the sum of the distances between the test point and the 3 neighbours.
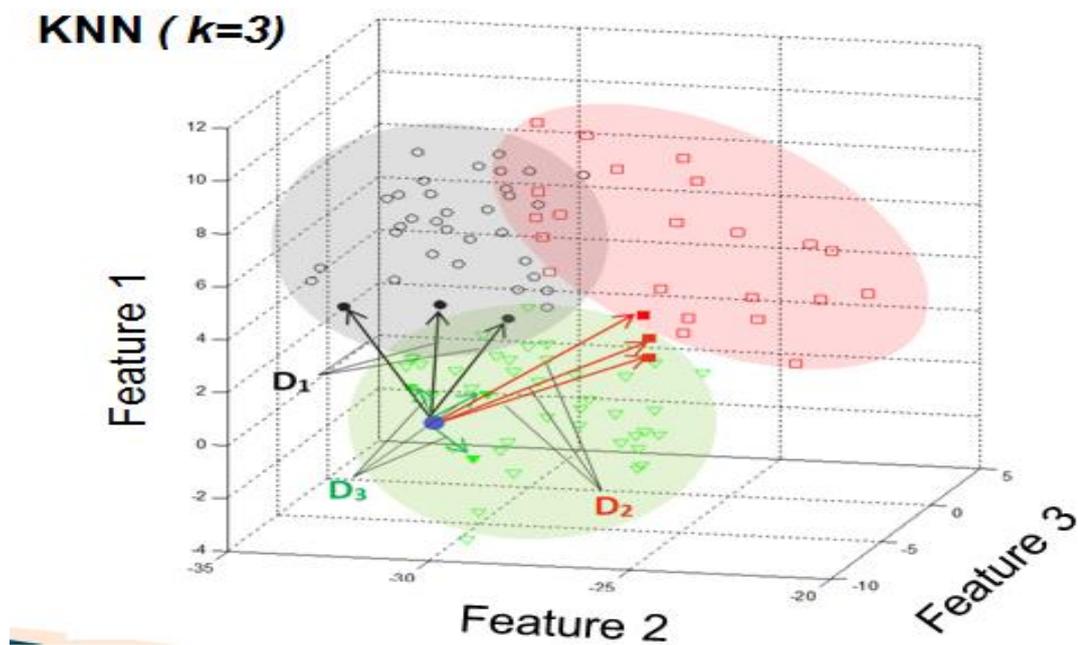
Figure 5.10 Distance between test point and training data using KNN algorithm.

1. K neighbours

In order to determine the optimal number of neighbours, the algorithm developed on Matlab was run for the same dataset while changing the number of neighbours and measuring the probability of correct recognition. In addition to this method, Orange canvas was used to change the number of neighbours in the K nearest neighbour widget and measure the classification accuracy at each time. Two methods were used in order to compare the performance and provide reliable results. Figure 5.12 shows the KNN widget which was used to change the number of neighbours and metrics. Figure 5.13 shows the probability of correct recognition with respect to number of neighbours for asphalt, gravel, and grass. As shown in figure 5.13, the probability of recognition decreases significantly when the number of neighbours exceeds 3 for grass surface. For the other surfaces, it declines after 3 neighbours and then increases when using more than 4 neighbours.

Figure 5.11: K Nearest Neighbours widget
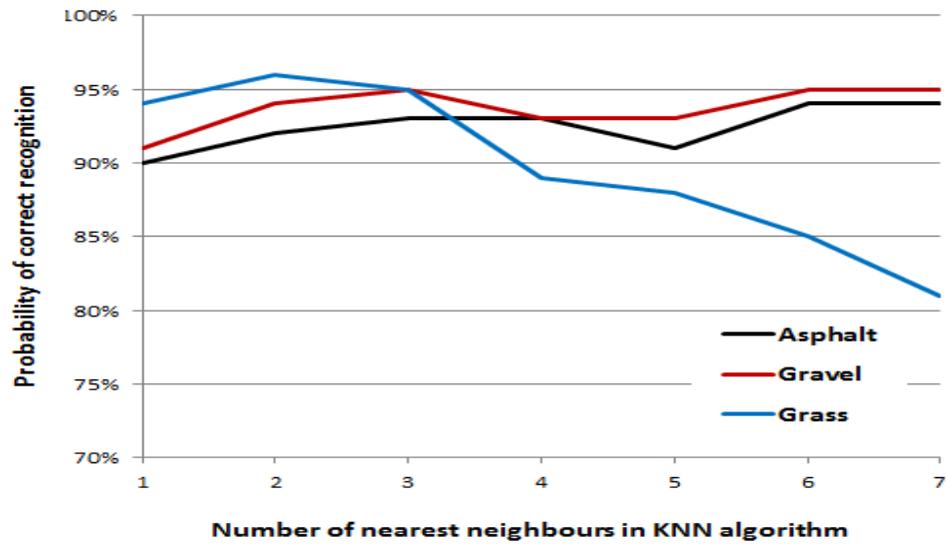
Figure 5.12: Probability of correct recognition with respect to number of neighbours.

This investigation showed that k=2 and k=3 results are very similar. This agrees with the simple empirical approach which states that the number of neighbours K should be equal to [112]:

$$K = \sqrt{p} \qquad (5.17)$$

Where $p = 6$ is the number of features (dimensionality of the data). Nevertheless, the number of Neighbours used for KNN was 2 neighbours since it compromises for the best performance amongst all surfaces.

2. The distance function (Metrics)

The distance function determines the distance between each pair of elements of a set. There is a need to measure the distance between the features vector of the tested point and the corresponding K closest neighbours in each class. Defining a distance function between the

132

pairs of features vectors depends on the features and data types of the variables. Many distance functions exist and can be used in classification (Euclidean distance, Manhattan, Hamming, and Chebyshev) but the most widely used for most learners is the Euclidean distance [60]. A further description of each distance function is found in Appendix C.

During the initial stage of the research, a case study approach was used to allow understanding of each metric. The study based upon changing the number of neighbours and metrics for KNN in Orange Canvas and measuring the performance for each target class. Table 5.8 shows the performance measures of Euclidean, Manhattan, and Maximal distance for KNN when gravel was considered the target class. Three features (P_a, P_{vv}, and Dt_a) were used in these results shown in table 5.8, but all the other optimal features were used and the same conclusion has been reached. As shown in table 5.8 (b), the metrics achieved very promising measures with sensitivity hitting 0.8. In order to determine the optimal metric, the algorithms developed on Matlab were run for the same dataset while changing the metrics and measuring the probability of correct recognition. The analyses showed that the Euclidean distance is a good compromise considering the large and small differences between the feature values. Nevertheless, the classification accuracy achieved by the other metrics was not significantly less than that achieved by the Euclidean distance. Euclidean distance was used for KNN and MLE, but a new distance function was proposed for the MDC algorithm. This function will be discussed in the next section.

Number of Neighbours/ Measures	Classification Accuracy	Sensitivity	Specificity	Precision
3	0.7717	0.8095	0.8	0.7727
4	0.7717	0.8095	0.8	0.7727
5	0.7717	0.8095	0.8	0.7727

| 6 | 0.7609 | 0.8095 | 0.78 | 0.7556 |

(a) Euclidean

Number of Neighbours/ Measures	Classification Accuracy	Sensitivity	Specificity	Precision
3	0.7500	0.8095	0.7800	0.7556
4	0.7500	0.8095	0.7800	0.7556
5	0.7500	0.8095	0.7800	0.7556
6	0.7609	0.8095	0.8000	0.7727

(b) Manhattan

Number of Neighbours/ Measures	Classification Accuracy	Sensitivity	Specificity	Precision
3	0.7717	0.8095	0.8000	0.7727
4	0.7717	0.8095	0.8000	0.7727
5	0.7609	0.7857	0.8000	0.7674
6	0.7391	0.7857	0.7600	0.7333

(c) Maximal

Table 5.8 Performance measures for different metrics (a) Euclidean, (b) Manhattan, (c) Maximal.

5.6.2 Minimum Distance Classification (MDC)

The Minimum Distance Classifier (MDC) classify unknown tested point to a class according to the minimum distance between the point and the class in multi-feature space. Figure 5.15 shows an example of MDC on our surface data with three different classes distributed according to three different features. D_1 ,D_2, and D_3 are the distances between the test point and the 3 means.

134
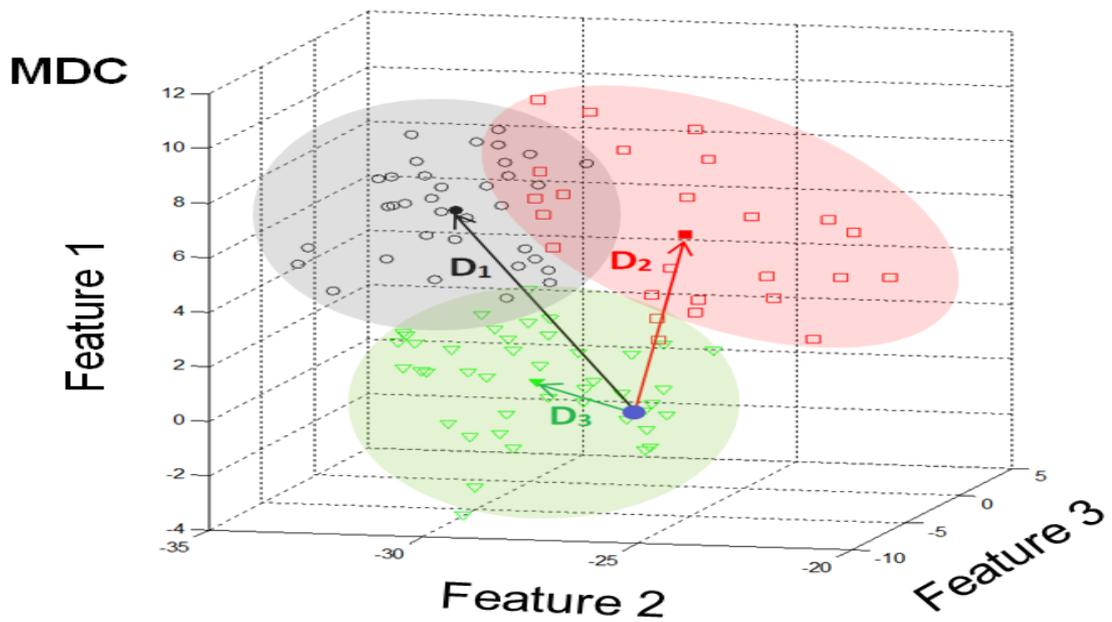
Figure 5.13 Distance between test point and training data using MDC algorithm.

The measure of distance (metrics) between the tested point and the means is the factor affecting the performance of MDC. Our approach is to obtain the best metrics so that the highest classification accuracy can be achieved. In order to determine the optimal metric, the algorithms developed on Matlab were run for the same dataset while changing the metrics and measuring the probability of correct recognition. In addition to the Euclidean, Manhattan, and Maximal, one more metric was used in the MDC classification. This metric is an adaptation of the Mahalanobis distance introduced by P.C. Mahalanobis in 1936. A further description of each distance function is found in Appendix C .

The new adapted distance metric was applied by estimating the mean for each class and then calculating the distance between the instances and the means averaged over the standard deviation of the class. The minimum distance means the least number of standard deviations

the test point is away from the mean of the class. The classification accuracy achieved by the adaptation of the Mahalanobis distance was significantly higher than that achieved by the Euclidean distance for MDC. The results will be demonstrated in the next chapter.

5.6.3 Classification Via Gaussian Distribution Model

The Matlab code developed for road surface identification via Gaussian Distribution Model using MLE method finds the values of parameters which makes the known likelihood distribution a maximum. It estimates the Maximum likelihood estimators (MLEs) for each class in the model. After estimating the MLEs, it calculates the probability density function pdf of the test point (Explained in Chapter 2). This is estimated for each dimension (feature) in the class. Based on the highest probability, the decision on belonging of the tested point to the closest class is taken. Like all the previous codes, MLE method for Gaussian Distribution Model returns a confusion matrix with classification accuracy for all the surfaces compared.

5.6.4 Multilayer Perceptron (MLP)

This Multilayer Perceptron is used instead of a Perceptron since it can distinguish non-linear data. A single layer Perceptron can only discriminate linearly separated data; a limitation that is overcome by the use of multiple layers in MLP [116]. The Matlab nprtool GUI was used to apply MLP on the road surface identification data. The tool uses two layer feedforward network (figure 5.16) of multiple nodes with sigmoid transfer functions in both hidden and output layer. The nodes in each layer, except the input layer, are processing units

(neurons). The number of output neurons is set to 5 which is the number of classes (surfaces) to classify, and the network will be trained with scaled conjugate gradient Backpropagation technique.
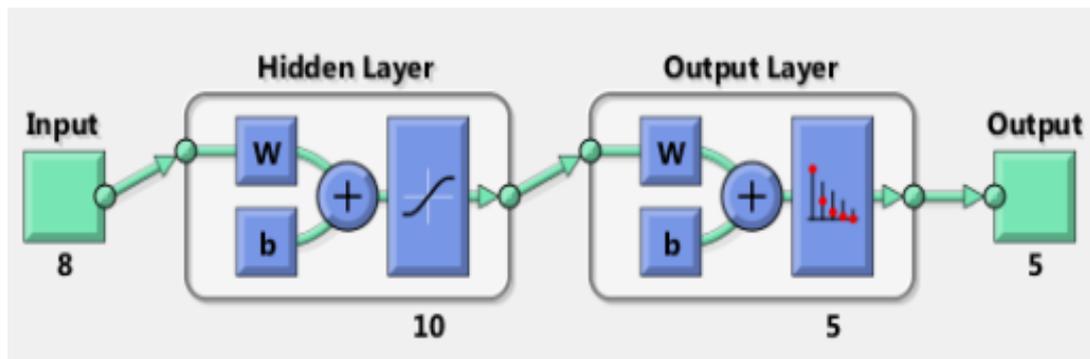
Figure 5.14Neural Network of road surface identification

1. Experimental Methodology

The training phase of the MLP algorithm consists of providing two matrices that define the pattern recognition problem. The first matrix named "Qinput" contains the full set of data with feature values, while the second one "Qtarget" consists of the labels of the instances of "Qinput". Table 5.9 shows a real example of the input database of the road surface identification MLP system. Table 5.9 (a) shows the first four rows of Qinput with 8 feature's values extracted from both ultrasonic and radar sensors. The first four rows of Qtarget which indicates the classes is shown in table 5.9 (b). Since the 5 surfaces (soft asphalt, rammed asphalt, grass, gravel, and sand) are being compared, each scalar target value is set to either 1 or 0, indicating which class the corresponding input belongs to. Both matrices has 612 rows (number of instances used for MLP) of data collected under dry weather.

P_a	Dt_a	Pt_a	Std_a	P_{vv}	P_{vh}	P_{hh}	Std_{vv}
-28.5030	1.889	0.2803	0.147	0.1047	0.1019	0.1171	0.8656
-28.2630	1.851	0.303	0.145	0.1672	0.1097	0.1515	0.6789
-28.0190	1.843	0.2948	0.142	0.1124	0.0992	0.0832	0.9481
-30.6480	1.379	0.225	0.153	0.0833	0.0326	0.0451	0.7592
…	…	…	…	…	…	…	…

Soft Asphalt	Rammed Asphalt	Grass	Gravel	Sand
0	0	0	1	0
0	0	0	1	0
0	0	0	1	0
1	0	0	0	0
…	…	…	…	…

(b)Qtarget

Table 5.9 Matrices for MLP input (a) Qinput , (b) Qtarget

After defining the data sets and importing the matrices into the input layer, the network is trained through the following steps:

1. Setting the training, validation, and testing data sets (figure 5.17). Qinput and Qtarget will be randomly divided into these 3 data sets. It is common in supervised classification to use 70 % of the original data for training. Validation data set which comprises 15% of the original data is used to monitor the error. After developing the model, 15 % of the data are used to test the model. This will give the true expected

performance of the model (Held-out-technique) [61]. For more details about dividing the data, readers are referred to [117].

2. Setting up the number of hidden neurons. This number was initially set to 10, and then optimized to 25. Further optimization of the number of neurons was not investigated, as it wasn't in our research interest and it will be proposed for future work. The problem of Overfitting can be overcome by reducing the number of neurons. As mentioned before, the number of output neurons is automatically set to the number of columns in Qtarget (5 classes).
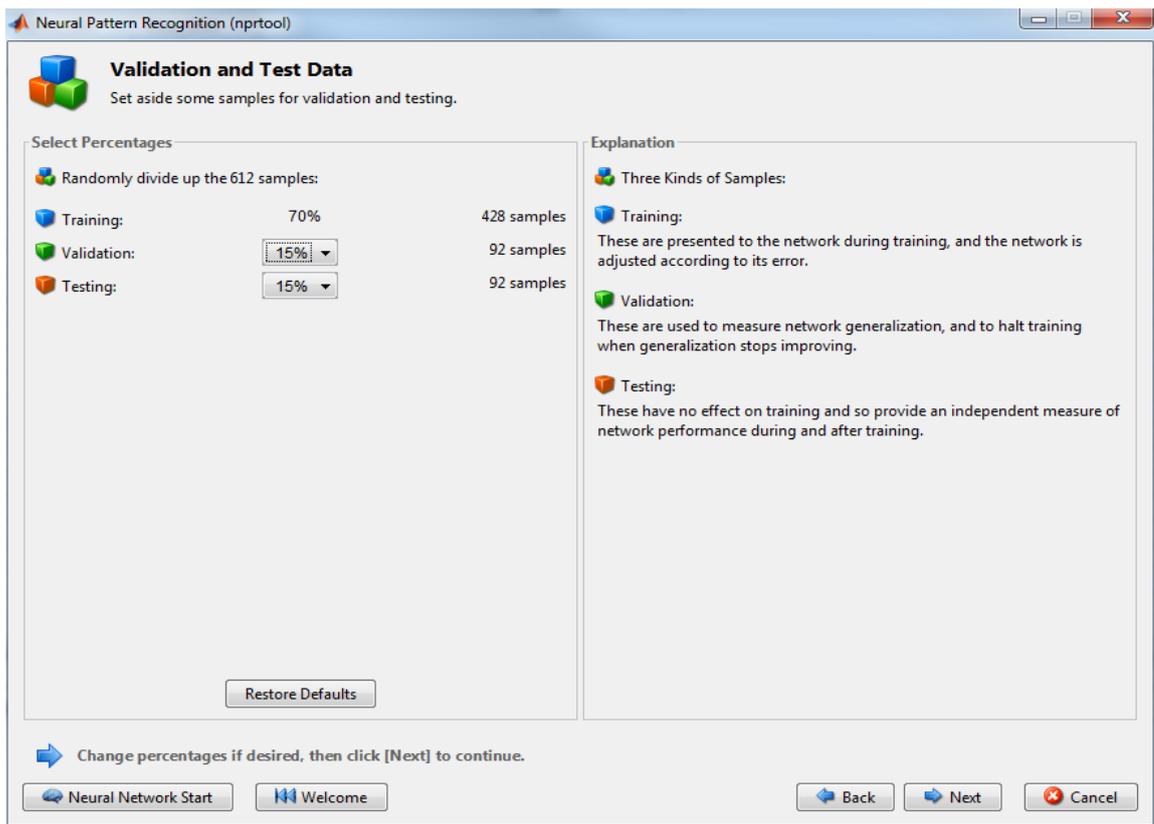
Figure 5.15: Dividing the data sets.

3. Training the network and updating the weight and bias values using the conjugate scaled gradient Backpropagation. This algorithm, based on Conjugate Gradient Methods, uses the second order information from the neural network and requires less memory and less consuming time as compared to normal Backpropagation. Conjugate Scaled Gradient Backpropagation is explained in details in [118]. Training automatically stops when there is no further improving; this is indicated by the error of the validation set. It took 34 iterations of training road surface identification data set.

4. Plotting the confusion matrix, ROC curve, and evaluating the network with a new set of data. Figure 5.16 shows confusion matrices for training, validation, and testing of the 5 surfaces. The green squares represent the correct prediction, while red boxes represent the incorrectly predicted percentage. The blue square illustrates the overall accuracy of the network. In this example the final matrix "Test Confusion Matrix" shows an average probability of 84.8% for 5 surfaces. The results achieved with retraining the algorithm will be shown and discussed in the next chapter.
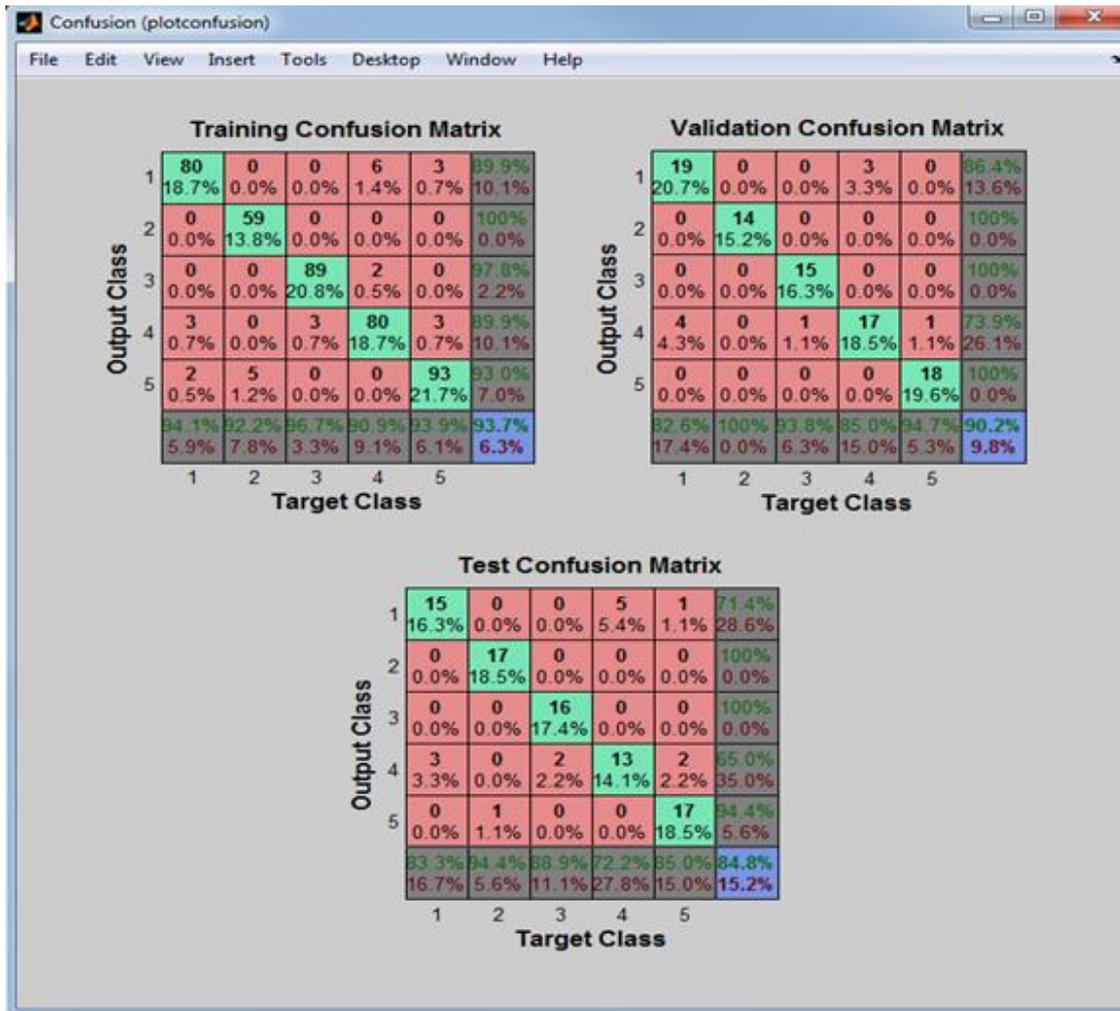
Figure 5.16 Confusion matrices of training, validation, and test data.

5.7 Conclusion

Using Orange Canvas and Matlab, features were extracted from real data experimentations and reduced by the use of filters and forward selection method wrapper. A set of optimal features was deduced for radar data, ultrasonic data, and both radar and ultrasonic. Four classification algorithms were applied to the data set, and two methods were used to evaluate the classifiers (confusion matrix and ROC curve). It was shown in this chapter that

performance of KNN classifier is affected by the distance functions and number of neighbors, while that of MDC and MLE depends on the distance function and the correlation between features. The number of neurons and the training of the network is expected to affect the performance of MLP classifier. Results of the performance and evaluation of the classifiers will be shown in the next chapter.

Chapter 6

Evaluation and Results

6.1   Introduction

The developed radar and ultrasonic system was positioned along more than 450 locations to ascertain repeatability of backscattering properties from surfaces. The database composed of the signatures was used to classify new unlabelled signatures and predict any new surface encountered using four different classification algorithms (KNN, MDC, MLE, and MLP). This chapter illustrates the results obtained by the classification algorithms and compare its performance and classification accuracy. In addition, the effect of exhaust gases, when it interferes with the echo signal, on the accuracy of measurements is shown.

6.2 Experimental Results

During the classification process, the performance of several data sets was investigated separately. Radar at different frequencies (5.8 GHz, 18 GHz, and 24 GHz), ultrasonic, and various combination of radar and ultrasonic data sets were analyzed, trained, and applied to the classification algorithms as explained in chapter 5. The summary of the results will be shown according to each data set, and the overall performance of the classifiers will be evaluated.

6.2.1 Ultrasonic Sensor Data

Based on the analysis of ultrasonic data, and using the optimal ultrasonic features by itself, the classification algorithms showed high classification accuracy. Figure 6.2 shows the confusion matrix of classification accuracy for KNN (K=2) using ultrasonic optimal features.

Features	Dt_a , Std_a				
	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	65%	0%	20%	9%	6%
Asphalt Smooth	0%	88%	0%	0%	12%
Grass	15%	0%	83%	2%	0%
Gravel	9%	0%	3%	88%	0%
Sand	9%	10%	0%	4%	77%
Average probability of correct recognition					81%

(a) 2 features

Features	Dt_a , Std_a , P_a				
	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	74%	0%	8%	16%	2%
Asphalt Smooth	0%	91%	0%	0%	9%
Grass	7%	0%	90%	3%	0%
Gravel	18%	0%	3%	78%	1%
Sand	2%	8%	0%	1%	89%
Average probability of correct recognition					84%

(b) 3 features

Figure 6.1 Confusion matrices of KNN using ultrasonic features (a) 2 features, (b) 3 features.

Using the 3 ultrasonic features of mean signal power P_a, signal duration above the threshold Dt_a, and signal standard deviation Std_a (figure 6.2 (b))

KNN	Dt_a, Std_a, P_a, Pt_a				
		Asphalt.			
	Asphalt	Smooth	Grass	Gravel	Sand
Asphalt	73%	0%	8%	16%	3%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	12%	0%	86%	2%	0%
Gravel	22%	0%	3%	73%	2%
Sand	9%	0%	0%	5%	86%
Average probability of correct recognition					84%

(a)

MLP	Dt_a, Std_a, P_a, Pt_a				
		Asphalt.			
	Asphalt	Smooth.	Grass	Gravel	Sand
Asphalt	84%	0%	10%	3%	3%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	12%	0%	85%	3%	0%
Gravel	12%	0%	1%	87%	0%
Sand	0%	0%	0%	0%	100%
Average probability of correct recognition					92%

(c)

MDC	Dt_a, Std_a, P_a, Pt_a				
		Asphalt.			
	Asphalt	Smooth	Grass	Gravel	Sand
Asphalt	70%	0%	5%	6%	19%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	16%	0%	84%	0%	0%
Gravel	27%	0%	5%	63%	5%
Sand	14%	0%	0%	0%	86%
Average probability of correct recognition					81%

(b)

MLE	Dt_a, Std_a, P_a, Pt_a				
		Asphalt.			
	Asphalt	Smooth.	Grass	Gravel	Sand
Asphalt	65%	0%	8%	11%	16%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	10%	0%	90%	0%	0%
Gravel	3%	0%	2%	95%	0%
Sand	10%	0%	0%	0%	90%
Average probability of correct recognition					88%

(d)

Figure 6.2 Confusion matrices of ultrasonic optimal features using 4 classifiers (a) KNN, (b) MDC, (c) MLP, (d) MLE.

improved the average propability of correct recognition by 3 %. The APCR achieved by two optimal features (Dt_a, Std_a) is 81 %, while that of 3 features (Dt_a, Std_a, P_a) is 84 %. Adding a fourth feature (power above threshold Pt_a) provided the optimal results out of

ultrasonic data set for all the classification algorithms. Figure 6.3 shows the confusion matrices for the four classifiers with four features. Good differentiation is demonstrated between smooth asphalt, grass, sand and other types of examined surfaces with the probability of correct recognition between 81% and 92%. However, asphalt and gravel are very similar in the acoustic properties, so their probability correct recognition is significantly lower than the other surfaces.

The performance of MDC method is slightly lower than the performance of KNN method (Figure 6.3 confusion matrix (b)). MLE provides better probability of correct surface recognition, than both KNN and MDC, especially in the case of gravel (figure 6.3 confusion matrix (d)). The best performance was achieved by MLP method (Figure 6.3 confusion matrix (c)) with average probability of 92 % when using 30 nodes in the hidden layer.

6.2.2 Radar Sensor Data

Based on the analysis of radar data, and using the optimal radar features by itself, the classification algorithms showed lower classification accuracy than that of ultrasonic analysis. Figure 6.4 shows the confusion matrix for KNN (k=2) using radar optimal features. Using the 4 radar features of vertical mean signal power P_{vv}, cross polarized mean signal power P_{vh}, horizontal mean signal power P_{hh}, and vertical standard deviation Std_{vv} (figure 6.4 (b)) at 5.8 GHz improved the average probability of correct recognition by 30% as compared to using the 3 relative features (P_{vv}/P_{hh}, P_{vh}/P_{hh}, and Std_{vv}) shown in figure 6.4 (a). The APCR achieved by three features is 50 %, while that of 4 optimal features (P_{vv}, P_{vh}, P_{hh}, Std_{vv}) is 80 %.

The results obtained with 5.8 GHz, 18GHz and 24 GHz radar data sets did not differ; therefore only 18 GHz results will be illustrated. Using the 4 radar features P_{vv}, P_{vh}, P_{hh}, and Std_{vv} (figure 6.5 (b)) at 18 GHz improved the average propability of correct recognition of KNN by 34% as compared to using the 3 relative features shown in figure 6.5 (a).

Features	P_{vv}/P_{hh}, P_{vh}/P_{hh}, Std_{vv}				
	Asphalt	Asphalt Sm.	Grass	Gravel	Sand
Asphalt	51%	13%	20%	13%	3%
Asphalt Smooth	2%	60%	22%	12%	4%
Grass	6%	18%	39%	31%	6%
Gravel	6%	9%	27%	50%	8%
Sand	0%	19%	19%	9%	53%
Average probability of correct recognition					50%

(a) 3 relative features

Features	P_{vv}, P_{vh}, P_{hh}, Std_{vv}				
	Asphalt	Asphalt Sm.	Grass	Gravel	Sand
Asphalt	74%	0%	11%	15%	0%
Asphalt Smooth	4%	83%	0%	13%	0%
Grass	0%	3%	81%	14%	2%
Gravel	3%	8%	5%	81%	3%
Sand	0%	0%	14%	5%	81%
Average probability of correct recognition					80%

(b) 4 optimal features

Figure 6.3 Confusion matrices of KNN using radar features at 5.8 GHz frequency (a) 3 relative features , (b) 4 optimal features.

The APCR achieved by three features $(P_{vv}/P_{hh}, P_{vh}/P_{hh}, Std_{vv})$ is 46 %, while that of 4 optimal features $(P_{vv}, P_{vh}, P_{hh}, Std_{vv})$ is 80 %. This showed that the results of classification based on the analysis of 18/24 GHz radar data are almost the same as those based on 5.8 GHz radar.

Features	$P_{vv}/P_{hh}, P_{vh}/P_{hh}, Std_{vv}$				
	Asphalt	Asphalt Sm.	Grass	Gravel	Sand
Asphalt	36%	15%	16%	20%	2%
Asphalt Smooth	8%	48%	20%	17%	6%
Grass	14%	24%	45%	14%	3%
Gravel	11%	20%	15%	50%	3%
Sand	10%	14%	14%	14%	48%
Average probability of correct recognition					46%

(a) 3 features

Features	$P_{vv}, P_{vh}, P_{hh}, Std_{vv}$				
	Asphalt	Asphalt Sm.	Grass	Gravel	Sand
Asphalt	70%	8%	5%	10%	7%
Asphalt Smooth	4%	92%	0%	0%	4%
Grass	3%	0%	81%	14%	2%
Gravel	10%	2%	7%	77%	3%
Sand	10%	10%	0%	0%	80%
Average probability of correct recognition					80%

(b) 4 features

Figure 6.4 Confusion matrices of KNN using radar features at 18 GHz frequency (a) 3 features , (b) 4 features.

Using the four optimal features yielded the best results out of 18 GHz radar data set for all the classification algorithms. Figure 6.6 shows the confusion matrices for the four classifiers

148

with four features. Although APCR is lower than that of ultrasonic data, good differentiation is demonstrated between smooth asphalt, grass, sand and other types of examined surfaces with the probability of correct recognition between 69% and 88%.

KNN — $P_{vv}, P_{vh}, P_{hh}, Std_{vv}$

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	75%	6%	3%	8%	8%
Asphalt Smooth	2%	98%	0%	0%	0%
Grass	3%	0%	75%	20%	2%
Gravel	17%	0%	8%	72%	3%
Sand	9%	5%	0%	0%	86%
Average probability of correct recognition					81%

(a)

MLP — $P_{vv}, P_{vh}, P_{hh}, Std_{vv}$

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	75%	0%	2%	15%	8%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	2%	0%	96%	2%	0%
Gravel	2%	0%	9%	84%	5%
Sand	17%	0%	0%	0%	83%
Average probability of correct recognition					88%

(c)

MDC — $P_{vv}, P_{vh}, P_{hh}, Std_{vv}$

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	49%	2%	0%	16%	33%
Asphalt Smooth	0%	90%	0%	0%	10%
Grass	2%	0%	67%	28%	3%
Gravel	24%	0%	19%	44%	13%
Sand	5%	0%	0%	0%	95%
Average probability of correct recognition					69%

(b)

MLE — $P_{vv}, P_{vh}, P_{hh}, Std_{vv}$

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	74%	3%	0%	6%	17%
Asphalt Smooth	0%	98%	0%	0%	2%
Grass	0%	0%	73%	27%	0%
Gravel	19%	0%	17%	64%	0%
Sand	0%	0%	0%	0%	100%
Average probability of correct recognition					81%

(d)

Figure 6.5 Confusion matrices radar 18 GHz optimal features using 4 classifiers (a) KNN, (b) MDC, (c) MLP, (d) MLE.

The performance of MDC method is the lowest with APCR of 69 % (Figure 6.6 confusion matrix (b)). KNN and MLE provide better probability of correct surface recognition than MDC. The difference is significant in terms of the 5 surfaces (figure 6.6 confusion matrix

(a) & (d)).The best performance was achieved by MLP method (figure 6.6 confusion matrix (c)) with average probability of 88 % when using 30 nodes in the hidden layer.

Since each separate sensor (radar or ultrasonic) couldn't provide reliable detection of all five surfaces, we examined the data obtained from two sensors. Starting with radar data set with two frequencies, figure 6.7 shows the confusion matrices for the four classifiers using a combination of optimal 5.8 and 18 GHz radar features. The features used ($P_{vv5.8}$, $P_{hh5.8}$, P_{vv18} , P_{hh18} P_{vh18}) provided better results than that of the separate sensors with probability of correct recognition between 82% and 97%.

KNN — $P_{vv5.8}, P_{hh5.8}, P_{vv18}, P_{hh18}$ P_{vh18}

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	89%	2%	0%	6%	3%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	92%	8%	0%
Gravel	6%	2%	5%	87%	0%
Sand	0%	0%	0%	0%	100%

Average probability of correct recognition **93%**

(a)

MLP — $P_{vv5.8}, P_{hh5.8}, P_{vv18}, P_{hh18}$ $P_{vh18}, , Std_{vv5.8}, Std_{vv18}$

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	94%	0%	0%	5%	1%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	98%	2%	0%
Gravel	3%	0%	2%	95%	0%
Sand	0%	0%	0%	0%	100%

Average probability of correct recognition **97%**

(c)

MDC — $P_{vv5.8}, P_{hh5.8}, P_{vv18}, P_{hh18}$ P_{vh18}

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	78%	2%	0%	20%	0%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	74%	26%	0%
Gravel	8%	0%	22%	70%	0%
Sand	5%	5%	0%	0%	90%

Average probability of correct recognition **82%**

(b)

MLE — $P_{vv5.8}, P_{hh5.8}, P_{vv18}, P_{hh18}$ P_{vh18}

	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	87%	2%	0%	5%	6%
Asphalt Smooth	0%	100%	0%	0%	2%
Grass	0%	0%	68%	32%	0%
Gravel	19%	0%	14%	67%	0%
Sand	0%	0%	0%	0%	100%

Average probability of correct recognition **84%**

(d)

Figure 6.6 Confusion matrices radar 5.8 +18 GHz optimal features using 4 classifiers (a) KNN, (b) MDC, (c) MLP, (d) MLE.

The performance of MDC method and MLE is almost similar with APCR of 82 % and 84 % respectively. KNN provides better probability of correct surface recognition than both MDC and MLE methods. The difference is significant in terms of the 5 surfaces. The best performance was achieved by MLP method (figure 6.7 confusion matrix (c)) with an average probability of 97 % when using 30 nodes in the hidden layer.

6.2.3 Radar and Ultrasonic Sensor Data

The below presented results are based on the analysis of both ultrasonic and 24 GHz radar data sets. Using 6 optimal radar and ultrasonic features, the classification algorithms reached the highest classificaiton accuracy of the 5 surfaces. Figure 6.8 shows the confusion matrices for MDC using the 6 features. Using Euclidean distance (figure 6.8 (a)), APCR achieved by MDC was 81%, but it significantly improved to 91% when Mahalanobis distance was used. The optimal features provided are vertical mean signal power P_{vv}, cross polarized mean signal power P_{vh}, horizontal mean signal power P_{hh}, mean ultrasonic signal power P_a, signal duration above the threshold Dt_a , and ultrasonic signal standard deviation Std_a. The accuracy of asphalt, gravel, and grass particulary improved with Mahalanobis distance. Mahalanobis distance is ,therefore, considered the distance function to be used with MDC classifier in all cases.

Figure 6.9 shows the confusion matrices for all the four classifiers with 6 optimal features. Excellent differentiation is demonstrated between all types of examined surfaces with the probability of correct recognition between 91 % and 99 %. The performance of MDC and KNN method is the lowest with APCR of 91% (Figure 6.9 confusion matrix (a) & (b)). MLE

151

provides better probability with correct surface recognition of 94%. A significant

improvement was noticed in the case of distinguishing sand from other surfaces

Features	$P_{vv}, P_{vh}, P_{hh},, Dt_a, Std_a, P_a$				
	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	70%	0%	5%	6%	19%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	2%	0%	84%	14%	0%
Gravel	25%	0%	3%	69%	3%
Sand	18%	0%	0%	0%	82%
Average probability of correct recognition					81%

(a) Euclidean distance

Features	$P_{vv}, P_{vh}, P_{hh},, Dt_a, Std_a, P_a$				
	Asphalt	Asphalt Smooth	Grass	Gravel	Sand
Asphalt	84%	0%	0%	5%	11%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	98%	2%	0%
Gravel	8%	0%	5%	87%	0%
Sand	14%	0%	0%	0%	86%
Average probability of correct recognition					91%

(b) Mahalanobis distance.

Figure 6.7 Confusion matrices of MDC using radar and ultrasonic features (a) Euclidean distance , (b)

Mahalanobis distance.

when MLE was used. The best performance was achieved by MLP method (confusion

matrix (c)) with average probability of 99 % when using 30 nodes in the hidden layer.

Distinguishing smooth asphalt, grass, and sand is 100% accurate using this method. Only

4% of asphalt was misclassified as gravel, and 2 % of gravel was misclassified as asphalt.

KNN	$P_{vv}, P_{vh}, P_{hh}, , Dt_a , Std_a ,$ P_a				
	Asphalt	Asp. Sm.	Grass	Gravel	Sand
Asphalt	92%	0%	2%	3%	3%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	2%	0%	98%	0%	0%
Gravel	8%	0%	3%	89%	0%
Sand	24%	0%	0%	0%	76%
Average probability of correct recognition					91%

(a)

MLP	$P_{vv}, P_{vh}, P_{hh}, , Dt_a , Std_a ,$ P_a				
	Asphalt	Asp. Sm.	Grass	Gravel	Sand
Asphalt	96%	0%	0%	4%	0%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	100%	0%	0%
Gravel	2%	0%	0%	98%	0%
Sand	0%	0%	0%	0%	100%
Average probability of correct recognition					99%

(c)

MDC	$P_{vv}, P_{vh}, P_{hh}, , Dt_a , Std_a ,$ P_a				
	Asphalt	Asp. Sm.	Grass	Gravel	Sand
Asphalt	84%	0%	5%	5%	11%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	98%	2%	0%
Gravel	8%	0%	5%	87%	0%
Sand	14%	0%	0%	0%	86%
Average probability of correct recognition					91%

(b)

MLE	$P_{vv}, P_{vh}, P_{hh}, , Dt_a , Std_a ,$ P_a				
	Asphalt	Asp. Sm.	Grass	Gravel	Sand
Asphalt	89%	0%	0%	0%	11%
Asphalt Smooth	0%	100%	0%	0%	0%
Grass	0%	0%	97%	3%	0%
Gravel	11%	0%	2%	85%	2%
Sand	5%	0%	0%	0%	95%
Average probability of correct recognition					94%

(d)

Figure 6.8 Confusion matrices of 24 GHz radar and ultrasonic features using 4 classifiers (a) KNN, (b) MDC, (c) MLP, (d) MLE.

Nevertheless, it was expected that distinguishing asphalt and gravel will give some error due to their similarity of backscattering properties in microwave and ultrasonic range [106].

6.3 Trade-off between Computational Efficiency and Recognition Accuracy

Computational complexity is very important in the road surface identification system, especially when the classification algorithms will be implemented on the vehicle's computer. However, this work is focused on the classification performance. The optimization of the software, which requires the proficiency in programming and computing, is outside the scope of this research. Discussion below should only be considered as a general sense estimation of expected computational performance based on the level of mathematical abstraction of the algorithms and, thus, expected numerical efforts. Parametric methods require far less computation than nonparametric methods and can be used for the initial rough classification between different surfaces. As observed earlier, MDC is less accurate than MLE, but using Mahalanobis distance function improved the accuracy of MDC significantly.

Once the neural network, in the MLP method, is trained and the structure is defined it doesn't require extensive computational resources [106]. Table 6.1 shows the program execution real time (in ms) of the classification algorithms when run in Matlab. The classifiers were run on a computer with Windows 7 64-bit operating system based on Intel Core i7-3770 CPU (3.40 GHz) and the training set used consisted of 1500 measurements. As observed, MDC classifier is the fastest with 1ms execution time. The training of MLP method was 1.5 ms and classification time was only 0.5 ms (fastest). This high speed of MLP is probably

due to its optimization (the standard Matlab nprtool was used). Table 6.1 shows that the slowest classification algorithm is KNN with 7 to 8ms execution time. Nevertheless, this execution time is much faster than the time taken to execute one measurement ($\sim$300 ms) of firing the signal and receiving back the reflection.

The shortcoming of the KNN method and a review of optimisation method is given in [122]. The use of Euclidean distance to its K nearest neighbours lead to a slower performance as compared to the MDC. Since the database is relatively small, training time of MLP was short. However, it should be noted that this time depends on the size of the database, number of inputs, outputs and hidden nodes in MLP. The results we have below cannot be considered as final since the execution time is influenced by the language chosen to implement it. Nevertheless, our executions which were implemented in the same language (MATLAB) confirms the considerations and back up the expected trend found in [122, 123]. Optimizing the execution time and full comparative implementations is suggested for future work.

Classifier	Parameters	Execution time (ms)
K Nearest Neighbour (KNN)	K=2	7.6
K Nearest Neighbour (KNN)	K=3	8.0
Minimum Distance Classification MDC	Euclidean metric	0.9
Minimum Distance Classification MDC	Mahalanobis metric	1.0
Multilayer Perceptron (MLP)	Training	1.5
Multilayer Perceptron (MLP)	Test	0.5

Table 6.1 Classification algorithms execution time

Class	Method	Average probability of correct recognition
Parametric	Minimum Distance Classification MDC (Euclidean)	81%
Parametric	Minimum Distance Classification MDC (Mahalanobis)	91%
Parametric	Maximum Likelihood Estimation(MLE)	94%
Non-parametric	K Nearest Neighbour (KNN)	91%
Non-parametric	Multilayer Perceptron (MLP)	99%

Table 6.2 Average probability of correct recognition of parametric and non-parametric classifiers.

Table 6.2 shows the average probability of correct recognition of the five surfaces for the parametric and non parametric methods when using the optimal radar and ultrasonic features.

Method	Ultrasonic and radar 24 GHz				
	Asphalt	*Smooth asphalt*	*Grass*	*Gravel*	*Sand*
KNN	92%	100%	98%	89%	76%
MDC	84%	100%	98%	87%	86%
MLE	89%	100%	97%	85%	95%
MLP	96%	100%	100%	98%	100%

Table 6.3 Average probability of correct recognition for each surface.

	Surface					Average probability of correct classification
	Asphalt	Smooth asphalt	Grass	Gravel	Sand	
Ultrasonic	84%	100%	85%	87%	100%	92%
Radar 5.8 GHz	79%	92%	96%	85%	84%	87%
Radar 24 GHz	75%	100%	96%	84%	83%	88%
Sonar and radar 5.8 GHz	95%	100%	99%	97%	98%	97%
Sonar and radar 24 GHz	96%	100%	100%	98%	100%	99%
Radar 5.8 GHz and radar 24 GHz	94%	100%	98%	95%	100%	97%
Sonar, radar 5.8 GHz and radar 24 GHz	93%	100%	98%	96%	96%	96%

Table 6.4 Average probability of correct recognition for each data set using MLP

Multilayer Perceptron proved to be the best method to adopt for this investigation with 99% of average probability of correct recognition. Table 6.3 shows the APCR of each surface for all the classifiers. Smooth asphalt and grass are well distinguished with almost 100% accuracy for all the classifiers. Gravel and asphalt are harder to distinguish but the use of MLP improved the correcct recognition to 96 and 98% respectively. KNN and MDC stuggles to distinguish Sand, but MLE achieves better classification (95%), while MLP provides an optimal 100% classification when it comes to sand.

Since the Multilayer Perceptron (MLP) had produced the best classification, a comparison of all the data sets for this method is shown for discussion in table 6.4. It can be shown that the use of two different sensors (ultrasonic and radar) results in a significant increase of the probability of correct recognition. The use of one sensor (ultrasonic or single frequency radar) merely provided a surface identification of 87% to 92%. The use of radar and ultrasonic provides the best performance (97%-99%). However, the combination of dual-frequency radar and ultrasonic has no advantages over combination of single frequency radar and ultrasonic sensor.

6.4 Influence of Exhaust Gases on Measurement's Accuracy

The accuracy of measurements is significantly affected by the vehicle exhaust gases. The experimental set-up described in chapter 4 was used to analyse the influence of the exhaust air on the surface recognition. Several tests were conducted on the basement parking asphalt, with or without running engine of two vehicles: Range Rover Sport and Land Rover Defender.

The exhaust gases are directed to the asphalt area or a carpet (rubber and soft side) two meters away from the sensors (figure 6.11). This will help to understand the error caused by air fluctuations directed towards the object of measurement. Figure 6.12 shows the ultrasonic echo vs. distance reflected from asphalt, soft carpet, and rubber carpet when the engine is ON/OFF. Looking at the 2 meters distance in figure 6.12 (a), when the engine is Off the carpet is clearly distinguished from the asphalt with a difference of more than 10 dBm. One the other hand, carpet is poorly distinguished when the engine is ON with a power difference of 5 dBm (figure 6.12 (b)).
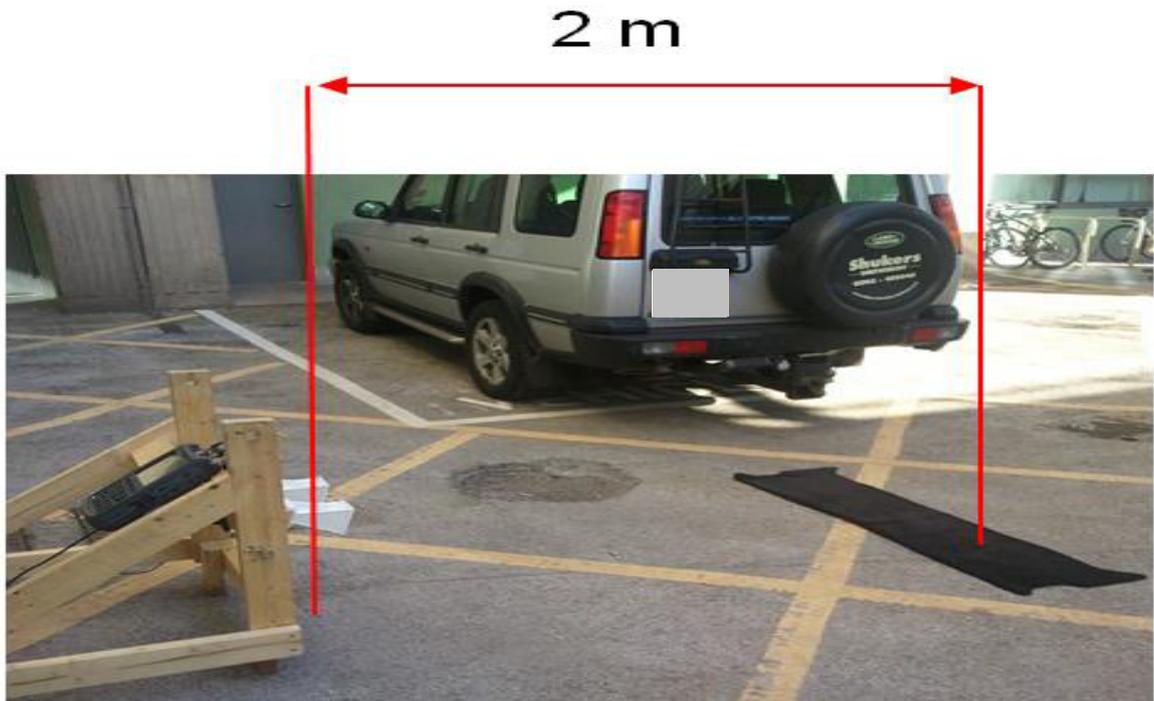
Figure 6.9 Experimental setup for analysis of exhaust gases.
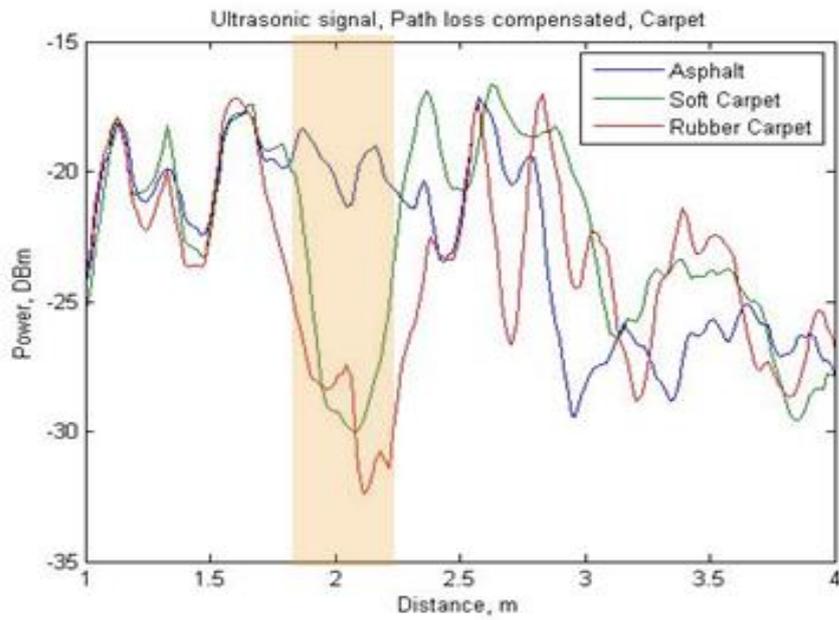
(a) Engine OFF
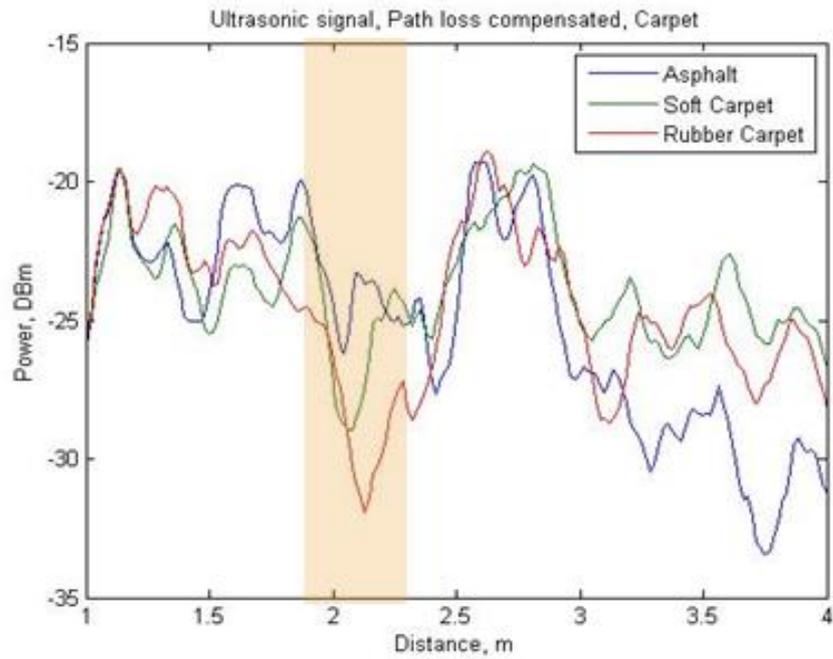
(b) Engine ON

Figure 6.10 Ultrasonic signal power with respect to distance for asphalt, soft carpet, and rubber carpet.

The power reflected from asphalt deteriorated from -20 to -25 dBm when engine is ON.

Asphalt	Power, dBm	Average signal relative stand. dev.	Spread between implementations
Engine OFF	-18.8	0.07	0.10
RR Sport ON	-19.1	0.07	0.65
RR Sport ON /GAS	-20.6	0.09	0.76
LR Defender ON	-20.8	0.08	0.66
LR Defender ON/ GAS	-22.0	0.08	1.04

Table 6.5 Influence of engine characteristics on the ultrasonic signal.

Table 6.5 shows the influence of the engine status of the vehicles on 3 ultrasonic parameters:

averaged signal power, normalised standard deviation, and the spread between

160

implementations. The results, averaged over fifty measurements, show that the power reflected decreases by 4 dBm (-18.8 to -22dBm) when the engine is ON and the air is directed towards the asphalt examined. The average relative standard deviation RTSD is calculated based on the signal envelope within the range gate of 2 meters averaged over 50 measurements. It characterizes the signal envelope. The spread between implementation refers to the deviation of features from one measurement to another. It is the standard deviation of the RTSD over 50 measurements, and it shows this signal changes from one measurement to another. This also increased in a significant way (0.1 to 1.04) when the engine and Gas in ON. Therefore, reflected ultrasonic signal signature and power depend on the engine characteristics (gas or No gas). Air currents from the working engine lead to increasing variance of measured parameters.

The same study was conducted on the radar characteristics. The influence of the engine status has been considered for all the parameters. Figure 6.12 shows the influence on 2 radar parameters (P_{vv}, Std_{vv}) at two frequencies (5.8 and 18 GHz). In this example, the 18GHz vertical polarization average power (dBm) reflected from asphalt is plotted against distance (meters) when the engine is ON/OFF. The difference between the green envelope (engine off) and the blue dotted envelope (engine on) is not significant. The other parameters were affected the same way.
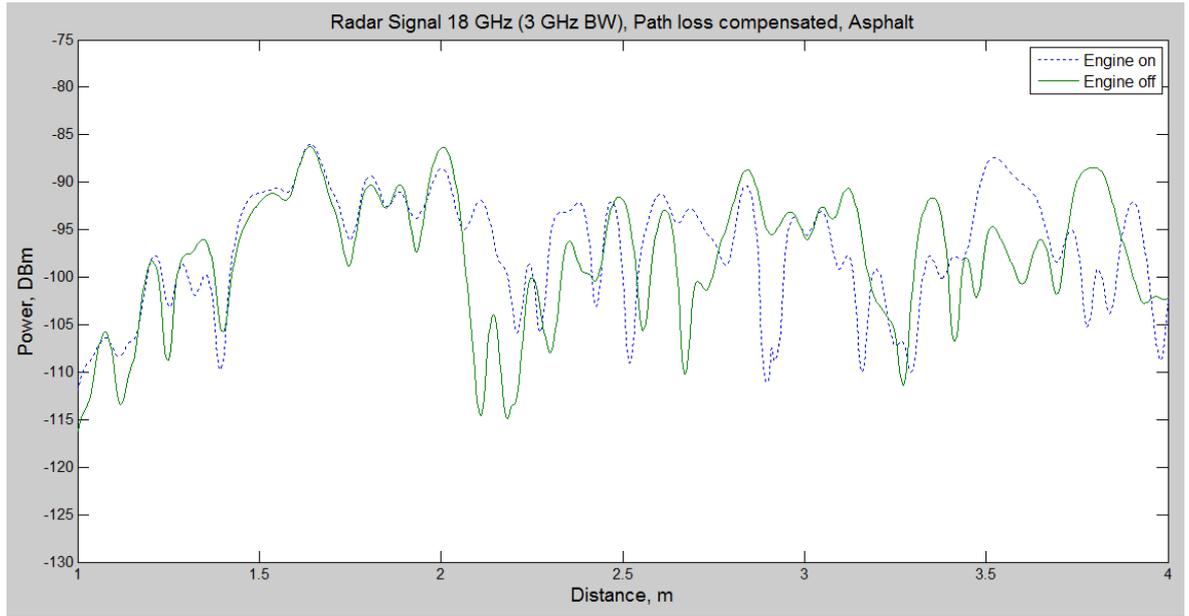
Figure 6.11 18 GHz radar signal average power (dBm) with respect to distance (meters) for asphalt.

Surface	Asphalt		Carpet Soft		Carpet Rubber	
Engine	OFF	GAS	OFF	GAS	OFF	GAS
P_{vv} dBm 5.8	-66.4	-67.3	-67.8	-67.9	-67.7	-67.8
Std_{vv} 18	0.75	0.86	1.02	1.01	1.03	1.04
P_{vv} dBm 18	-93.1	-92.4	-92.5	-92.3	-91.6	-91.2
Std_{vv} 18	0.76	0.75	0.86	0.79	0.74	0.72

Table 6.6 Influence of engine characteristics on the radar signal.

Working engine almost has no influence on the backscattered signal power. Table 6.5 shows the influence of the engine status of the vehicles on 2 radar parameters: averaged signal power and standard deviation (P_{vv}, Std_{vv}) at 5.8 GHz and 18 GHz. As shown in table 6.5, the influence of engine gas is negligible for both frequencies and for the three different

surfaces (asphalt, soft carpet, and rubber carpet). Therefore, the exhausted air has almost no influence on the radar signal power and shape.

6.5 Conclusion

After analysing the performance of four common classification algorithms in the case of surface classification, it was shown that combining radar and ultrasonic sensors can allow a confident distinguishing between different road surfaces. The Multilayer Perceptron method is one of the more practical ways to classify surfaces with a 99 % probability of correct recognition. Other methods like MLE, KNN, and MDC had lower performance, but they require less computation. Each of the method was optimized by using the features extracted from the data set, in addition to the distance functions, number of neighbours, number of neurons, etc. Moreover, a study was conducted on the effect of exhaust gases on the measurement's accuracy. Our results showed that exhaust gases decrease the average power of the reflected ultrasonic signal by 4 dBm. It also increases the data spread from one measurement to another. Microwave sensors are practically devoid of this shortcoming. Therefore, the diversity of sensors can provide optimal solution for surface classification.

Chapter 7

Conclusions and Future Work

7.1   Summary

In the present study, the backscattering properties of road and off-road surfaces radiated by microwave and ultrasonic signals were investigated for the purpose of surface classification. Different probing signals have been used with different carrier frequencies, bandwidths and polarisations. A set of optimal features have been identified to be used in the statistical classification. Four methods of supervised classification were applied to distinguish the following type of surfaces: smooth and rough asphalt, grass, gravel, and sand.

The fundamentals of radar, ultrasonic sensors and pattern recognition were presented in Chapter two. It was stated that surfaces can be characterized by the change of polarisation of the radar electromagnetic waves and change of the echo energy of the ultrasonic signals. The surfaces of interest were reviewed showing the materials and thickness of the roads and off-roads. A survey of literature concerning radar and ultrasonic sensors revealed a lack in the analysis of sensor fusion for surface recognition.

Chapter 2 also presents a literature review of the classification algorithms used for surface recognition. The statistical classification methods including two parametric methods

(Minimum Distance Classification MDC, Maximum Likelihood Estimation MLE) and two non-parametric methods (K nearest Neighbor KNN, Multilayer Perceptron MLP) proved to be suitable for the surface classification using radar and ultrasonic data. Higher accuracy and better performance can be achieved by modifying certain parameters and reducing the data by applying feature extraction. Nevertheless, the goal of most previous studies was to find a difference in the backscattered signals rather than implementing a significant statistically representative number of measurements and studying the efficiency of the statistical classification algorithms. Furthermore, a lack of research about experimenting on different types of surfaces, real surfaces under real weather conditions, and influence of external factors was reported.

In chapter three, the dependence of the backscattered signals on the surface properties (roughness, material, etc.) was analyzed. Surface clutter and radar cross section were presented as two parameters that characterize the surfaces. It was shown that the scattering coefficient highly depends on the dielectric permittivity of the surface, the roughness and polarization. A model of land clutter was developed to show the effect of different surfaces (different properties) and grazing angle on the reflected voltage. The reflected voltage is proportional to the scattering coefficient of the signal. The second part of the chapter highlighted the clutter models used by the previous studies to compare the experimental results. The study showed that the measurement of polarization ratios for different surfaces is the best method to differentiate these surfaces. The average power extracted from the reflected voltage S21 at a certain time gate for different polarizations ratio (VV/HH, VH/HH, and HV/HH) would give promising features for surface recognition.

A full experimentation set of the developed radar and ultrasonic system was described in Chapter four. The experiments took place, in real case scenarios, at several locations around

University of Birmingham and Gaydon, UK. The proposed experimental setup showed that the system is easy to build and can be used to acquire data and visualize it using simple data collection methods. A preliminary study has been conducted on the frequencies and the bandwidths used by the sensors. Moreover, investigation of the effect of grazing angle, exhaust gases, vehicle movement, and weather conditions on the measurements took place. For automotive application, we are restricted by the height of the sensors depending on the bumper of the car and the area ahead of the vehicle required to be covered (5 to 10 meters). This limited the use of a grazing angle between 4.5 and 10 degrees. In the future, as the area to be covered is required to increase up to 15 meters, the grazing angle would be reaching 1.5 to 3 degrees. Studies showed that the influence of the factors (vehicle movement, exhaust gases, etc.) can be reduced. In most cases, weather is not expected to attenuate the signal significantly. Weather conditions depend on many factors like temperature, humidity, and size of water droplets. It is mostly transparent to RF radiations and its effect can be overcome by the diversity of sensors used (Ultrasonic and RF). Collection of a full dataset with different weather conditions, providing more thorough experimental evaluation, is suggested for future work.

In chapter five, a complete classification system has been developed. The classification system consist of three phases: Features selection, training phase and classification phase. A full set of raw variables and extracted features was presented. Features were reduced by the use of two feature selection methods: filters and forward selection method wrapper. A set of optimal features was deduced for radar data, ultrasonic data, and fusion of both radar and ultrasonic. Orange canvas was introduced for data visualization and analysis. Using Orange Canvas and Matlab, the four classification algorithms were applied to the data set, and two methods were used to evaluate the classifiers (confusion matrix and ROC curve). The values of parameters affecting the performance of each classifier were analyzed. It was shown the

KNN is affected by the distance functions and the number of neighbours, while MDC and MLE depends on the distance function and the correlation between features. The optimal number of Neighbours has been assessed and it was shown that 2 neighbours yields the best performance, in most cases, for the surfaces used. The performance measures of the distance functions: Euclidean, Manhattan, Maximal distance, and adaptation of Mahalanobis were provided. Euclidean and Mahalanobis distance offered the highest classification accuracy. Nevertheless, the classification accuracy achieved by the other metrics was not significantly less. Euclidean distance was used for KNN, and MLE, while Mahalanobis function was proposed for the MDC algorithm. On the other hand, MLP method's phases were described. The training phase of the MLP algorithm consists of providing two matrices that define the pattern recognition problem. One matrix contains the full set of unlabelled data with feature values, while the other consists of the labels of the instances. The network is then trained and weights are updated using the conjugate scaled gradient Backpropagation. The performance of MLP performance depends on the number of neurons used and the way the network is trained.

Chapter six illustrates the results obtained by the classification algorithms and evaluates its performance and classification accuracy. Several data sets were investigated separately: ultrasonic sensor data, radar data, and the fusion of both. Confusion matrices were used to show the classification accuracy of KNN, MDC, MLE and MLP for the fives surfaces (asphalt, smooth asphalt, grass, gravel, and sand). Classification based on ultrasonic data on its own showed good differentiation between smooth asphalt, grass, and sands. However, asphalt and gravel has lower classification accuracy due to their similarity in the acoustic properties. The maximum average probability of recognition was achieved by the MLP method (92 %). Classification based on radar data on its own showed lower accuracy than

that of the ultrasonic analysis. The results obtained with 18GHz and 24 GHz radar data sets did not significantly differ; therefore only 18 GHz results were illustrated. Using four optimal features, good differentiation was demonstrated the five surface. The best performance was achieved by MLP method with average probability of recognition hitting 88% using 30 nodes in the hidden layer.

Since each separate sensor (radar or ultrasonic) couldn't provide reliable detection of all five surfaces, we examined the data obtained by the fusion of sensors. The classification based on the fusion of radar data with two frequencies (5.8 GHz and 18 GHz) provided better performance than that of a separate sensor. Again, the best performance achieved was by the MLP method with average probability of recognition of 97 %. KNN provided better recognition (93%) than that of MDC and MLE (82 % and 84% respectively). On the other hand, the classification based on the fusion of ultrasonic and radar data reached the highest classification accuracy achieved in this project. Excellent differentiation was demonstrated between all types of examined surfaces. The performance of MDC provided an APCR of 91%. MLE provided better probability of correct recognition (94%) with a significant improvement especially in terms of sand discrimination. MLP proved to be the best classification method by achieving an average probability of 99 %. Distinguishing smooth asphalt, grass, and sand is 100% accurate using this method. Only 4% of asphalt was misclassified as gravel, and 2 % of gravel was misclassified as asphalt. This was expected due to the similarity of the two surfaces.

The second part of Chapter Six demonstrates the tests conducted to understand the influence of the vehicle exhaust air on the accuracy of measurements. The results showed that the echo signal decreases by 4 dBm in the ultrasonic case (when exhaust gases are directed towards measuring surfaces). The reflected ultrasonic signal signature and power depend on the

engine characteristics (gas or No gas). Air currents from the working engine lead to increasing variance of measured parameters. On the other hand, the influence of engine gas is negligible in the radar case. The exhausted air has almost no influence on the radar signal power and shape. Therefore, the diversity of sensors helps to overcome the influence of exhaust gases on the measurement's accuracy and provides an optimal solution for remote sensing of surface conditions.

7.2 Discussion

Although this research deals with surface classification under stationary conditions, it is necessary to develop practical approach that can be used in the near future. The complexity of the classification algorithms is one of the important factors that affect the application of road surface identification especially when the algorithms will be implemented on the vehicle's computer. The analysis of the parametric and non-parametric classifier showed that parametric methods require far less computation and time complexity. The execution time of each classifier was measured on a training set of 1500 measurements. MDC method was the fastest with 1 ms execution time, while KNN was the slowest with 8 ms execution time. MLP had a very high speed of 1.5 ms for training and 0.5 ms for testing. The four classifiers proved to be applicable in terms of time complexity, since even the slowest KNN method allows about ten measurements at 1 m distance in the case of driving with a speed of 50 km per hour.

This thesis proves the feasibility of the fusion of radar and ultrasonic data for surface classification. The results given in this thesis demonstrate that a surface can be classified

with excellent accuracy using the Multilayer Perceptron neural network. These two sensors provide sufficiently high sensitivity for surface classification and are robust to external conditions. The classification performance is significantly improved by reducing the features and optimizing the algorithms. The diversity of sensors improves the performance reliability of surface recognition and reduces the errors caused by the air fluctuations and weather conditions.

A new experimental setup was used for the ultrasonic sensor on its own. The experiments were conducted while mounting the ultrasonic sensor on a vehicle moving with a speed 0.5 m/s (2 km/hour). Three measurements were taken per second, and averaging was applied by a sliding window over 10 sweeps. Feature optimizations have not been done at this stage and the features used were (average power, standard deviation, power above threshold, and duration above threshold). Classification algorithms were instantly applied while taking the measurements using a new dynamic algorithm. Figure 7.1 shows a representation of the correct recognition of the system with respect to time. The surface of interest is highlighted in red. It is shown in figure 7.1(a) that MLE correctly predicted the surface grass between 0 and 55 seconds. MLE achieved a correct recognition of 97% (achieved 90% in the stationary case), while KNN achieved a correct recognition of 83 % (87% in the stationary case). Figure (b) shows that KNN achieved 100 % for the recognition of smooth asphalt, while MLE achieved 94%
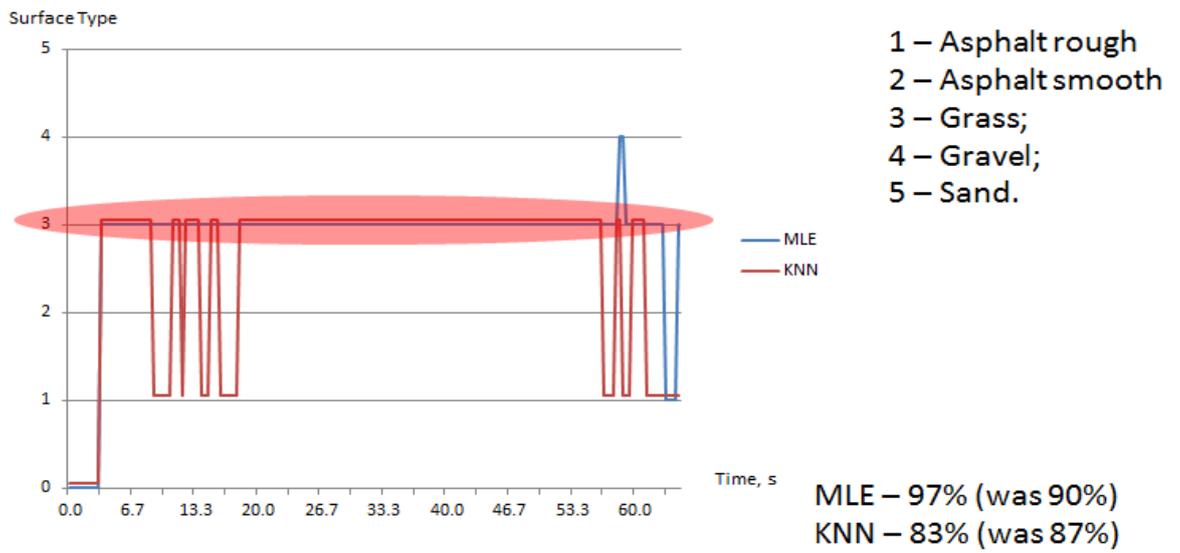
(a) Grass
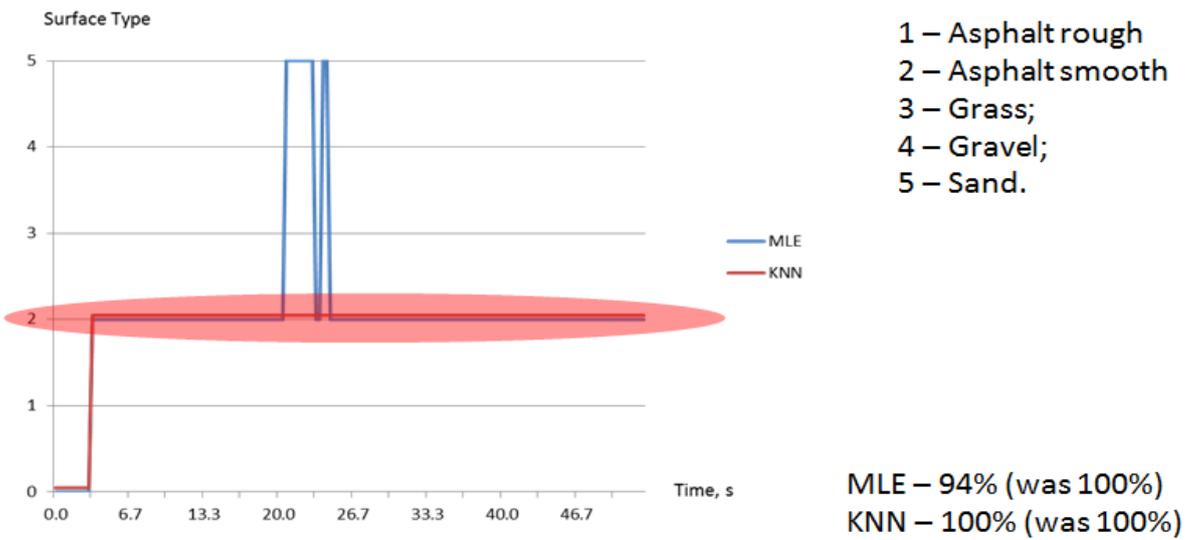
(b) Smooth asphalt

Figure 7.1 Probability of correct recognition in the dynamic case.

7.3 Future Work

Although the main concept of this research has been proven, the goal is to take this system much further to provide road surface classification in the dynamic conditions. The radar and ultrasonic system proposed clearly demonstrates its competent performance in the classification of surfaces. Despite this, several aspects could be investigated to explore the potential of this system for further improvement. We have initiated the work on some of these aspects in a hope that a dynamic surface classification system will be available in the automotive market soon.

1. The smaller the number of surfaces, the easier to distinguish between them. A priori knowledge of the environment can help to increase the system performance by ignoring some of the surfaces. For example, there no need to consider ice or snow when the outside temperature is significantly above zero.

2. Surfaces covered with snow are not easily distinguishable at 5.8 and 24 GHz. The snow depth of 3 to 5 cm doesn't allow the radar, and even the ultrasonic sensor, used to penetrate its layer. Nevertheless, detecting snow is feasible since the backscattering properties of snow differ from the properties of clear asphalt and ice on asphalt. Although this task was on the research plan, there was a lack of training data due to the lack of snow and ice in the UK during the research time. An initial study has been conducted on the classification between asphalt, snow and ice. The study was based on the few measurements which took place on snow and ice covered surfaces at the same locations used in this research. There is a need to perform more measurements on surfaces covered with ice and snow.

3. Finally, this work is dedicated to defining and selecting the parameters which will enhance the recognition performance. The collection of data which took 2 years composed a statistical representative of the surfaces. There is a need to collect more data for the following future investigations:

I.   Performing the dynamic measurements on more surface and more classification algorithms. Performing dynamic tests using ultrasonic to analyse the dependence of performance vs. speed and weather conditions.

II.   Developing a program for MLP neural network, as Matlab toolbox does not work in dynamic condition.

III. Updating the training database, collecting more data, applying feature selection, and optimizing the classification algorithms to achieve as high classification accuracy as the static condition.

IV.   Performing experimental evaluations to analyse the effect of weather conditions on the performance of the system.

V.   Providing statistical estimation of the effect of grazing angle by collecting data using the following grazing angles:

- $5°$ grazing angle for short range sensing up to 5 meters.

- $3°$ grazing angle for longer sensing distnaces where range between minimum and maximum grazing angles will be even smaller.

Chapter 8

References

[1] "Remote Sensing Technology for Automotive Safety." [Online]. Available: http://www.microwavejournal.com/articles/5689-remote-sensing-technology-for-automotive-safety. [Accessed: 26-Jan-2015].

[2] REEVES, *Manual of Remote Sensing, Volume 1*, First Edition edition. American Society of Photogrammetry, 1975.

[3] "Barry's Remote Sensing Page." [Online]. Available: http://www.sarracenia.com/astronomy/remotesensing/primer0120.html. [Accessed: 26-Jan-2015].

[4] M. I. Skolnik, *Introduction to Radar Systems*, 3 edition. Boston: McGraw-Hill Higher Education, 2002.

[5] "RAC Foundation - the independent motoring charity." [Online]. Available: http://www.racfoundation.org/. [Accessed: 26-Jan-2015].

[6] "Reported road accident statistics - Commons Library Standard Note," *UK Parliament*. [Online]. Available: http://www.parliament.uk/business/publications/research/briefing-papers/SN02198/reported-road-accident-statistics. [Accessed: 26-Jan-2015].

[7] "Jaguar Land Rover's new 'sensing' technology to beat drowsy driving," *coventrytelegraph*. [Online]. Available: http://www.coventrytelegraph.net/news/coventry-news/jaguar-land-rover-unveils-new-8414322. [Accessed: 12-Feb-2015].

[8] "The Factors That Cause Car Crashes and Traffic Accidents." [Online]. Available: http://www.driving-test-success.com/causes-car-crash.htm. [Accessed: 12-Feb-2015].

[9] L. Vlacic, M. Parent, and F. Harashima, Eds., *Intelligent Vehicle Technologies: Theory and Applications*. Warrendale, PA: Society of Automotive Engineers Inc, 2001.

[10] "Bistatic radar - Telecom ABC." [Online]. Available: http://www.telecomabc.com/b/bistatic-radar.html. [Accessed: 26-Jan-2015].

[11] "Microwave Engineering 3rd Edition David M.Pozar [www.itepress.com]," *Scribd*. [Online]. Available: https://www.scribd.com/doc/113200280/Microwave-Engineering-3rd-Edition-David-M-Pozar-www-itepress-com. [Accessed: 26-Jan-2015].

[12] M. I. Skolnik, *Radar Handbook, Third Edition*, 3 edition. New York: McGraw-Hill Professional, 2008.

[13] K. S. F. T. Ulaby, "Statistical properties of the Mueller matrix of distributed targets," *Radar Signal Process. IEE Proc. F*, no. 2, pp. 136 – 146, 1992.

[14] V. V. Viikari, T. Varpula, and M. Kantanen, "Road-Condition Recognition Using 24-GHz Automotive Radar," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 4, pp. 639–648, Dec. 2009.

[15] V. Viikari, T. Varpula, and M. Kantanen, "automotive radar technology for detecting road conditions. Backscattering properties of dry, wet, and icy asphalt," in *Radar Conference, 2008. EuRAD 2008. European*, 2008, pp. 276–279.

[16] B. R. J. Snuttjer and R. M. Narayanan, "Millimeter-wave backscatter measurements in support of surface navigation applications," in *Geoscience and Remote Sensing Symposium, 1996. IGARSS '96. "Remote Sensing for a Sustainable Future.", International*, 1996, vol. 1, pp. 506–508 vol.1.

[17] L. Giubbolini and G. Wanielik, "A microwave coherent polarimeter with velocity gating for road surface monitoring," in *13th International Conference on Microwaves, Radar and Wireless Communications. 2000. MIKON-2000*, 2000, vol. 1, pp. 355–358 vol.1.

[18] K. Sarabandi, E. S. Li, and A. Nashashibi, "Modeling and measurements of scattering from road surfaces at millimeter-wave frequencies," *IEEE Trans. Antennas Propag.*, vol. 45, no. 11, pp. 1679–1688, Nov. 1997.

[19] J. Lin, C. R. Liu, J. Li, and X. Chen, "Measurement of concrete highway rough surface parameters by an X-band scatterometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1188–1196, Jun. 2004.

[20] J. Hakli, J. Saily, P. Koivisto, I. Huhtinen, T. Dufva, A. Rautiainen, H. Toivanen, and K. Nummila, "Road surface condition detection using 24 GHz automotive radar technology," in *Radar Symposium (IRS), 2013 14th International*, 2013, vol. 2, pp. 702–707.

[21] B. I. Tapkan, S. Yoakum-Stover, and R. F. Kubichek, "ACTIVE MICROWAVE REMOTE SENSING OF ROAD SURFACE CONDITIONS," in *Transportation Research Board Conference Proceedings*, 1997.

[22] J. H. Y. Il Young Song, "Classification of road surface status using a 94 GHz dual-channel polarimetric radiometer," *Int. J. Remote Sens. - INT J REMOTE SENS*, vol. 33, no. 18, pp. 5746–5767, 2012.

[23] R. Finkele, A. Schreck, and G. Wanielik, "Polarimetric road condition classification and data visualisation," in *Geoscience and Remote Sensing Symposium, 1995. IGARSS '95. "Quantitative Remote Sensing for Science and Applications", International*, 1995, vol. 3, pp. 1786–1788 vol.3.

[24] H. Rudolf, G. Wanielik, and A. J. Sieber, "Road condition recognition using microwaves," in *, IEEE Conference on Intelligent Transportation System, 1997. ITSC '97*, 1997, pp. 996–999.

[25] N. Kees and J. Detlefsen, "Road surface classification by using a polarimetric coherent radar module at millimeter waves," in *Microwave Symposium Digest, 1994., IEEE MTT-S International*, 1994, pp. 1675–1678 vol.3.

[26] W. Hetzner, "Recognition of road conditions with active and passive millimetre-wave sensors," vol. 38, no. 7/8, pp. 179–185, 1984.

[27] R. Finkele, "Detection of ice layers on road surfaces using a polarimetric millimetre wave sensor at 76 GHz," *Electron. Lett.*, no. 13, pp. 1153 – 1154, 1997.

[28] W. P. G. Magerl, "Remote Sensing Of Road Condition," pp. 2137–2140, 1991.

[29] P. H. R Schubert, "Microwave Doppler Sensors Measuring Vehicle Speed and Travelled Distance: Realistic System Tests in Railroad Environment."

[30] B. B. Kelly, B. Hardel, T. Moseley, and J. Heimdahl, "Soil Moisture Estimation Using Air-Launched GPR Techniques.," Apr-2009. [Online]. Available: http://digital.library.wisc.edu/1793/36099.

[31] Herry Schulz, B. Kelly, "Monitoring Near-Surface Soil Water Content Using Air-Launched GPR Techniques.," *Univ. Wis.*

[32] B. Ma, S. Lakshmanan, and A. O. Hero, "Simultaneous detection of lane and pavement boundaries using model-based multisensor fusion," *IEEE Trans. Intell. Transp. Syst.*, 2000.

[33] P. S. H. Shiu Hang Tsang, "Advance Path Measurement for Automotive Radar Applications," *Intell. Transp. Syst. IEEE Trans. On*, no. 3, pp. 273 – 281, 2006.

[34] J. P. Andreas Eidehall, "Toward Autonomous Collision Avoidance by Steering," *Intell. Transp. Syst. IEEE Trans. On*, no. 1, pp. 84 – 94, 2007.

[35] R. Schneider, G. Wanielik, and J. Wenger, "Radar arrangement for road condition detection in motor vehicle," German Paternt DE 19715999, Oct.1998.

[36] R. Finkele, A. Schreck, and G. Wanielik, "Multi-sensor advance detection of road conditions in front of vehicle using narrow beam

millimeter radar and infrared radar beam targeting a road section in front of vehicle," German Patent DE 19 932 094, Jan. 2001.

[37]   H. S. Kim, "Road surface sensing device," Korean Patent KR 2001 047 234, Jun.2001.

[38]   V. Viikari and T. Varpula, "A method for road condition recognition," EP2216659 A1, 11-Aug-2010.

[39]   L. Kleeman and R. Kuc, "Sonar Sensing," in *Springer Handbook of Robotics*, B. S. Prof and O. K. Prof, Eds. Springer Berlin Heidelberg, 2008, pp. 491–519.

[40]   J. Ogilvy, *Theory of Wave Scattering From Random Rough Surfaces,*. Bristol, England ; Philadelphia: CRC Press, 1991.

[41]   P. Probert Smith and K. Zografos, "Sonar for recognising the texture of pathways," *Robot. Auton. Syst.*, vol. 51, no. 1, pp. 17–28, Apr. 2005.

[42]   G. Kao and P. Probert, "Feature Extraction from a Broadband Sonar Sensor for Mapping Structured Environments Efficiently," *Int. J. Robot. Res.*, vol. 19, no. 10, pp. 895–913, Oct. 2000.

[43]   KINSLER, *FUNDAMENTALS OF ACOUSTICS FOURTH 4TH EDITION*, Fourth Edition edition. BBS, 2009.

[44]   H. E. Bass and F. D. Shields, "Absorption of sound in air - High-frequency measurements," *Acoust. Soc. Am. J.*, vol. 62, pp. 571–576, Sep. 1977.

[45]   P. J. McKerrow and B. J. Kristiansen, "Classifying surface roughness with CTFM ultrasonic sensing," *Fac. Inform. - Pap.*, 2006.

[46]   K. Zografos and P. P. Smith, "Rough surfaces classification for environmental perception of a mobile robot using CTFM sonar imaging and neural networks," *IEEE Adv. Sess. 2001*, pp. 585–590.

[47]   O. Bozma and R. Kuc, "Characterizing The Environment Using Echo Energy, Duration, And Range: The Endura Method," in *, Proceedings of the 1992 lEEE/RSJ International Conference on Intelligent Robots and Systems, 1992*, 1992, vol. 2, pp. 813–820.

[48]   O. Bozma and R. Kuc, "A physical model-based analysis of heterogeneous environments using sonar-ENDURA method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 5, pp. 497–506, May 1994.

[49]   Z. Politis and P. J. P. Smith, "Classification of Textured Surfaces for Robot Navigation Using Continuous Transmission Frequency-Modulated Sonar Signatures," *Int. J. Robot. Res.*, vol. 20, no. 2, pp. 107–128, Feb. 2001.

[50]   P. N. Pathirana and A. Zaknich, "Surface identification by acoustic reflection characteristics using time delay spectrometry and artificial neural networks," in *, International Conference on Neural Networks,1997*, 1997, vol. 1, pp. 31–36 vol.1.

[51] J. E. Wilhjelm, P. C. Pedersen, and S. M. Jacobsen, "The influence of roughness, angle, range, and transducer type on the echo signal from planar interfaces," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 48, no. 2, pp. 511–521, Mar. 2001.

[52] N. Roy, G. Dudek, and P. Freedman, "Surface sensing and classification for efficient mobile robot navigation," in *, 1996 IEEE International Conference on Robotics and Automation, 1996. Proceedings*, 1996, vol. 2, pp. 1224–1228 vol.2.

[53] K. Sasaki and M. Takano, "Classification Of Objects' Surface By Acoustic Transfer Function," in *, Proceedings of the 1992 lEEE/RSJ International Conference on Intelligent Robots and Systems, 1992*, 1992, vol. 2, pp. 821–828.

[54] H. Kobayashi, T. Kimura, and M. Negishi, "Ultrasonic ground speedometer utilizing doppler effect," US5054003 A, 01-Oct-1991.

[55] "Road Structure - explains pavements, rigid, flexible and composite roads and the various layers of the carrigeway." [Online]. Available: http://incatrad.com/highway.htm. [Accessed: 03-Feb-2015].

[56] "Asphalt," *Wikipedia, the free encyclopedia*. 16-Feb-2015.

[57] G. of C. and A.-F. Canada, "Canadian Soil Information Service," 07-Feb-2013. [Online]. Available: http://sis.agr.gc.ca/cansis/publications/manuals/1976-glossary/index.html. [Accessed: 03-Feb-2015].

[58] Ö. Bozma and R. Kuc, "Characterizing pulses reflected from rough surfaces using ultrasound," *J. Acoust. Soc. Am.*, vol. 89, no. 6, pp. 2519–2531, Jun. 1991.

[59] A. Rosenfeld, A. C. Kak, and A. C. Kak, *Digital Picture Processing, Volume 1, Second Edition*, 2 edition. New York: Morgan Kaufmann, 1982.

[60] I. H. Witten, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Francisco, Calif: Morgan Kaufmann Publishers In, 1999.

[61] J. Demšar, *Data Mining: Material for Lectures on Data Mining at Kyoto University, Dept. of Health Informatics in July 2010*. 2010.

[62] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley & Sons, 2012.

[63] M. James, *Classification Algorithms*. Wiley, 1985.

[64] B. G. Batchelor, *Pattern Recognition: Ideas in Practice*, 1978 edition. New York: Springer, 1978.

[65] A. G. Karegowda, M. A. Jayaram, and A. S. Manjunath, "Cascading K-means Clustering and K-Nearest Neighbor Classifier for Categorization of Diabetic Patients," *Int. J. Eng. Adv. Technol. IJEAT*, vol. 1, no. 2, Feb. 2012.

[66]   B. K. Lavine, "Clustering and Classification of Analytical Data," in *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd, 2006.

[67]   D. J. Hand, *Discrimination and classification*. J. Wiley, 1981.

[68]   C. M. Bishop, *Neural Networks for Pattern Recognition*. Clarendon Press, 1995.

[69]   S. O. Haykin, *Neural Networks: A Comprehensive Foundation*, 2 edition. Upper Saddle River, N.J.: Pearson, 1997.

[70]   I. Millington and J. Funge, *Artificial Intelligence for Games*. CRC Press, 2009.

[71]   J. P. Paulo and J. L. B. Coelho, "A HYBRID METHOD FOR IDENTIFICATION OF ROAD PAVEMENT TYPES USING A BAYSEAN ANALYSIS AND NEURAL NETWORKS," *Cairo*, no. ICSV 17 :: International Congress on Sound & Vibration, 2010.

[72]   J. P. Paulo, J. L. B. Coelho, and M. A. T. Figueiredo, "Statistical classification of road pavements using near field vehicle rolling noise measurements," *J. Acoust. Soc. Am.*, vol. 128, no. 4, pp. 1747–1754, Oct. 2010.

[73]   J. Alonso, J. M. López, I. Pavón, M. Recuero, C. Asensio, G. Arcas, and A. Bravo, "On-board wet road surface identification using tyre/road noise and Support Vector Machines," *Appl. Acoust.*, vol. 76, pp. 407–415, Feb. 2014.

[74]   V. S. Il Young Song, "Robust urban road surface monitoring system using Bayesian classification with outlier rejection algorithm," pp. 1118–1122, 2012.

[75]   B. R. Mahafza, *Radar Systems Analysis and Design Using MATLAB Third Edition*, 3 edition. Boca Raton ; London: Chapman and Hall/CRC, 2013.

[76]   G. Brooker, *Introduction to Sensors for Ranging and Imaging*, First edition. Raleigh, NC: SciTech Publishing, 2009.

[77]   M. W. Long, *Radar Reflectivity of Land and Sea*, 2 edition. Dedham, MA: Artech House Publishers, 1983.

[78]   D. K. Barton, *Radar Clutter*. Dedham, Mass: Artech Print on Demand, 1975.

[79]   M. Skolnik, *Radar Handbook, Third Edition*. McGraw Hill Professional, 2008.

[80]   V. V. Viikari, T. Varpula, and M. Kantanen, "Road-Condition Recognition Using 24-GHz Automotive Radar," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 4, pp. 639–648, Dec. 2009.

[81]   F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Volume II: Radar Remote Sensing and Surface Scattering and Emission Theory*. Reading, Mass.: Artech House Publishers, 1986.

[82]   Y. Dong, "Models of Land Clutter vs Grazing Angle, Spatial Distribution and Temporal Distribution - L-Band VV Polarisation Perspective," Mar. 2004.

[83]   F. T. Ulaby, P. P. Batlivala, and M. C. Dobson, "Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, and Soil Texture: Part I-Bare Soil," *IEEE Trans. Geosci. Electron.*, vol. 16, no. 4, pp. 286–295, Oct. 1978.

[84]   F. T. Ulaby and M. C. Dodson, *Handbook of Radar Scattering Statistics for Terrain*. Norwood, MA: Artech House Publishers, 1989.

[85]   D. A. Boyarskii, V. V. Tikhonov, and N. Y. Komarova, "Model of Dielectric Constant of Bound Water in Soil for Applications of Microwave Remote Sensing - Abstract," *J. Electromagn. Waves Appl.*, vol. 16, no. 3, pp. 411–412, Jan. 2002.

[86]   G. T. Ruck, *Radar cross section handbook*. Plenum Press, 1970.

[87]   E. F. Knott, J. Shaeffer, and M. Tuley, *Radar Cross Section, Second Edition*. SciTech Publishing, 2004.

[88]   A. J. Gatesman, T. M. Goyette, J. C. Dickinson, R. H. Giles, J. Waldman, J. Sizemore, R. M. Chase, and W. E. Nixon, "Polarimetric backscattering behavior of ground clutter at X, Ka, and W-band," presented at the Algorithms for Synthetic Aperture Radar Imagery XII, 2005, vol. 5808, pp. 428–439.

[89]   R. E. Clapp, Massachusetts Institute of Technology., and Radiation Laboratory., *A theoretical and experimental study of radar ground return*. Cambridge, Mass.: Radiation Laboratory, Massachusetts Institute of Technology, 1946.

[90]   L. Spetner and I. Katz, "Two statistical models for radar terrain return," *IRE Trans. Antennas Propag.*, vol. 8, no. 3, pp. 242–246, May 1960.

[91]   P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, New edition edition. Norwood, MA: Artech House Publishers, 1987.

[92]   "SRF08 Ultra sonic range finder." [Online]. Available: http://www.robot-electronics.co.uk/htm/srf08tech.shtml. [Accessed: 11-Feb-2015].

[93]   "The Wireless Telegraphy (Automotive Short Range Radar) (Exemption) Regulations 2013." [Online]. Available: http://www.legislation.gov.uk/uksi/2013/1437/made. [Accessed: 11-Feb-2015].

[94]   S. W. Smith, *Digital Signal Processing: A Practical Guide for Engineers and Scientists*. Newnes, 2003.

[95]   G. Brooker, *Introduction to sensors for ranging and imaging*. [Raleigh, N.C.]: SciTech Pub., 2011.

[96] retro7, "VW Suspension Tuning | RetroRims VW Alloy & Steel Wheel Blog." .

[97] N. J. Veck, "Atmospheric transmission and natural illumination (visible to microwave regions)," *GEC J. Res.*, vol. 3, pp. 209–223, 1985.

[98] D. Deirmendjian, "Extinction of Submillimeter Waves by Clouds and Rain," *Int. Conf. Submillimeter Waves Their Appl.*, Jun. 1974.

[99] C.-C. Chen, "Attenuation of Electromagnetic Radiation by Haze, Fog, Clouds, and Rain," 1975. [Online]. Available: http://www.rand.org/pubs/reports/R1694.html. [Accessed: 03-Nov-2014].

[100] R. J. Donaldson, "The measurement of cloud liquid-water content by radar," *J. Meteorol.*, vol. 12, no. 3, pp. 238–244, Jun. 1955.

[101] D. E. Kerr, *Propagation of Short Radio Waves*. IET, 1951.

[102] R. G. Eldridge, "Haze and Fog Aerosol Distributions.," *J. Atmospheric Sci.*, vol. 23, pp. 605–613, Sep. 1966.

[103] "ISO 9613-1:1993 - Acoustics -- Attenuation of sound during propagation outdoors -- Part 1: Calculation of the absorption of sound by the atmosphere." [Online]. Available: http://www.iso.org/iso/catalogue_detail.htm?csnumber=17426. [Accessed: 11-Feb-2015].

[104] "Absorption." [Online]. Available: http://www.acs.psu.edu/drussell/Demos/Absorption/Absorption.html. [Accessed: 03-Nov-2014].

[105] "Calculation of absorption of sound by the atmosphere." [Online]. Available: http://resource.npl.co.uk/acoustics/techguides/absorption/. [Accessed: 11-Feb-2015].

[106] A. Bystrov, M. Abbas, E. Hoare, Thuy-Yung Tran, N. Clarke, M. Gashinova, and M. Cherniakov, "Analysis of algorithms of road surface classification," presented at the 2015 IEEE International Radar Conference.

[107] S. S. Stevens, "On the Theory of Scales of Measurement," *Science*, vol. 103, no. 2684, pp. 677–680, Jun. 1946.

[108] B. Azhagusundari and A. Selvadoss Thanamani, "Feature Selection based on Information Gain," *Int. J. Innov. Technol. Explor. Eng. IJITEE ISSN*, vol. 2, no. 2, pp. 2278–3075, anuary 2013.

[109] A. Janecek, W. Gansterer, M. A. Demel, and G. F. Ecker, "On the relationship between feature selection and classification accuracy," in *Journal of Machine Learning Research Workshop and Conference Proceedings 4*, -, 2008, vol. 91–106, pp. 90–105.

[110] "Orange Data Mining." [Online]. Available: http://orange.biolab.si/. [Accessed: 11-Feb-2015].

[111] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

[112] P. Hall, B. U. Park, and R. J. Samworth, "Choice of neighbor order in nearest-neighbor classification," *Ann. Stat.*, vol. 36, no. 5, pp. 2135–2152, Oct. 2008.

[113] "Chebychev Distance Metric." [Online]. Available: http://www.improvedoutcomes.com/docs/WebSiteDocs/Clustering/Clustering_Parameters/Chebychev_Distance_Metric.htm. [Accessed: 11-Feb-2015].

[114] "Mahalanobis Metric." [Online]. Available: http://www.cs.princeton.edu/courses/archive/fall08/cos436/Duda/PR_Mahal/M_metric.htm. [Accessed: 11-Feb-2015].

[115] S. Joken, N. Inoue, and Y. Yamashita, "Numerical analysis of Mahalanobis metric in vector space," in *19th International Conference on Pattern Recognition, 2008. ICPR 2008*, 2008, pp. 1–4.

[116] E. Alpaydin, *Introduction to Machine Learning*, 2nd Revised edition edition. Cambridge, Mass: MIT Press, 2010.

[117] "Divide Data for Optimal Neural Network Training - MATLAB & Simulink - MathWorks United Kingdom." [Online]. Available: http://uk.mathworks.com/help/nnet/ug/divide-data-for-optimal-neural-network-training.html. [Accessed: 11-Feb-2015].

[118] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Netw.*, vol. 6, no. 4, pp. 525–533, 1993.

[119] R. P. L. Durgabai, " Feature Selection using ReliefF Algorithm" in *International Journal of Advanced Research in Computer and Communication, vol. 3, issue 10, Oct.2014*.

[120] B. Azhagusundari, A. S. Thanamani, " Feature Selection based on Information Gain" in *International Journal of Innovative Technology and Exploring Engineering, vol. 2, issue 2, ISSN 2278-3075, Jan.2013*.

[121] S. Lei, " A Feature Selection Method Based on Information Gain and Genetic Algorithm" in *2012 International Conference on Computer Science and Electronics Engineering, pp. 355- 358, 2012*.

[122] N. Bhatia et al, Survey of Nearest Neighbor Techniques. International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010.

[123] MF. Amasyali, OK. Ersoy , Comparison of single and ensemble classifiers in terms of accuracy and execution time. In: International symposium on innovation intelligent system applications. doi:10.1109/INISTA.2011.5946119, 2011

Appendix A: Structure of Roads
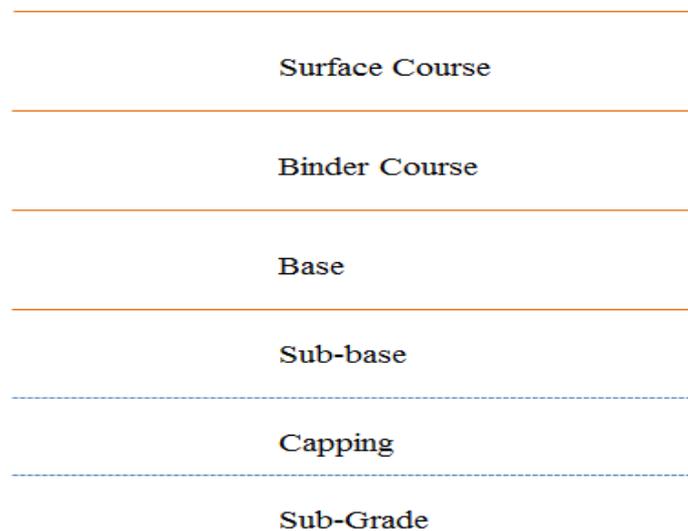
Structure of a flexible road[55].

1.   **The Sub-base:** The Sub-base is the first layer of the road structure. This layer is meant to protect the surface beneath from the damage caused by rain or sun. The materials used in the sub-bases, commonly termed Type 1, are made of crushed rocks, slags, crushed concrete, or non-plastic shale. Sand might constitute up to 10% of the sub-base. Other materials like bituminous-bound materials and concrete can be used for the construction of sub-bases.

2.   **Base**: The Base is the second layer of the road structure. Type 1 is also the most commonly used material. Other materials like slag bound material are used as well. The materials must be compacted and spread evenly to protect them from the drying or wetting changes. The layers constituting the Base are of 110mm-225mm thickness.

3. **Binder course and Surface course:** These two layers are known as the surfacing of the road. Usually surfacing is made of two course binder and surface. The binder course, with typical thickness of 45 mm and 105mm, is the base of the thinner surface course. It uses stones of 20, 28 or 40 mm size and materials including macadam and rolled asphalt. Surface courses are made of a range of bituminous materials. Nevertheless, the traffic intensity plays a role in selecting the surface course material. In the UK, the most common surface materials used are hot rolled asphalt (HRA) and stone mastic asphalt (SMA). HRA is a combination of asphaltic cement with crushed rocks, slag or gravel. It has a thickness of 40mm. SMA reduces surface noise and its thickness is between 20 and 40 mm [55].

The lifetime of a road in the UK is 40 years. Nevertheless, traffic increase is a major factor affecting the life time of a road. Major reconstruction should be carried out in the middle of the road life time to maintain its life.

Appendix B: Implementation of Classification Algorithms

K Nearest Neighbor (KNN)

1. Let x_i (A_i, B_i, C_i) be the test point with three features A, B and C. let D be the set of training samples $y_j(A_j, B_j, C_j)$ belonging to M classes (D1, D2,…, DM). K is initially set as the number of nearest neighbors.

2. For each test point x_i , the algorithm computes the distance between features $d(x_i, y_j)$ using a certain distance function for every training sample y_j of D .

3. Finds the K closet training sample y_j in each of the M classes to the test point x_i .

4. Classifies x_i according to the class of the nearest neighbours with the minimum distance $d(x_i, y_j)$.

Minimum Distance Classification (MDC)

1. Let x_i (A_i, B_i, C_i) be the test point with three features A, B and C. let D be the set of

training samples $y_j(A_j, B_j, C_j)$ belonging to belonging to M classes (D1, D2,…, DM).

2. Represent each training class by a mean vector m_l :

$$m_l = \frac{\Sigma_{y \in D_l}\, y}{N_l}$$

where N_l is the number of training samples y_j in the class D_l, and $l=1,2,…,$M.

3. For each test point x_i , the algorithm computes the distance between features d(x_i, m_l) using a certain distance function for every class D.

4. Classify x_i according to the class D_l of the nearest mean with the minimum distance d(x_i, m_l).

Maximum Likelihood Estimation (MLE)

1. It assumes the p (y/D_l) can be determined by a known parameter vector θ_l that consists of mean and covariance matrix μ_l and $\sigma^2{}_l$, respectively.

2. It uses the training samples to obtain estimates for the unknown parameter vectors $(\theta_1, \theta_2,…,\theta_l)$. These parameters of different classes are assumed functionally independent.

3. Since the samples y_j are drawn independently, the likelihood of θ with respect to the samples is given by the likelihood function:

$$P(D/ \theta) = f (y_1, y_2, ..., y_j/\theta) = \prod_{k=1}^{j} p \ (y_k/ \theta)$$

4. It is more convenient to work with the logarithm of the likelihood. The value that maximizes the log-likelihood will also maximize the likelihood since the algorithm is monotonically increasing:

$$L (\theta) = \ln P(D/ \theta) = \sum_{k=1}^{j} \ln p \ (y_k/ \theta)$$

5. The best estimate for mean and variance is the one that maximizes the probability to obtain the samples actually observed [62]. The mean and variance that maximize the log-likelihood occurs when the partial derivative $L(\theta)$ equals to zero. Solving the equation of derivative with respect to the desired variable:

$$\nabla_\theta \ L = 0$$

yields the MLE estimates of the mean and variance :

$$\hat{\mu} = \frac{1}{N_l} \sum_{k=1}^{j} y_k$$

and

$$\hat{\sigma}^2 = \frac{1}{N_l} \sum_{k=1}^{j} \ (y_k - \hat{\mu})^2$$

Appendix C: Distance Functions

1. Euclidean distance: let 'a' be an instance of features $(a_1, a_2, ..., a_k)$, and 'b' another instance of features $(b_1, b_2, ..., b_k)$, where k is the number of features. The Euclidean distance between the two instances is equal to the square root of the sum of squares of the differences between the values:

$$D = \sqrt{\sum_{i=1}^{k}(a_i - b_i)^2} \tag{1}$$

Some operations do not include the square root when comparing the distances; others take higher powers rather than squaring. This increases the effect of high differences, and dwarfs that of low differences.

2. Manhattan distance: This is an alternative distance of Euclidean where the difference between the values is not squared. The metric, called City-clock metric in some references, is the distance from one data point to the other following a grid-like path. The formula for Manhattan distance between two instances 'a' and 'b' is given by:

$$D = \sum_{i=1}^{k}|(a_i - b_i)| \tag{2}$$

Figure 1 shows the different paths of the two distance functions for two points a and b.
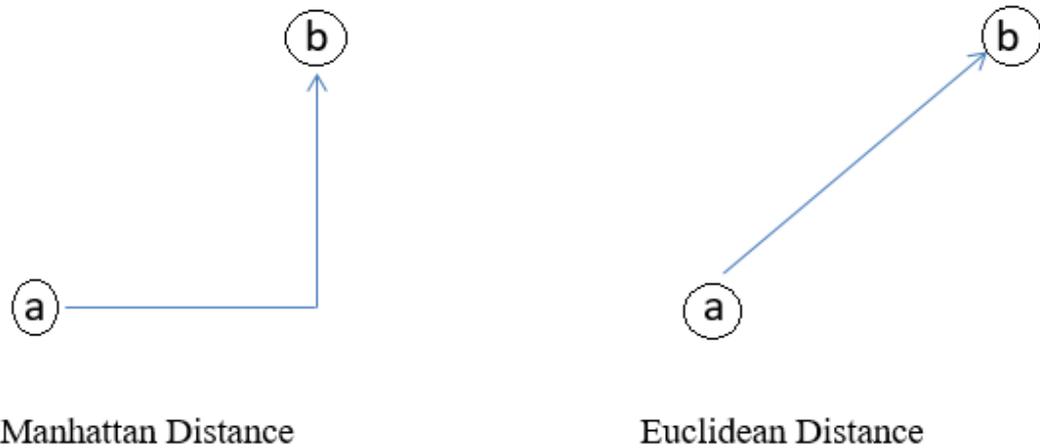
Figure 1 Manhattan and Euclidean metrics.

3. Chebyshev distance: or Maximal distance is the maximum distance between the two instances in any single dimension. Using this metric considers the individual difference between features rather than the whole difference. The formula of Chebyshev is given by[113]:

$$D = max_i \left(|a_i - b_i| \right) \tag{3}$$

4. Mahalanobis measures the number of standard deviations the test point is away from the mean of the class. It considers the variance of each variable and then estimates the covariance matrix of each class. The formula of Mahalanobis between an instance a $(a_1, a_2, ..., a_k)$ and a class of mean m $(m_1, m_2, ..., m_k)$, is given by:

$$D = \sqrt{(a - m)^T S^{-1}(a - m)} \qquad (4)$$

Where *S* is the covariance matrix. By using the covariance matrix, the Mahalanobis considers the correlation between the features. It scales the coordinate's axes by transforming the data into uncorrelated data and then computes the Euclidean distance for the transformed data. In case all the features were uncorrelated and variances are the same in all directions, the covariance matrix is diagonal and Mahalanobis distance turns to be the same as a normalized Euclidean distance [114]. For more details, readers are referred to [115].

Appendix D : WBH1-18 S Horn Antennas Datasheet

Miniature Single Polarisation
1 - 18 GHz Horn Antenna

Model Number WBH1-18S

- World's smallest 1 - 18 GHz antenna: just 96 x 90 x 148 mm
- Gain 1.3 - 13 dBi across the band
- Stable, clean radiation patterns
- Robust, weatherproof design
- Suitable for use as a Reflector Feed or Stand Alone antenna
- Applications include Radio Astronomy, Spectrum Monitoring or Surveillance and EMC Emission tests

Fitting into the palm of a hand, and with an aperture size at just over a quarter wavelength at the lowest frequency, this must be the smallest and lightest wide band 1 GHz horn in the market. Despite this, it punches above its weight offering good gain, VSWR and radiation pattern performance right up to 18 GHz.

The small size has been achieved from a total re-design of the launcher region so that only the aperture now determines the minimum frequency. Its compact nature and low frontal profile will prove useful for minimally invasive situations such as electric field probing.

The antenna is ruggedly constructed from aluminium and engineering plastics. Top, bottom and back pieces are machined out of solid aluminium for low mass and high mechanical stability.

Accessories	Other models available
QTP-A - Economy Antenna Tripod Unit	WBH2-18# - Standard Gain 2 - 18 GHz Horn Antenna
QTP-B - Standard Antenna Tripod Unit	WBHDP0.9-18S 0.9 - 18 GHz Dual Polarised Horn
	WBHDP0.9-18S 0.9 - 18 GHz DP Offset Reflector Antenna

Typical Specification

Frequency	1 - 18 GHz
Gain	1.3 - 13 dBi
Antenna Factor	28.9 - 43.7 dB/m
3dB Beamwidth	280° - 20°
Typical VSWR	2 : 1
Construction	Aluminium / Composite Material
Power Handling	40 Watts, c.w.
Max. Dimensions	96 x 90 x 148 mm long (inclusive of connector)
Connector Type	SMA
Mounting	2 x M5 tapped holes, 38 mm centres
Weight	0.7 kg (1.54 lbs)
Colour	RAL 9016 Traffic White

Typical Gain / Antenna Factor

------- Antenna Factor   ------- Gain
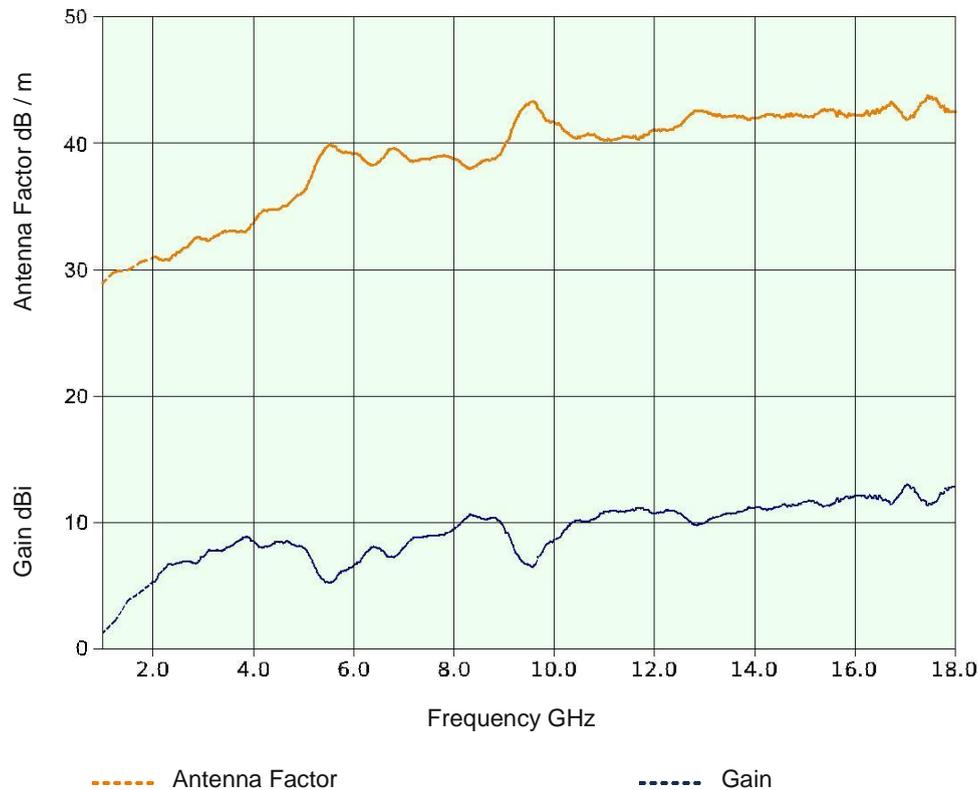
Measured Beamwidths

3 dB Beamwidth
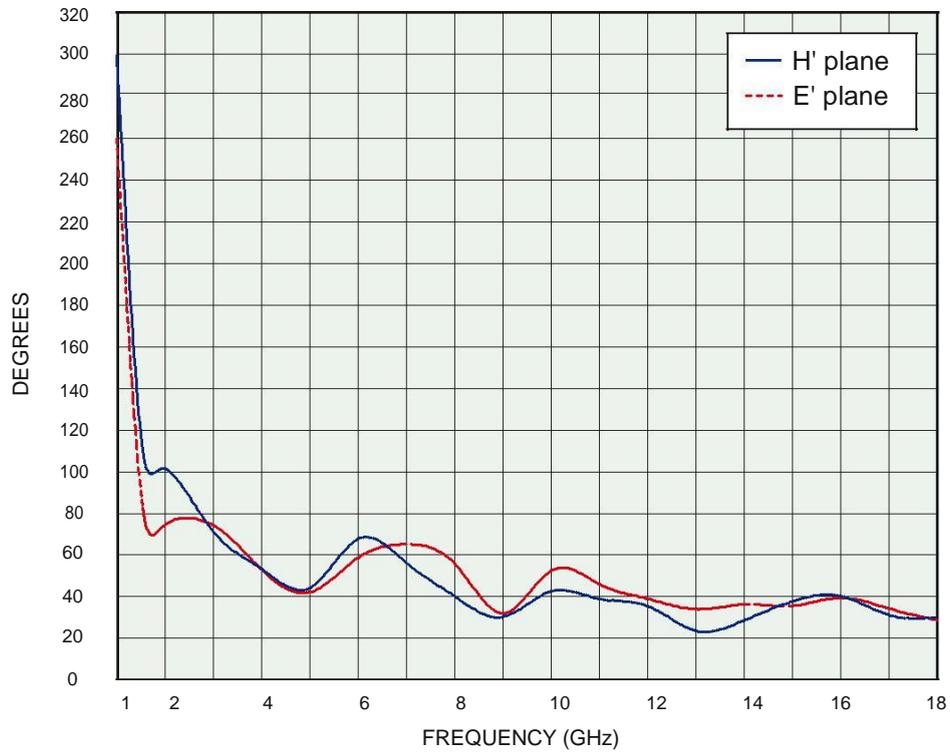
Rear View

The photo above shows the back of the antenna, with weight reduction applied to the mounting plate. The photo also shows the two M5 mounting holes, spaced on 38 mm centres.

Antenna Patterns

Key

H Plane

E Plane
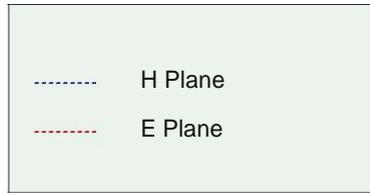
1 GHz
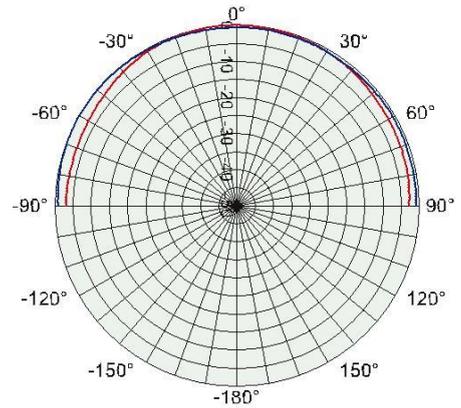
Angle (Degrees)

6 GHz
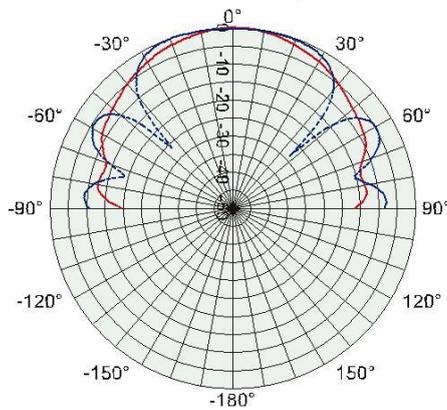
Angle (Degrees)

10 GHz
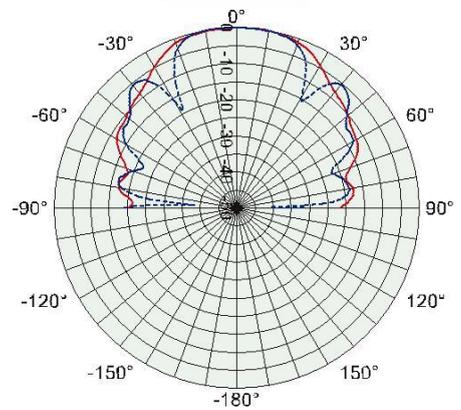
Angle (Degrees)

14 GHz
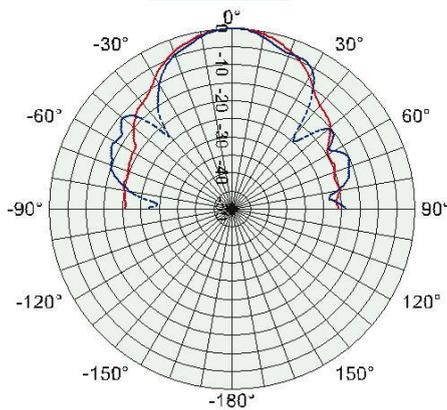
Angle (Degrees)
(Degrees)

18 GHz
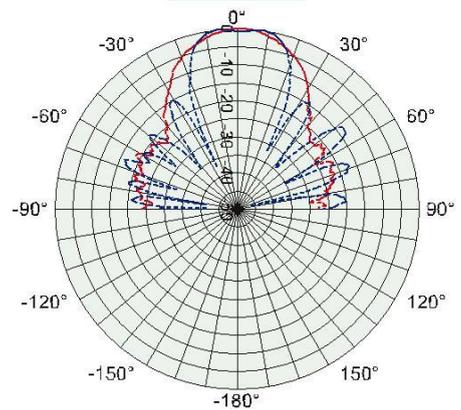
Angle

Designed and Manufactured in England to the Highest Standards

194

Appendix E: Matlab Codes for Acquiring and Visualizing Data

Data_acq_Ac_Radar_15.m

```matlab
%Matlab version (R2014a)
%File name: Data_acq_Ac_Radar_15.m
%Purpose: Data acquiring.


close all;
clear all;

%Enter test number

File_name = input('\n Enter test number \n');

if isempty(File_name)
    FN=1;
    else
    FN=File_name;
end

% Configuring ultrasonic Sensor

%Set ultrasonic data acquisition parameters

S_Rate = 10000;                         %Sampling frequency 10 kHz
Samples = 400;                          %Number of samples per frame 40
nframe = 10;                            %Number of frames

%Set the range of measurement

Range = 340*Samples/(2*S_Rate);        %Maximum ultrasonic range
distance = linspace(0, Range, 400);

% Configuring Radar Sensor

%Create TCP/IP object, address and port number
FF = tcpip('192.168.0.1',5025);
%Set buffer size and timeout
set(FF, 'InputBufferSize', 200000);
set(FF, 'Timeout', 10);
%Open connection
fopen(FF);

%Request VNA ID
fprintf(FF, '*IDN?');
response3 = fscanf(FF);
% Print instrument ID
fprintf(1, 'Connected to: %s\n', response3);

%Specify active directory
fprintf(FF, 'MMEM:CDIR "[SDCARD]:"');
```

```matlab
index_over=1;                        %Tests at different polarisations

while index_over~=0

%Set the ultrasonic sensor
ai = analoginput('mcc',0);
chan = addchannel(ai,[0 1]);
set(ai, 'SampleRate', S_Rate);        %Sampling frequency
set(ai, 'SamplesPerTrigger', Samples); %Number of samples per frame
set(ai, 'TriggerType', 'Software');
set(ai, 'TriggerCondition', 'Falling');
set(ai, 'TriggerConditionValue', 3.0);
set(ai, 'TriggerChannel', chan(1));

%Acquire the ultrasonic data

i = 1;
echo_data_aver = 0;

while (i<nframe+1);                   %Getting Ultrasonic data
start(ai);
Data = getdata(ai);
echo_data = Data(:,2);
echo_data_aver = echo_data_aver + echo_data;
i = i+1;
end;

delete(ai);

%Average the ultrasonic data
echo_data_aver = echo_data_aver./nframe;
echo_data_aver_dB = -(2.324 - echo_data_aver)/0.025;

%Plotting and closing plot
plot(distance, echo_data_aver_dB, 'r');
%plot(distance, echo_data_aver, 'r');
pause(2);
close;

%Write ultrasonic echo file to disk
FilenameA = strcat(num2str(FN),'_A','.dat') ;
fidA = fopen(FilenameA,'w');
fwrite(fidA, echo_data_aver,'float32') ;
fclose(fidA);

%Obtaining Radar Data
%Set radar data acquisition parameters

fcentr = [5.8, 18.2];        %Center Frequencies (only used in file names)
fim = [5.725, 17.9];         %Start Frequencies
fem = [5.875, 18.1];         %Stop Frequencies
fnum = length(fcentr);       %Number of Frequences

i = 1;
while(i<fnum+1);             %Tests at different frequencies

fprintf(FF, 'INST:SEL "NA";*OPC?');          %Reset the instrument
fprintf(FF, 'SENS:SWE:POIN 1601');           %Set point number

%Set start and stop frequencies
```

```matlab
formatSpec1 = '%5.3e';
fi = num2str(fim(i)*1000000000, formatSpec1);   %Start frequency, Hz
fe = num2str(fem(i)*1000000000, formatSpec1);   %Stop frequency, Hz
fprintf(FF, ['SENS:FREQ:START ' fi '']);
fprintf(FF, ['SENS:FREQ:STOP ' fe '']);

fprintf(FF, 'INIT:CONT 0;*OPC?');               %Set to HOLD mode, wait
fprintf(FF, 'INIT:IMM; *OPC?');      %Trigger  one measurement

% Save trace to a s2p. file
FileName = strcat(num2str(FN),'_',num2str(fcentr(i)),'.s2p');
fprintf(FF, ['MMEM:STOR:SNP:DATA "' FileName '"']);

i = i+1;

end

pause (19);
beep;
pause (1);
beep;

index_over=input('\n\n\n Going to make next measurement? - enter number
of next trial, if total finish - enter 0 \n');

FN=index_over;

end;

...........................................................................................
```

Data_Import_NewVNA_5_Graphs

```matlab
%Matlab version (R2014a)
%File name: Data_Import_NewVNA_5_Graphs.m
%Purpose: Data Visualization.


%Disconnect and delete instrument objects

close all
clear all

Filename1 = strcat('1441_5.8.s2p');
Filename2 = strcat('1441_18.2.s2p');

npoints=1601;                          %Number of data points
ggw = 2;                               %Gaussian window parameter,  2-
standard, 0-no window
lengt = 100;                           %Sampling scale from: 1 to: npoints
distmin =0.5;                          %Minimum distance, m
distmax = 5;                           %Maximum distance, m

gw = gausswin(npoints,ggw);            %Gaussian window
```

```matlab
% 1 signal
F_S12 = fopen(Filename1, 'r');
formatSpec = '%f %f %f %f %f %f %f %f %f';
FA_S12 = textscan(F_S12,formatSpec,'HeaderLines',12);
FB_S12 = cell2mat(FA_S12);


freq_S12 = FB_S12(:,1);
%amplitude and phase
amD_S12 = FB_S12(:,4);                  %Power in dB
ph_S12 = FB_S12(:,5);
am_S12 = 10.^(amD_S12./20).*gw;     %Power Linear
%am_S12 = 10.^(amD_S12./20);
re_S12 = am_S12.*cosd(ph_S12);
im_S12 = am_S12.*sind(ph_S12);

% 2 signal
F2_S12 = fopen(Filename2, 'r');
formatSpec = '%f %f %f %f %f %f %f %f %f';
FA2_S12 = textscan(F2_S12,formatSpec,'HeaderLines',12);
FB2_S12 = cell2mat(FA2_S12);


freq2_S12 = FB2_S12(:,1);
%amplitude and phase
amD2_S12 = FB2_S12(:,4);                         %Power in dB
ph2_S12 = FB2_S12(:,5);
am2_S12 = 10.^(amD2_S12./20).*gw;                %Power Linear
%am_S12 = 10.^(amD_S12./20);
re2_S12 = am2_S12.*cosd(ph2_S12);
im2_S12 = am2_S12.*sind(ph2_S12);



%Set up frequency range%
%1 Signal

fmin = min(freq_S12)/10e8;                    %Minimum frequency
fmax = max(freq_S12)/10e8;                    %Maximum frequency
freqn = linspace(fmin, fmax, npoints);        %Frequency range
tau = 1./(fmax-fmin);
tim = linspace(0.0, lengt*tau, npoints);      %Time range
dist = tim*0.3/2;                             %Distance to the surface
distm = max(dist);
dmax = round(npoints*distmax/distm);
dmin = round(npoints*distmin/distm);
d1 = round(npoints/distm);
tmax = dmax*2/0.3;
tmin = dmin*2/0.3;

%2 Signal

fmin2 = min(freq2_S12)/10e8;                     %Minimum frequency
fmax2 = max(freq2_S12)/10e8;                     %Maximum frequency
freqn2 = linspace(fmin2, fmax2, npoints);        %Frequency range
tau2 = 1./(fmax2-fmin2);
tim2 = linspace(0.0, lengt*tau2, npoints);       %Time range
dist2 = tim2*0.3/2;                              %Distance to the
surface
distm2 = max(dist2);
dmax2 = round(npoints*distmax/distm2);
dmin2 = round(npoints*distmin/distm2);
d1_2 = round(npoints/distm2);
```

198

```matlab
tmax2 = dmax2*2/0.3;
tmin2 = dmin2*2/0.3;


%Path loss compenstaion
nn = 1;
Loss = zeros(1, npoints);
while nn < npoints + 1;
%Loss(nn) = 1;
Loss(nn) = -5 + 40*log10(nn/512);
nn = nn +1;
end;

%restore 1 signal in time domain%
l = 1;
t_S12 = 0;
while(l<npoints);
t_S12 = t_S12 + am_S12(l)*cos(2*pi*freqn(l)*tim + ph_S12(l)*pi/180);
l = l+1;
end;

%Calculating the signal envelope and power%

th_S12 = hilbert(t_S12./npoints);
ty_S12=20*log10(abs(th_S12));
%ty_S12 = abs(th_S12);
Atm_S12 = mean(ty_S12(dmin:dmax));
tyy_S12 = smooth((ty_S12),1);
L_S12 = smooth((Loss),1);

%restore 2 signal in time domain%

l = 1;
t2_S12 = 0;
while(l<npoints);
t2_S12 = t2_S12 + am2_S12(l)*cos(2*pi*freqn2(l)*tim2 +
ph2_S12(l)*pi/180);
l = l+1;
end;

%Calculating the signal envelope and power%
th2_S12 = hilbert(t2_S12./npoints);
ty2_S12=20*log10(abs(th2_S12));
%ty_S12 = abs(th_S12);
Atm2_S12 = mean(ty2_S12(dmin:dmax));
tyy2_S12 = smooth((ty2_S12),1);


tyy_S12 = tyy_S12;
tyy2_S12 =tyy2_S12;

figure;
plot(dist, tyy_S12);
hold on;
%plot(dist(dmin:dmax), ty_S12(dmin:dmax));
plot(dist2, tyy2_S12, 'k');
xlabel('Distance, m');
ylabel('Power, DBm');
xlim([0 5]);
%ylim([-60 0]);
title('Sand 03 March 2014 18GHz (3GHz VV)');
```

```matlab
legend('file1', 'file2');
```

....

Data_Import_Radar_200

```matlab
%Matlab version (R2014a)
%File name: Data_Import_Radar_200.m
%Purpose: Import Radar Data
%Disconnect and delete instrument objects

close all
clear all

File_name=input('\n \n Enter file name \n')

if isempty(File_name)
    FN1=1;
    else
    FN1=File_name;

end

fcentr = [5.8, 18.2];                   %Center Frequencies

fnum = length(fcentr);                  %Number of Frequences
npoints=1601;                           %Number of data points
ggw = 2;               %Gaussian window parameter,  2-standard, 0-no window
lengt = 100;                            %Sampling scale from: 1 to: npoints
distmin = 1.5;                          %Minimum distance, m
distmax = 4.0;                          %Maximum distance, m

Ampl = zeros(49, 9);                    %Matrix of results
gw = gausswin(npoints,ggw);             %Gaussian window

ij = 1;

while ij < 25;
    FN = FN1 + ij - 1;
k = 1;
while (k<4);

    i = 1;
while(i<fnum+1);
    Filename1 =
strcat(num2str(FN),num2str(k),'_',num2str(fcentr(i)),'.s2p');

%import data%
%import backscattered signal 1%
F_S12 = fopen(Filename1, 'r');
formatSpec = '%f %f %f %f %f %f %f %f %f';
FA_S12 = textscan(F_S12,formatSpec,'HeaderLines',12);
FB_S12 = cell2mat(FA_S12);


freq_S12 = FB_S12(:,1);
```

```matlab
%amplitude and phase
amD_S12 = FB_S12(:,4);                          %Power in dB
ph_S12 = FB_S12(:,5);
am_S12 = 10.^(amD_S12./20).*gw;                 %Power Linear
%am_S12 = 10.^(amD_S12./20);
%re_S12 = am_S12.*cosd(ph_S12);
%im_S12 = am_S12.*sind(ph_S12);


%Set up frequency range%

fmin = min(freq_S12)/10e8;                      %Minimum frequency
fmax = max(freq_S12)/10e8;                      %Maximum frequency
freqn = linspace(fmin, fmax, npoints);          %Frequency range
tau = 1./(fmax-fmin);
tim = linspace(0.0, lengt*tau, npoints);        %Time range
dist = tim*0.3/2;                               %Distance to the
surface
distm = max(dist);
dmax = round(npoints*distmax/distm);
dmin = round(npoints*distmin/distm);
%tmax = dmax*2/0.3;
%tmin = dmin*2/0.3;


%restore signal in time domain%
l = 1;
t_S12 = 0;

while(l<npoints);
t_S12 = t_S12 + am_S12(l)*cos(2*pi*freqn(l)*tim + ph_S12(l)*pi/180);
l = l+1;
end;


%Calculating the signal envelope and power%%
%th_S12 = hilbert(t_S12./npoints);

th_S12 = hilbert(t_S12);
%ty_S12=20*log10(abs(th_S12));
ty_S12 = abs(th_S12);
Atm_S12 = mean(ty_S12(dmin:dmax));

%tyy_S12 = smooth(ty_S12, 5);
%plot(dist(dmin:dmax), ty_S12(dmin:dmax));

Ampl(i+2*ij-1, k+1) = Atm_S12;
Ampl(i+2*ij-1, 1) = fcentr(i);
%Ampl(i+2*ij, 9) = FN;

if (k == 1)
    Ampl(i+2*ij-1, 8)= std(ty_S12(dmin:dmax))/Atm_S12;
    Ampl(i+2*ij-1, 9) = FN;
end

i = i+1;
end

k= k+1;
end

ij = ij +1;
end;
```

```matlab
Ampl(:,6) = 0;
Ampl(:,7) = 0;

%Writting average radar signal amplitude to disk

FilenameA = strcat(num2str(FN1),'_','D21.csv') ;
    csvwrite(FilenameA, Ampl);

% Read input file data content (without header).

fid_in = fopen(FilenameA, 'r') ;            %Open input file for reading.
fgetl(fid_in) ;                             % Skip header line in input
file.
content = fread(fid_in) ;                   % Read rest of input file.
fclose(fid_in) ;                            % Close input file.

% Write output file: new header + previous data content.

fid_out = fopen(FilenameA, 'w') ;           %Open input file for writing.
fprintf(fid_out, 'FREQ_BW,VV,VH,HH,HV,DT, PT,STD_HH,FN\n') ;    % Output
new header.
fwrite(fid_out, content) ;                  % Output previous data content.
fclose(fid_out) ;                           % Close output file.
fclose all;
```

....

Data_Import_Acoustic_200

```matlab
%Matlab version (R2014a)
%File name: Data_Import_Acoustic_200.m
%Purpose: Import Ultrasonic Data

close all;
clear all;

Filenumber = 145;
Filename1 = Filenumber;
Filename2 = Filename1 + 1;
Filename3 = Filename1 + 2;
Filename4 = Filename1 + 3;

Filename11 = strcat(num2str(Filename1),'1_A.dat');
Filename12 = strcat(num2str(Filename2),'1_A.dat');
Filename13 = strcat(num2str(Filename3),'1_A.dat');
Filename14 = strcat(num2str(Filename4),'1_A.dat');

A = zeros(8, 6);                            %ultrasonic results

FN = Filename1;

%Set ultrasonic data acquisition parameters
```

202

```matlab
S_Rate = 10000;                         %Sampling frequency 10 kHz
Samples = 400;                          %Number of samples per frame 40
nframe = 10;                            %Number of frames


%Set the range of measurement
Range = 340*Samples/(2*S_Rate);         %Maximum ultrasonic range
distance = linspace(0, Range, 400);
dist_min = 1.5;                         %Minimum distance
dist_max = 4.0;                         %Maximum distance
sample_min = round(dist_min*Samples/Range);
sample_max = round(dist_max*Samples/Range);


%Import echo signal%

F_11 = fopen(Filename11,'r');
FA_1 = fread(F_11,'float32');
F_12 = fopen(Filename12,'r');
FA_2 = fread(F_12,'float32');
F_13 = fopen(Filename13,'r');
FA_3 = fread(F_13,'float32');
F_14 = fopen(Filename14,'r');
FA_4 = fread(F_14,'float32');


%Amplitude in dB
%0dB = 2.324V, 0.25mV / dB


FA_1DB = -(2.324 - FA_1)/0.025;
FA_2DB = -(2.324 - FA_2)/0.025;
FA_3DB = -(2.324 - FA_3)/0.025;
FA_4DB = -(2.324 - FA_4)/0.025;


FA_1DM = mean(FA_1DB(sample_min : sample_max, :));
FA_2DM = mean(FA_2DB(sample_min : sample_max, :));
FA_3DM = mean(FA_3DB(sample_min : sample_max, :));
FA_4DM = mean(FA_4DB(sample_min : sample_max, :));


FA_1AM = mean(FA_1(sample_min : sample_max, :));
FA_2AM = mean(FA_2(sample_min : sample_max, :));
FA_3AM = mean(FA_3(sample_min : sample_max, :));
FA_4AM = mean(FA_4(sample_min : sample_max, :));


FA_1ST = std(FA_1(sample_min : sample_max, :))/FA_1AM;
FA_2ST = std(FA_2(sample_min : sample_max, :))/FA_2AM;
FA_3ST = std(FA_3(sample_min : sample_max, :))/FA_3AM;
FA_4ST = std(FA_4(sample_min : sample_max, :))/FA_4AM;


%Path loss compenstaion

nn = 1;
Loss = zeros(Samples, 1);
while nn < Samples + 1;
Loss(nn) = -30 + 40*log10((nn)/10);
nn = nn +1;
end;


FA_1DBPL = FA_1DB + Loss;
FA_1APL = db2mag(FA_1DBPL);
FA_2DBPL = FA_2DB + Loss;
FA_2APL = db2mag(FA_2DBPL);
FA_3DBPL = FA_3DB + Loss;
```

```matlab
FA_3APL = db2mag(FA_3DBPL);
FA_4DBPL = FA_4DB + Loss;
FA_4APL = db2mag(FA_4DBPL);

plot (distance, FA_1DBPL, distance, FA_1DB);

%Signal energy and power above the threshold

PA1 = 0;
DA1 = 0;
PA2 = 0;
DA2 = 0;
PA3 = 0;
DA3 = 0;
PA4 = 0;
DA4 = 0;
Thresh = 0.15;
for ni = sample_min : sample_max;
    if (FA_1APL(ni) > Thresh)
        DA1 = DA1 + 1;
        PA1 = PA1 + FA_1APL(ni);
    end;
    if (FA_2APL(ni) > Thresh)
        DA2 = DA2 + 1;
        PA2 = PA2 + FA_2APL(ni);
    end;
    if (FA_1APL(ni) > Thresh)
        DA3 = DA3 + 1;
        PA3 = PA3 + FA_3APL(ni);
    end;
    if (FA_4APL(ni) > Thresh)
        DA4 = DA4 + 1;
        PA4 = PA4 + FA_4APL(ni);
    end;
end;

DA1m = DA1*Range/Samples;
PA1a = PA1/(sample_max - sample_min);
DA2m = DA2*Range/Samples;
PA2a = PA2/(sample_max - sample_min);
DA3m = DA3*Range/Samples;
PA3a = PA3/(sample_max - sample_min);
DA4m = DA4*Range/Samples;
PA4a = PA4/(sample_max - sample_min);

%Data to write

A(1,1) = FA_1AM;
A(1,2) = FA_1DM;
A(1,3) = DA1m;
A(1,4) = PA1a;
A(1,5) = FA_1ST;
A(1,6) = Filename1;
A(2,:) = A(1,:);
A(3,1) = FA_2AM;
A(3,2) = FA_2DM;
A(3,5) = FA_2ST;
A(3,3) = DA2m;
A(3,4) = PA2a;
A(3,6) = Filename2;
A(4,:) = A(3,:);
A(5,1) = FA_3AM;
```

204

```matlab
A(5,2) = FA_3DM;
A(5,5) = FA_3ST;
A(5,3) = DA3m;
A(5,4) = PA3a;
A(5,6) = Filename3;
A(6,:) = A(5,:);
A(7,1) = FA_4AM;
A(7,2) = FA_4DM;
A(7,5) = FA_4ST;
A(7,3) = DA4m;
A(7,4) = PA4a;
A(7,6) = Filename4;
A(8,:) = A(7,:);

%Write ultrasonic echo file to disk
FilenameB = strcat(num2str(FN),'_D.csv') ;
csvwrite(FilenameB, A);

%Write header of ultrasonic echo file
fid_in = fopen(FilenameB, 'r');
content=fread(fid_in);
fclose(fid_in);
fid_out = fopen(FilenameB, 'w');
fprintf(fid_out, 'Pow_mV,Pow_dB,Dur_Thr_m,Pow_Thr_mV,St_dev,File_No\n');
fwrite(fid_out, content);
fclose(fid_out);
```

Appendix F: Matlab Codes for Classification Algorithms

MDC_Classifier

```matlab
%Matlab version (R2014a)
%File name: MDC_Classifier.m
%Purpose: MDC Classification Algorithm.


close all;
clear all;
clc;

knn = 1;

%Selecting the surfaces for comparison

Sname1 = 'Asph';
Sname2 = 'Asph2';
%Sname2 = 'Asph';
Sname3 = 'Grass';
%Sname3 = 'Grav';
Sname4 = 'Grav';
Sname5 = 'Sand';
%Sname5 = 'Grav';



%Importing Excel file with all data%



filename = 'Database_202Neural.xls';
sheet = 1;
[Data, Text, Alldata] = xlsread(filename);

%Selecting necessary data for the scatters%

Data_S = Data (2:517, [2,8,9,10,27,28,29,30,12, 17,18,19,20,21,22,23]);
%2 - 669

%Plotting the scatters in 3D

Scatter1 = 'P';
Scatter2 = 'DT';
Scatter3 = 'Std';
Scatter4 = 'VV18';
Scatter5 = 'HH18';
Scatter6 = 'VH18';

%P, DT, PT, Std, VV5, VH5, Std5, VVHH5, VHHH5  /18

%1 - Record number
%2 - Surface: Gravel 1; Asphalt 3; Grass 5
%3 - Covered: No info 0; Dry 1; Wet 2; Snow 3; Ice 4; Frost 5
%4 - Details: No info 0; Soft 1; Rammed 2; Wet 3;
%5 - Ac Power
```

```matlab
%6 - Ac Dur Thresh
%7 - Ac Power Thresh
%8 - Ac STD
%9 - Frequency
%10 - 17 VV
%11 - 18 VH
%12 - 19 HH
%13 - 20 HV
%14 - 21 DT
%15 - 22 PT
%16 - 23 Std


idx_5 = (Data_S(:,9) == 5.8);
Data_5 = Data_S(idx_5,:);
idx_18 = (Data_S(:,9) == 18);
Data_18 = Data_S(idx_18,:);




%Sorting by surfaces ACOUSTIC AND 5.8GHz


idx_Asph = (Data_5(:,2)==3); %Asphalt
Data_Asph5 = Data_5(idx_Asph,:);
idx_Grass = (Data_5(:,2)==5); %Grass
Data_Grass5 = Data_5(idx_Grass,:);
idx_Grav = (Data_5(:,2)==1); %Gravel
Data_Grav5 = Data_5(idx_Grav,:);
idx_Asph2 = (Data_5(:,2)==8); %Asph Smooth
Data_Asph25 = Data_5(idx_Asph2,:);
%idx_Sand = (Data_A(:,2)==7); %Sand(ALL))
idx_Sand = (Data_5(:,2)==7 & Data_5(:,4)==1); %Sand(Smooth only)
Data_Sand5 = Data_5(idx_Sand,:);


%Sorting by surfaces 18GHz


idx_Asph = (Data_18(:,2)==3); %Asphalt
Data_Asph18 = Data_18(idx_Asph,:);
idx_Grass = (Data_18(:,2)==5); %Grass
Data_Grass18 = Data_18(idx_Grass,:);
idx_Grav = (Data_18(:,2)==1); %Gravel
Data_Grav18 = Data_18(idx_Grav,:);
idx_Asph2 = (Data_18(:,2)==8); %Asph Smooth
Data_Asph218 = Data_18(idx_Asph2,:);
%idx_Sand = (Data_18(:,2)==7); %Sand(ALL))
idx_Sand = (Data_18(:,2)==7 & Data_18(:,4)==1); %Sand(Smooth only)
Data_Sand18 = Data_18(idx_Sand,:);




%Asphalt


Asph_P = Data_Asph5(:,5);
Asph_DT = Data_Asph5(:,6);
Asph_PT = Data_Asph5(:,7);
Asph_Std = Data_Asph5(:,8);
Asph_VV5 = Data_Asph5(:,10);
Asph_VH5 = Data_Asph5(:,11);
Asph_HH5 = Data_Asph5(:,12);
Asph_VVHH5 = Data_Asph5(:,10)./Data_Asph5(:,12);
Asph_VHHH5 = Data_Asph5(:,11)./Data_Asph5(:,12);
Asph_DT5 = Data_Asph5(:,14);
Asph_PT5 = Data_Asph5(:,15);
Asph_Std5 = Data_Asph5(:,16);
```

```matlab
Asph_VV18 = Data_Asph18(:,10);
Asph_VH18 = Data_Asph18(:,11);
Asph_HH18 = Data_Asph18(:,12);
Asph_VVHH18 = Data_Asph18(:,10)./Data_Asph18(:,12);
Asph_VHHH18 = Data_Asph18(:,11)./Data_Asph18(:,12);
Asph_DT18 = Data_Asph18(:,14);
Asph_PT18 = Data_Asph18(:,15);
Asph_Std18 = Data_Asph18(:,16);

%Grass

Grass_P = Data_Grass5(:,5);
Grass_DT = Data_Grass5(:,6);
Grass_PT = Data_Grass5(:,7);
Grass_Std = Data_Grass5(:,8);
Grass_VV5 = Data_Grass5(:,10);
Grass_VH5 = Data_Grass5(:,11);
Grass_HH5 = Data_Grass5(:,12);
Grass_VVHH5 = Data_Grass5(:,10)./Data_Grass5(:,12);
Grass_VHHH5 = Data_Grass5(:,11)./Data_Grass5(:,12);
Grass_DT5 = Data_Grass5(:,14);
Grass_PT5 = Data_Grass5(:,15);
Grass_Std5 = Data_Grass5(:,16);
Grass_VV18 = Data_Grass18(:,10);
Grass_VH18 = Data_Grass18(:,11);
Grass_HH18 = Data_Grass18(:,12);
Grass_VVHH18 = Data_Grass18(:,10)./Data_Grass18(:,12);
Grass_VHHH18 = Data_Grass18(:,11)./Data_Grass18(:,12);
Grass_DT18 = Data_Grass18(:,14);
Grass_PT18 = Data_Grass18(:,15);
Grass_Std18 = Data_Grass18(:,16);

%Gravel

Grav_P = Data_Grav5(:,5);
Grav_DT = Data_Grav5(:,6);
Grav_PT = Data_Grav5(:,7);
Grav_Std = Data_Grav5(:,8);
Grav_VV5 = Data_Grav5(:,10);
Grav_VH5 = Data_Grav5(:,11);
Grav_HH5 = Data_Grav5(:,12);
Grav_VVHH5 = Data_Grav5(:,10)./Data_Grav5(:,12);
Grav_VHHH5 = Data_Grav5(:,11)./Data_Grav5(:,12);
Grav_DT5 = Data_Grav5(:,14);
Grav_PT5 = Data_Grav5(:,15);
Grav_Std5 = Data_Grav5(:,16);
Grav_VV18 = Data_Grav18(:,10);
Grav_VH18 = Data_Grav18(:,11);
Grav_HH18 = Data_Grav18(:,12);
Grav_VVHH18 = Data_Grav18(:,10)./Data_Grav18(:,12);
Grav_VHHH18 = Data_Grav18(:,11)./Data_Grav18(:,12);
Grav_DT18 = Data_Grav18(:,14);
Grav_PT18 = Data_Grav18(:,15);
Grav_Std18 = Data_Grav18(:,16);

%Sand

Sand_P = Data_Sand5(:,5);
Sand_DT = Data_Sand5(:,6);
Sand_PT = Data_Sand5(:,7);
Sand_Std = Data_Sand5(:,8);
```

```matlab
Sand_VV5 = Data_Sand5(:,10);
Sand_VH5 = Data_Sand5(:,11);
Sand_HH5 = Data_Sand5(:,12);
Sand_VVHH5 = Data_Sand5(:,10)./Data_Sand5(:,12);
Sand_VHHH5 = Data_Sand5(:,11)./Data_Sand5(:,12);
Sand_DT5 = Data_Sand5(:,14);
Sand_PT5 = Data_Sand5(:,15);
Sand_Std5 = Data_Sand5(:,16);
Sand_VV18 = Data_Sand18(:,10);
Sand_VH18 = Data_Sand18(:,11);
Sand_HH18 = Data_Sand18(:,12);
Sand_VVHH18 = Data_Sand18(:,10)./Data_Sand18(:,12);
Sand_VHHH18 = Data_Sand18(:,11)./Data_Sand18(:,12);
Sand_DT18 = Data_Sand18(:,14);
Sand_PT18 = Data_Sand18(:,15);
Sand_Std18 = Data_Sand18(:,16);

%Asphalt Smooth

Asph2_P = Data_Asph25(:,5);
Asph2_DT = Data_Asph25(:,6);
Asph2_PT = Data_Asph25(:,7);
Asph2_Std = Data_Asph25(:,8);
Asph2_VV5 = Data_Asph25(:,10);
Asph2_VH5 = Data_Asph25(:,11);
Asph2_HH5 = Data_Asph25(:,12);
Asph2_VVHH5 = Data_Asph25(:,10)./Data_Asph25(:,12);
Asph2_VHHH5 = Data_Asph25(:,11)./Data_Asph25(:,12);
Asph2_DT5 = Data_Asph25(:,14);
Asph2_PT5 = Data_Asph25(:,15);
Asph2_Std5 = Data_Asph25(:,16);
Asph2_VV18 = Data_Asph218(:,10);
Asph2_VH18 = Data_Asph218(:,11);
Asph2_HH18 = Data_Asph218(:,12);
Asph2_VVHH18 = Data_Asph218(:,10)./Data_Asph218(:,12);
Asph2_VHHH18 = Data_Asph218(:,11)./Data_Asph218(:,12);
Asph2_DT18 = Data_Asph218(:,14);
Asph2_PT18 = Data_Asph218(:,15);
Asph2_Std18 = Data_Asph218(:,16);


G1 = eval(['Grass_' num2str(Scatter1)]);
G2 = eval(['Grass_' num2str(Scatter2)]);
G3 = eval(['Grass_' num2str(Scatter3)]);
G4 = eval(['Grass_' num2str(Scatter4)]);
G5 = eval(['Grass_' num2str(Scatter5)]);
G6 = eval(['Grass_' num2str(Scatter6)]);
G1p = mle(G1);
G2p = mle(G2);
G3p = mle(G3);
G4p = mle(G4);
G5p = mle(G5);
G6p = mle(G6);

%Build the matrix%

S_Grass = [G1, G2, G3, G4, G5, G6];
Sm_Grass = [G1p(1), G2p(1), G3p(1), G4p(1), G5p(1), G6p(1)];
Sp_Grass = [G1p(2), G2p(2), G3p(2), G4p(2), G5p(2), G6p(2)];
%S_Grass = [G1, G2, G3];
```

```matlab
V1 = eval(['Grav_' num2str(Scatter1)]);
V2 = eval(['Grav_' num2str(Scatter2)]);
V3 = eval(['Grav_' num2str(Scatter3)]);
V4 = eval(['Grav_' num2str(Scatter4)]);
V5 = eval(['Grav_' num2str(Scatter5)]);
V6 = eval(['Grav_' num2str(Scatter6)]);
V1p = mle(V1);
V2p = mle(V2);
V3p = mle(V3);
V4p = mle(V4);
V5p = mle(V5);
V6p = mle(V6);


%Build the matrix%
S_Grav = [V1, V2, V3, V4, V5, V6];
Sm_Grav = [V1p(1), V2p(1), V3p(1), V4p(1), V5p(1), V6p(1)];
Sp_Grav = [V1p(2), V2p(2), V3p(2), V4p(2), V5p(2), V6p(2)];
%S_Grav = [V1, V2, V3];


A1 = eval(['Asph_' num2str(Scatter1)]);
A2 = eval(['Asph_' num2str(Scatter2)]);
A3 = eval(['Asph_' num2str(Scatter3)]);
A4 = eval(['Asph_' num2str(Scatter4)]);
A5 = eval(['Asph_' num2str(Scatter5)]);
A6 = eval(['Asph_' num2str(Scatter6)]);
A1p = mle(A1);
A2p = mle(A2);
A3p = mle(A3);
A4p = mle(A4);
A5p = mle(A5);
A6p = mle(A6);


%Build the matrix%
S_Asph = [A1, A2, A3, A4, A5, A6];
Sm_Asph = [A1p(1), A2p(1), A3p(1), A4p(1), A5p(1), A6p(1)];
Sp_Asph = [A1p(2), A2p(2), A3p(2), A4p(2), A5p(2), A6p(2)];
%S_Asph = [A1, A2, A3];


S1 = eval(['Sand_' num2str(Scatter1)]);
S2 = eval(['Sand_' num2str(Scatter2)]);
S3 = eval(['Sand_' num2str(Scatter3)]);
S4 = eval(['Sand_' num2str(Scatter4)]);
S5 = eval(['Sand_' num2str(Scatter5)]);
S6 = eval(['Sand_' num2str(Scatter6)]);
S1p = mle(S1);
S2p = mle(S2);
S3p = mle(S3);
S4p = mle(S4);
S5p = mle(S5);
S6p = mle(S6);


%Build the matrix%
S_Sand = [S1, S2, S3, S4, S5, S6];
Sm_Sand = [S1p(1), S2p(1), S3p(1), S4p(1), S5p(1), S6p(1)];
Sp_Sand = [S1p(2), S2p(2), S3p(2), S4p(2), S5p(2), S6p(2)];
%S_Sand = [S1, S2, S3];


U1 = eval(['Asph2_' num2str(Scatter1)]);
U2 = eval(['Asph2_' num2str(Scatter2)]);
U3 = eval(['Asph2_' num2str(Scatter3)]);
```

210

```matlab
U4 = eval(['Asph2_' num2str(Scatter4)]);
U5 = eval(['Asph2_' num2str(Scatter5)]);
U6 = eval(['Asph2_' num2str(Scatter6)]);
U1p = mle(U1);
U2p = mle(U2);
U3p = mle(U3);
U4p = mle(U4);
U5p = mle(U5);
U6p = mle(U6);

%Build the matrix%
S_Asph2 = [U1, U2, U3, U4, U5, U6];
Sm_Asph2 = [U1p(1), U2p(1), U3p(1), U4p(1), U5p(1), U6p(1)];
Sp_Asph2 = [U1p(2), U2p(2), U3p(2), U4p(2), U5p(2), U6p(2)];
%S_Asph2 = [U1, U2, U3];

S1 = eval(['S_' num2str(Sname1)]);
S2 = eval(['S_' num2str(Sname2)]);
S3 = eval(['S_' num2str(Sname3)]);
S4 = eval(['S_' num2str(Sname4)]);
S5 = eval(['S_' num2str(Sname5)]);
S1m = eval(['Sm_' num2str(Sname1)]);
S2m = eval(['Sm_' num2str(Sname2)]);
S3m = eval(['Sm_' num2str(Sname3)]);
S4m = eval(['Sm_' num2str(Sname4)]);
S5m = eval(['Sm_' num2str(Sname5)]);
S1p = eval(['Sp_' num2str(Sname1)]);
S2p = eval(['Sp_' num2str(Sname2)]);
S3p = eval(['Sp_' num2str(Sname3)]);
S4p = eval(['Sp_' num2str(Sname4)]);
S5p = eval(['Sp_' num2str(Sname5)]);

%Length of arrays
LS(1) = length(S1);
LS(2) = length(S2);
LS(3) = length(S3);
LS(4) = length(S4);
LS(5) = length(S5);

%Applying MDC method (Based on Kmeans algorithm)

%Calculating of confusion matrix

A = zeros(5, 5);                      %Confusion matrix
ERR = zeros(5, 5);

kj = 1;
while kj < 6;
    knnkj = zeros(1,5);
    knnkj = knnkj + knn;
    %knnkj(kj) = knn + 1;
    i = 1;

    DSKJ = zeros(5, 1);

    while i < LS(kj) + 1;
        SKJ = eval(['S' num2str(kj)]);
        newpoint = SKJ(i,:);
        %
        jp = 1;
        D1p = 0;
```

```matlab
        D2p = 0;
        D3p = 0;
        D4p = 0;
        D5p = 0;
        jp_length = length(newpoint);
        while jp < jp_length + 1
            %D1p = normpdf(X, mu, sigma);
            %D1p = D1p + log10(normpdf(newpoint(jp), S1m(jp), S1p(jp)));
            %D2p = D2p + log10(normpdf(newpoint(jp), S2m(jp), S2p(jp)));
            %D3p = D3p + log10(normpdf(newpoint(jp), S3m(jp), S3p(jp)));
            %D4p = D4p + log10(normpdf(newpoint(jp), S4m(jp), S4p(jp)));
            %D5p = D5p + log10(normpdf(newpoint(jp), S5m(jp), S5p(jp)));
            %K-means method
            D1p = D1p + abs(newpoint(jp)- S1m(jp));
            D2p = D2p + abs(newpoint(jp)- S2m(jp));
            D3p = D3p + abs(newpoint(jp)- S3m(jp));
            D4p = D4p + abs(newpoint(jp)- S4m(jp));
            D5p = D5p + abs(newpoint(jp)- S5m(jp));

                        jp = jp +1;
        end;

        DSKJ(1) = abs(D1p);
        DSKJ(2) = abs(D2p);
        DSKJ(3) = abs(D3p);
        DSKJ(4) = abs(D4p);
        DSKJ(5) = abs(D5p);

        DSKM = min(DSKJ);

        index = find(DSKJ == min(DSKJ(:)));
        ERR(kj, index) = ERR(kj, index) + 1;


    A(kj, 1) = ERR(kj, 1)/LS(kj);
    A(kj, 2) = ERR(kj, 2)/LS(kj);
    A(kj, 3) = ERR(kj, 3)/LS(kj);
    A(kj, 4) = ERR(kj, 4)/LS(kj);
    A(kj, 5) = ERR(kj, 5)/LS(kj);
    %A(kj, kj) = ERR(kj, kj)/LS(kj);
        i = i +1;
    end;

    kj = kj + 1;

end;



%Write confusion matrix to disk

FilenameB=strcat(num2str(Scatter1),'_',num2str(Scatter2),'_',num2str(Scatter3),'_MLC', '.csv') ;

%FilenameB = strcat('Confusion_Matrix1.csv') ;
csvwrite(FilenameB, A);

%Write header of ultrasonic echo file

fid_in = fopen(FilenameB, 'r');
content=fread(fid_in);
```

212

```
fclose(fid_in);
fid_out = fopen(FilenameB, 'w');
fprintf(fid_out, 'Asph,Asph_Sm,Grass,Gravel,Sand\n');
fwrite(fid_out, content);
fclose(fid_out);
```

....

MLE_Classifier

```
%Matlab version (R2014a)
%File name: MLE_Classifier.m
%Purpose: MLE Classification Algorithm.


close all;
clear all;
clc;

knn = 1;

%Selecting the surfaces for comparison
Sname1 = 'Asph';
Sname2 = 'Asph2';
%Sname2 = 'Asph';
Sname3 = 'Grass';
%Sname3 = 'Grav';
Sname4 = 'Grav';
Sname5 = 'Sand';
%Sname5 = 'Grav';



%Importing Excel file with all data%

%filename = 'Database_2013_3.xls';
filename = 'Database_202Neural.xls';
sheet = 1;
[Data, Text, Alldata] = xlsread(filename);

%Selecting necessary data for the scatters%

Data_S = Data (2:517, [2,8,9,10,27,28,29,30,12, 17,18,19,20,21,22,23]);
%2 - 669

%Plotting the scatters in 3D
Scatter1 = 'VV18';
Scatter2 = 'Std';
Scatter3 = 'HH18';
Scatter4 = 'VH18';
Scatter5 = 'P';
Scatter6 = 'DT';
%P, DT, PT, Std, VV5, VH5, Std5, VVHH5, VHHH5  /18
```

```matlab
%1 - Record number
%2 - Surface: Gravel 1; Asphalt 3; Grass 5
%3 - Covered: No info 0; Dry 1; Wet 2; Snow 3; Ice 4; Frost 5
%4 - Details: No info 0; Soft 1; Rammed 2; Wet 3;
%5 - Ac Power
%6 - Ac Dur Thresh
%7 - Ac Power Thresh
%8 - Ac STD
%9 - Frequency
%10 - 17 VV
%11 - 18 VH
%12 - 19 HH
%13 - 20 HV
%14 - 21 DT
%15 - 22 PT
%16 - 23 Std


idx_5 = (Data_S(:,9) == 5.8);
Data_5 = Data_S(idx_5,:);
idx_18 = (Data_S(:,9) == 18);
Data_18 = Data_S(idx_18,:);



%Sorting by surfaces ACOUSTIC AND 5.8GHz

idx_Asph = (Data_5(:,2)==3); %Asphalt
Data_Asph5 = Data_5(idx_Asph,:);
idx_Grass = (Data_5(:,2)==5); %Grass
Data_Grass5 = Data_5(idx_Grass,:);
idx_Grav = (Data_5(:,2)==1); %Gravel
Data_Grav5 = Data_5(idx_Grav,:);
idx_Asph2 = (Data_5(:,2)==8); %Asph Smooth
Data_Asph25 = Data_5(idx_Asph2,:);
%idx_Sand = (Data_A(:,2)==7); %Sand(ALL))
idx_Sand = (Data_5(:,2)==7 & Data_5(:,4)==1); %Sand(Smooth only)
Data_Sand5 = Data_5(idx_Sand,:);


%Sorting by surfaces 18GHz

idx_Asph = (Data_18(:,2)==3); %Asphalt
Data_Asph18 = Data_18(idx_Asph,:);
idx_Grass = (Data_18(:,2)==5); %Grass
Data_Grass18 = Data_18(idx_Grass,:);
idx_Grav = (Data_18(:,2)==1); %Gravel
Data_Grav18 = Data_18(idx_Grav,:);
idx_Asph2 = (Data_18(:,2)==8); %Asph Smooth
Data_Asph218 = Data_18(idx_Asph2,:);
%idx_Sand = (Data_18(:,2)==7); %Sand(ALL))
idx_Sand = (Data_18(:,2)==7 & Data_18(:,4)==1); %Sand(Smooth only)
Data_Sand18 = Data_18(idx_Sand,:);



%Asphalt

Asph_P = Data_Asph5(:,5);
Asph_DT = Data_Asph5(:,6);
Asph_PT = Data_Asph5(:,7);
Asph_Std = Data_Asph5(:,8);
Asph_VV5 = Data_Asph5(:,10);
```

```matlab
Asph_VH5 = Data_Asph5(:,11);
Asph_HH5 = Data_Asph5(:,12);
Asph_VVHH5 = Data_Asph5(:,10)./Data_Asph5(:,12);
Asph_VHHH5 = Data_Asph5(:,11)./Data_Asph5(:,12);
Asph_DT5 = Data_Asph5(:,14);
Asph_PT5 = Data_Asph5(:,15);
Asph_Std5 = Data_Asph5(:,16);
Asph_VV18 = Data_Asph18(:,10);
Asph_VH18 = Data_Asph18(:,11);
Asph_HH18 = Data_Asph18(:,12);
Asph_VVHH18 = Data_Asph18(:,10)./Data_Asph18(:,12);
Asph_VHHH18 = Data_Asph18(:,11)./Data_Asph18(:,12);
Asph_DT18 = Data_Asph18(:,14);
Asph_PT18 = Data_Asph18(:,15);
Asph_Std18 = Data_Asph18(:,16);

%Grass

Grass_P = Data_Grass5(:,5);
Grass_DT = Data_Grass5(:,6);
Grass_PT = Data_Grass5(:,7);
Grass_Std = Data_Grass5(:,8);
Grass_VV5 = Data_Grass5(:,10);
Grass_VH5 = Data_Grass5(:,11);
Grass_HH5 = Data_Grass5(:,12);
Grass_VVHH5 = Data_Grass5(:,10)./Data_Grass5(:,12);
Grass_VHHH5 = Data_Grass5(:,11)./Data_Grass5(:,12);
Grass_DT5 = Data_Grass5(:,14);
Grass_PT5 = Data_Grass5(:,15);
Grass_Std5 = Data_Grass5(:,16);
Grass_VV18 = Data_Grass18(:,10);
Grass_VH18 = Data_Grass18(:,11);
Grass_HH18 = Data_Grass18(:,12);
Grass_VVHH18 = Data_Grass18(:,10)./Data_Grass18(:,12);
Grass_VHHH18 = Data_Grass18(:,11)./Data_Grass18(:,12);
Grass_DT18 = Data_Grass18(:,14);
Grass_PT18 = Data_Grass18(:,15);
Grass_Std18 = Data_Grass18(:,16);

%Gravel

Grav_P = Data_Grav5(:,5);
Grav_DT = Data_Grav5(:,6);
Grav_PT = Data_Grav5(:,7);
Grav_Std = Data_Grav5(:,8);
Grav_VV5 = Data_Grav5(:,10);
Grav_VH5 = Data_Grav5(:,11);
Grav_HH5 = Data_Grav5(:,12);
Grav_VVHH5 = Data_Grav5(:,10)./Data_Grav5(:,12);
Grav_VHHH5 = Data_Grav5(:,11)./Data_Grav5(:,12);
Grav_DT5 = Data_Grav5(:,14);
Grav_PT5 = Data_Grav5(:,15);
Grav_Std5 = Data_Grav5(:,16);
Grav_VV18 = Data_Grav18(:,10);
Grav_VH18 = Data_Grav18(:,11);
Grav_HH18 = Data_Grav18(:,12);
Grav_VVHH18 = Data_Grav18(:,10)./Data_Grav18(:,12);
Grav_VHHH18 = Data_Grav18(:,11)./Data_Grav18(:,12);
Grav_DT18 = Data_Grav18(:,14);
Grav_PT18 = Data_Grav18(:,15);
Grav_Std18 = Data_Grav18(:,16);
```

```matlab
%Sand

Sand_P = Data_Sand5(:,5);
Sand_DT = Data_Sand5(:,6);
Sand_PT = Data_Sand5(:,7);
Sand_Std = Data_Sand5(:,8);
Sand_VV5 = Data_Sand5(:,10);
Sand_VH5 = Data_Sand5(:,11);
Sand_HH5 = Data_Sand5(:,12);
Sand_VVHH5 = Data_Sand5(:,10)./Data_Sand5(:,12);
Sand_VHHH5 = Data_Sand5(:,11)./Data_Sand5(:,12);
Sand_DT5 = Data_Sand5(:,14);
Sand_PT5 = Data_Sand5(:,15);
Sand_Std5 = Data_Sand5(:,16);
Sand_VV18 = Data_Sand18(:,10);
Sand_VH18 = Data_Sand18(:,11);
Sand_HH18 = Data_Sand18(:,12);
Sand_VVHH18 = Data_Sand18(:,10)./Data_Sand18(:,12);
Sand_VHHH18 = Data_Sand18(:,11)./Data_Sand18(:,12);
Sand_DT18 = Data_Sand18(:,14);
Sand_PT18 = Data_Sand18(:,15);
Sand_Std18 = Data_Sand18(:,16);

%Asphalt Smooth

Asph2_P = Data_Asph25(:,5);
Asph2_DT = Data_Asph25(:,6);
Asph2_PT = Data_Asph25(:,7);
Asph2_Std = Data_Asph25(:,8);
Asph2_VV5 = Data_Asph25(:,10);
Asph2_VH5 = Data_Asph25(:,11);
Asph2_HH5 = Data_Asph25(:,12);
Asph2_VVHH5 = Data_Asph25(:,10)./Data_Asph25(:,12);
Asph2_VHHH5 = Data_Asph25(:,11)./Data_Asph25(:,12);
Asph2_DT5 = Data_Asph25(:,14);
Asph2_PT5 = Data_Asph25(:,15);
Asph2_Std5 = Data_Asph25(:,16);
Asph2_VV18 = Data_Asph218(:,10);
Asph2_VH18 = Data_Asph218(:,11);
Asph2_HH18 = Data_Asph218(:,12);
Asph2_VVHH18 = Data_Asph218(:,10)./Data_Asph218(:,12);
Asph2_VHHH18 = Data_Asph218(:,11)./Data_Asph218(:,12);
Asph2_DT18 = Data_Asph218(:,14);
Asph2_PT18 = Data_Asph218(:,15);
Asph2_Std18 = Data_Asph218(:,16);


G1 = eval(['Grass_' num2str(Scatter1)]);
G2 = eval(['Grass_' num2str(Scatter2)]);
G3 = eval(['Grass_' num2str(Scatter3)]);
G4 = eval(['Grass_' num2str(Scatter4)]);
G5 = eval(['Grass_' num2str(Scatter5)]);
G6 = eval(['Grass_' num2str(Scatter6)]);
G1p = mle(G1);
G2p = mle(G2);
G3p = mle(G3);
G4p = mle(G4);
G5p = mle(G5);
G6p = mle(G6);
```

216

```matlab
%Build the matrix%
S_Grass = [G1, G2, G3, G4, G5, G6];
Sm_Grass = [G1p(1), G2p(1), G3p(1), G4p(1), G5p(1), G6p(1)];
Sp_Grass = [G1p(2), G2p(2), G3p(2), G4p(2), G5p(2), G6p(2)];
%S_Grass = [G1, G2, G3];


V1 = eval(['Grav_' num2str(Scatter1)]);
V2 = eval(['Grav_' num2str(Scatter2)]);
V3 = eval(['Grav_' num2str(Scatter3)]);
V4 = eval(['Grav_' num2str(Scatter4)]);
V5 = eval(['Grav_' num2str(Scatter5)]);
V6 = eval(['Grav_' num2str(Scatter6)]);
V1p = mle(V1);
V2p = mle(V2);
V3p = mle(V3);
V4p = mle(V4);
V5p = mle(V5);
V6p = mle(V6);

%Build the matrix%
S_Grav = [V1, V2, V3, V4, V5, V6];
Sm_Grav = [V1p(1), V2p(1), V3p(1), V4p(1), V5p(1), V6p(1)];
Sp_Grav = [V1p(2), V2p(2), V3p(2), V4p(2), V5p(2), V6p(2)];
%S_Grav = [V1, V2, V3];

A1 = eval(['Asph_' num2str(Scatter1)]);
A2 = eval(['Asph_' num2str(Scatter2)]);
A3 = eval(['Asph_' num2str(Scatter3)]);
A4 = eval(['Asph_' num2str(Scatter4)]);
A5 = eval(['Asph_' num2str(Scatter5)]);
A6 = eval(['Asph_' num2str(Scatter6)]);
A1p = mle(A1);
A2p = mle(A2);
A3p = mle(A3);
A4p = mle(A4);
A5p = mle(A5);
A6p = mle(A6);

%Build the matrix%
S_Asph = [A1, A2, A3, A4, A5, A6];
Sm_Asph = [A1p(1), A2p(1), A3p(1), A4p(1), A5p(1), A6p(1)];
Sp_Asph = [A1p(2), A2p(2), A3p(2), A4p(2), A5p(2), A6p(2)];
%S_Asph = [A1, A2, A3];

S1 = eval(['Sand_' num2str(Scatter1)]);
S2 = eval(['Sand_' num2str(Scatter2)]);
S3 = eval(['Sand_' num2str(Scatter3)]);
S4 = eval(['Sand_' num2str(Scatter4)]);
S5 = eval(['Sand_' num2str(Scatter5)]);
S6 = eval(['Sand_' num2str(Scatter6)]);
Sa1p = mle(S1);
Sa2p = mle(S2);
Sa3p = mle(S3);
Sa4p = mle(S4);
Sa5p = mle(S5);
Sa6p = mle(S6);

%Build the matrix%
S_Sand = [S1, S2, S3, S4, S5, S6];
Sm_Sand = [Sa1p(1), Sa2p(1), Sa3p(1), Sa4p(1), Sa5p(1), Sa6p(1)];
```

```matlab
Sp_Sand = [Sa1p(2), Sa2p(2), Sa3p(2), Sa4p(2), Sa5p(2), Sa6p(2)];
%S_Sand = [S1, S2, S3];


U1 = eval(['Asph2_' num2str(Scatter1)]);
U2 = eval(['Asph2_' num2str(Scatter2)]);
U3 = eval(['Asph2_' num2str(Scatter3)]);
U4 = eval(['Asph2_' num2str(Scatter4)]);
U5 = eval(['Asph2_' num2str(Scatter5)]);
U6 = eval(['Asph2_' num2str(Scatter6)]);
U1p = mle(U1);
U2p = mle(U2);
U3p = mle(U3);
U4p = mle(U4);
U5p = mle(U5);
U6p = mle(U6);


%Build the matrix%
S_Asph2 = [U1, U2, U3, U4, U5, U6];
Sm_Asph2 = [U1p(1), U2p(1), U3p(1), U4p(1), U5p(1), U6p(1)];
Sp_Asph2 = [U1p(2), U2p(2), U3p(2), U4p(2), U5p(2), U6p(2)];
%S_Asph2 = [U1, U2, U3];


S1 = eval(['S_' num2str(Sname1)]);
S2 = eval(['S_' num2str(Sname2)]);
S3 = eval(['S_' num2str(Sname3)]);
S4 = eval(['S_' num2str(Sname4)]);
S5 = eval(['S_' num2str(Sname5)]);
S1m = eval(['Sm_' num2str(Sname1)]);
S2m = eval(['Sm_' num2str(Sname2)]);
S3m = eval(['Sm_' num2str(Sname3)]);
S4m = eval(['Sm_' num2str(Sname4)]);
S5m = eval(['Sm_' num2str(Sname5)]);
S1p = eval(['Sp_' num2str(Sname1)]);
S2p = eval(['Sp_' num2str(Sname2)]);
S3p = eval(['Sp_' num2str(Sname3)]);
S4p = eval(['Sp_' num2str(Sname4)]);
S5p = eval(['Sp_' num2str(Sname5)]);
h=normplot(A6)
%Length of arrays
LS(1) = length(S1);
LS(2) = length(S2);
LS(3) = length(S3);
LS(4) = length(S4);
LS(5) = length(S5);


%Applying MLE method


%Calculating of confusion matrix

A = zeros(5, 5);                        %Confusion matrix
ERR = zeros(5, 5);


kj = 1;


while kj < 6;
    knnkj = zeros(1,5);
    knnkj = knnkj + knn;
    %knnkj(kj) = knn + 1;
    i = 1;

    DSKJ = zeros(5, 1);
```

```matlab
    while i < LS(kj) + 1;
        SKJ = eval(['S' num2str(kj)]);
        newpoint = SKJ(i,:);
        %
        jp = 1;
        D1p = 0;
        D2p = 0;
        D3p = 0;
        D4p = 0;
        D5p = 0;
        jp_length = length(newpoint);

while jp < jp_length + 1
            %D1p = normpdf(X, mu, sigma);
            D1p = D1p + log10(normpdf(newpoint(jp), S1m(jp), S1p(jp)));
            D2p = D2p + log10(normpdf(newpoint(jp), S2m(jp), S2p(jp)));
            D3p = D3p + log10(normpdf(newpoint(jp), S3m(jp), S3p(jp)));
            D4p = D4p + log10(normpdf(newpoint(jp), S4m(jp), S4p(jp)));
            D5p = D5p + log10(normpdf(newpoint(jp), S5m(jp), S5p(jp)));
            %K-means method
            %D1p = D1p + abs(newpoint(jp)- S1m(jp))/S1p(jp);
            %D2p = D2p + abs(newpoint(jp)- S2m(jp))/S2p(jp);
            %D3p = D3p + abs(newpoint(jp)- S3m(jp))/S3p(jp);
            %D4p = D4p + abs(newpoint(jp)- S4m(jp))/S4p(jp);
            %D5p = D5p + abs(newpoint(jp)- S5m(jp))/S5p(jp);

                        jp = jp +1;
        end;


        DSKJ(1) = D1p;
        DSKJ(2) = D2p;
        DSKJ(3) = D3p;
        DSKJ(4) = D4p;
        DSKJ(5) = D5p;

        DSKM = max(DSKJ);

        index = find(DSKJ == max(DSKJ(:)));
        ERR(kj, index) = ERR(kj, index) + 1;


    A(kj, 1) = ERR(kj, 1)/LS(kj);
    A(kj, 2) = ERR(kj, 2)/LS(kj);
    A(kj, 3) = ERR(kj, 3)/LS(kj);
    A(kj, 4) = ERR(kj, 4)/LS(kj);
    A(kj, 5) = ERR(kj, 5)/LS(kj);
    %A(kj, kj) = ERR(kj, kj)/LS(kj);
        i = i +1;
    end;

    kj = kj + 1;

end;


%Write confusion matrix to disk
```

```
FilenameB =
strcat(num2str(Scatter1),'_',num2str(Scatter2),'_',num2str(Scatter3),'_ML
C', '.csv') ;
%FilenameB = strcat('Confusion_Matrix1.csv') ;
csvwrite(FilenameB, A);

%Write header of acoustic echo file
fid_in = fopen(FilenameB, 'r');
content=fread(fid_in);
fclose(fid_in);
fid_out = fopen(FilenameB, 'w');
fprintf(fid_out, 'Asph,Asph_Sm,Grass,Gravel,Sand\n');
fwrite(fid_out, content);
fclose(fid_out);
```

....

KNN_Classifier

```
%Matlab version (R2014a)
%File name: KNN_Classifier.m
%Purpose: KNN Classification Algorithm.

close all;
clear all;
clc;

knn = 2;

%Selecting the surfaces for comparison

Sname1 = 'Asph';
Sname2 = 'Asph2';
%Sname2 = 'Asph';
Sname3 = 'Grass';
%Sname3 = 'Grav';
Sname4 = 'Grav';
Sname5 = 'Sand';
%Sname5 = 'Grav';


%Importing Excel file with all data%

%filename = 'Database_2013_3.xls';
filename = 'Database_202Neural.xls';
sheet = 1;
[Data, Text, Alldata] = xlsread(filename);

%Selecting necessary data for the scatters%
```

```matlab
Data_S = Data (2:517, [2,8,9,10,27,28,29,30,12, 17,18,19,20,21,22,23]);
%2 - 669
%Plotting the scatters in 3D
Scatter1 = 'P';
Scatter2 = 'DT';
Scatter3 = 'P';
Scatter4 = 'P';
Scatter5 = 'DT';
Scatter6 = 'DT';
%P, DT, PT, Std, VV5, VH5, Std5, VVHH5, VHHH5  /18


%1 - Record number
%2 - Surface: Gravel 1; Asphalt 3; Grass 5
%3 - Covered: No info 0; Dry 1; Wet 2; Snow 3; Ice 4; Frost 5
%4 - Details: No info 0; Soft 1; Rammed 2; Wet 3;
%5 - Ac Power
%6 - Ac Dur Thresh
%7 - Ac Power Thresh
%8 - Ac STD
%9 - Frequency
%10 - 17 VV
%11 - 18 VH
%12 - 19 HH
%13 - 20 HV
%14 - 21 DT
%15 - 22 PT
%16 - 23 Std

idx_5 = (Data_S(:,9) == 5.8);
Data_5 = Data_S(idx_5,:);
idx_18 = (Data_S(:,9) == 18);
Data_18 = Data_S(idx_18,:);



%Sorting by surfaces ACOUSTIC AND 5.8GHz

idx_Asph = (Data_5(:,2)==3); %Asphalt
Data_Asph5 = Data_5(idx_Asph,:);
idx_Grass = (Data_5(:,2)==5); %Grass
Data_Grass5 = Data_5(idx_Grass,:);
idx_Grav = (Data_5(:,2)==1); %Gravel
Data_Grav5 = Data_5(idx_Grav,:);
idx_Asph2 = (Data_5(:,2)==8); %Asph Smooth
Data_Asph25 = Data_5(idx_Asph2,:);
%idx_Sand = (Data_A(:,2)==7); %Sand(ALL))
idx_Sand = (Data_5(:,2)==7 & Data_5(:,4)==1); %Sand(Smooth only)
Data_Sand5 = Data_5(idx_Sand,:);


%Sorting by surfaces 18GHz

idx_Asph = (Data_18(:,2)==3); %Asphalt
Data_Asph18 = Data_18(idx_Asph,:);
idx_Grass = (Data_18(:,2)==5); %Grass
Data_Grass18 = Data_18(idx_Grass,:);
idx_Grav = (Data_18(:,2)==1); %Gravel
Data_Grav18 = Data_18(idx_Grav,:);
idx_Asph2 = (Data_18(:,2)==8); %Asph Smooth
Data_Asph218 = Data_18(idx_Asph2,:);
%idx_Sand = (Data_18(:,2)==7); %Sand(ALL))
idx_Sand = (Data_18(:,2)==7 & Data_18(:,4)==1); %Sand(Smooth only)
```

```matlab
Data_Sand18 = Data_18(idx_Sand,:);



%Asphalt

Asph_P = Data_Asph5(:,5);
Asph_DT = Data_Asph5(:,6);
Asph_PT = Data_Asph5(:,7);
Asph_Std = Data_Asph5(:,8);
Asph_VV5 = Data_Asph5(:,10);
Asph_VH5 = Data_Asph5(:,11);
Asph_HH5 = Data_Asph5(:,12);
Asph_VVHH5 = Data_Asph5(:,10)./Data_Asph5(:,12);
Asph_VHHH5 = Data_Asph5(:,11)./Data_Asph5(:,12);
Asph_DT5 = Data_Asph5(:,14);
Asph_PT5 = Data_Asph5(:,15);
Asph_Std5 = Data_Asph5(:,16);
Asph_VV18 = Data_Asph18(:,10);
Asph_VH18 = Data_Asph18(:,11);
Asph_HH18 = Data_Asph18(:,12);
Asph_VVHH18 = Data_Asph18(:,10)./Data_Asph18(:,12);
Asph_VHHH18 = Data_Asph18(:,11)./Data_Asph18(:,12);
Asph_DT18 = Data_Asph18(:,14);
Asph_PT18 = Data_Asph18(:,15);
Asph_Std18 = Data_Asph18(:,16);

%Grass

Grass_P = Data_Grass5(:,5);
Grass_DT = Data_Grass5(:,6);
Grass_PT = Data_Grass5(:,7);
Grass_Std = Data_Grass5(:,8);
Grass_VV5 = Data_Grass5(:,10);
Grass_VH5 = Data_Grass5(:,11);
Grass_HH5 = Data_Grass5(:,12);
Grass_VVHH5 = Data_Grass5(:,10)./Data_Grass5(:,12);
Grass_VHHH5 = Data_Grass5(:,11)./Data_Grass5(:,12);
Grass_DT5 = Data_Grass5(:,14);
Grass_PT5 = Data_Grass5(:,15);
Grass_Std5 = Data_Grass5(:,16);
Grass_VV18 = Data_Grass18(:,10);
Grass_VH18 = Data_Grass18(:,11);
Grass_HH18 = Data_Grass18(:,12);
Grass_VVHH18 = Data_Grass18(:,10)./Data_Grass18(:,12);
Grass_VHHH18 = Data_Grass18(:,11)./Data_Grass18(:,12);
Grass_DT18 = Data_Grass18(:,14);
Grass_PT18 = Data_Grass18(:,15);
Grass_Std18 = Data_Grass18(:,16);

%Gravel

Grav_P = Data_Grav5(:,5);
Grav_DT = Data_Grav5(:,6);
Grav_PT = Data_Grav5(:,7);
Grav_Std = Data_Grav5(:,8);
Grav_VV5 = Data_Grav5(:,10);
Grav_VH5 = Data_Grav5(:,11);
Grav_HH5 = Data_Grav5(:,12);
Grav_VVHH5 = Data_Grav5(:,10)./Data_Grav5(:,12);
Grav_VHHH5 = Data_Grav5(:,11)./Data_Grav5(:,12);
Grav_DT5 = Data_Grav5(:,14);
```

222

```matlab
Grav_PT5 = Data_Grav5(:,15);
Grav_Std5 = Data_Grav5(:,16);
Grav_VV18 = Data_Grav18(:,10);
Grav_VH18 = Data_Grav18(:,11);
Grav_HH18 = Data_Grav18(:,12);
Grav_VVHH18 = Data_Grav18(:,10)./Data_Grav18(:,12);
Grav_VHHH18 = Data_Grav18(:,11)./Data_Grav18(:,12);
Grav_DT18 = Data_Grav18(:,14);
Grav_PT18 = Data_Grav18(:,15);
Grav_Std18 = Data_Grav18(:,16);


%Sand

Sand_P = Data_Sand5(:,5);
Sand_DT = Data_Sand5(:,6);
Sand_PT = Data_Sand5(:,7);
Sand_Std = Data_Sand5(:,8);
Sand_VV5 = Data_Sand5(:,10);
Sand_VH5 = Data_Sand5(:,11);
Sand_HH5 = Data_Sand5(:,12);
Sand_VVHH5 = Data_Sand5(:,10)./Data_Sand5(:,12);
Sand_VHHH5 = Data_Sand5(:,11)./Data_Sand5(:,12);
Sand_DT5 = Data_Sand5(:,14);
Sand_PT5 = Data_Sand5(:,15);
Sand_Std5 = Data_Sand5(:,16);
Sand_VV18 = Data_Sand18(:,10);
Sand_VH18 = Data_Sand18(:,11);
Sand_HH18 = Data_Sand18(:,12);
Sand_VVHH18 = Data_Sand18(:,10)./Data_Sand18(:,12);
Sand_VHHH18 = Data_Sand18(:,11)./Data_Sand18(:,12);
Sand_DT18 = Data_Sand18(:,14);
Sand_PT18 = Data_Sand18(:,15);
Sand_Std18 = Data_Sand18(:,16);


%Asphalt Smooth
Asph2_P = Data_Asph25(:,5);
Asph2_DT = Data_Asph25(:,6);
Asph2_PT = Data_Asph25(:,7);
Asph2_Std = Data_Asph25(:,8);
Asph2_VV5 = Data_Asph25(:,10);
Asph2_VH5 = Data_Asph25(:,11);
Asph2_HH5 = Data_Asph25(:,12);
Asph2_VVHH5 = Data_Asph25(:,10)./Data_Asph25(:,12);
Asph2_VHHH5 = Data_Asph25(:,11)./Data_Asph25(:,12);
Asph2_DT5 = Data_Asph25(:,14);
Asph2_PT5 = Data_Asph25(:,15);
Asph2_Std5 = Data_Asph25(:,16);
Asph2_VV18 = Data_Asph218(:,10);
Asph2_VH18 = Data_Asph218(:,11);
Asph2_HH18 = Data_Asph218(:,12);
Asph2_VVHH18 = Data_Asph218(:,10)./Data_Asph218(:,12);
Asph2_VHHH18 = Data_Asph218(:,11)./Data_Asph218(:,12);
Asph2_DT18 = Data_Asph218(:,14);
Asph2_PT18 = Data_Asph218(:,15);
Asph2_Std18 = Data_Asph218(:,16);



G1 = eval(['Grass_' num2str(Scatter1)]);
G2 = eval(['Grass_' num2str(Scatter2)]);
G3 = eval(['Grass_' num2str(Scatter3)]);
G4 = eval(['Grass_' num2str(Scatter4)]);
```

```matlab
G5 = eval(['Grass_' num2str(Scatter5)]);
G6 = eval(['Grass_' num2str(Scatter6)]);
%Build the matrix%
S_Grass = [G1, G2, G3, G4, G5, G6];
%S_Grass = [G1, G2, G3];

V1 = eval(['Grav_' num2str(Scatter1)]);
V2 = eval(['Grav_' num2str(Scatter2)]);
V3 = eval(['Grav_' num2str(Scatter3)]);
V4 = eval(['Grav_' num2str(Scatter4)]);
V5 = eval(['Grav_' num2str(Scatter5)]);
V6 = eval(['Grav_' num2str(Scatter6)]);
%Build the matrix%
S_Grav = [V1, V2, V3, V4, V5, V6];
%S_Grav = [V1, V2, V3];

A1 = eval(['Asph_' num2str(Scatter1)]);
A2 = eval(['Asph_' num2str(Scatter2)]);
A3 = eval(['Asph_' num2str(Scatter3)]);
A4 = eval(['Asph_' num2str(Scatter4)]);
A5 = eval(['Asph_' num2str(Scatter5)]);
A6 = eval(['Asph_' num2str(Scatter6)]);
%Build the matrix%
S_Asph = [A1, A2, A3, A4, A5, A6];
%S_Asph = [A1, A2, A3];

S1 = eval(['Sand_' num2str(Scatter1)]);
S2 = eval(['Sand_' num2str(Scatter2)]);
S3 = eval(['Sand_' num2str(Scatter3)]);
S4 = eval(['Sand_' num2str(Scatter4)]);
S5 = eval(['Sand_' num2str(Scatter5)]);
S6 = eval(['Sand_' num2str(Scatter6)]);
%Build the matrix%
S_Sand = [S1, S2, S3, S4, S5, S6];
%S_Sand = [S1, S2, S3];

U1 = eval(['Asph2_' num2str(Scatter1)]);
U2 = eval(['Asph2_' num2str(Scatter2)]);
U3 = eval(['Asph2_' num2str(Scatter3)]);
U4 = eval(['Asph2_' num2str(Scatter4)]);
U5 = eval(['Asph2_' num2str(Scatter5)]);
U6 = eval(['Asph2_' num2str(Scatter6)]);
%Build the matrix%
S_Asph2 = [U1, U2, U3, U4, U5, U6];
%S_Asph2 = [U1, U2, U3];

S1 = eval(['S_' num2str(Sname1)]);
S2 = eval(['S_' num2str(Sname2)]);
S3 = eval(['S_' num2str(Sname3)]);
S4 = eval(['S_' num2str(Sname4)]);
S5 = eval(['S_' num2str(Sname5)]);

%Length of arrays
LS(1) = length(S1);
LS(2) = length(S2);
LS(3) = length(S3);
LS(4) = length(S4);
LS(5) = length(S5);

%Applying KNN method
```

```matlab
%Calculation of confusion matrix

A = zeros(5, 5);                                %Confusion matrix
ERR = zeros(5, 5);

kj = 1;

while kj < 6;
    knnkj = zeros(1,5);
    knnkj = knnkj + knn;
    knnkj(kj) = knn + 1;
    i = 1;

    DSKJ = zeros(5, 1);

    while i < LS(kj) + 1;
        SKJ = eval(['S' num2str(kj)]);
        newpoint = SKJ(i,:);
        [IDX1, D1] = knnsearch (S1, newpoint, 'k', knnkj(1));
        [IDX2, D2] = knnsearch (S2, newpoint, 'k', knnkj(2));
        [IDX3, D3] = knnsearch (S3, newpoint, 'k', knnkj(3));
        [IDX4, D4] = knnsearch (S4, newpoint, 'k', knnkj(4));
        [IDX5, D5] = knnsearch (S5, newpoint, 'k', knnkj(5));

        DSKJ(1) = sum(D1);
        DSKJ(2) = sum(D2);
        DSKJ(3) = sum(D3);
        DSKJ(4) = sum(D4);
        DSKJ(5) = sum(D5);

        DSKM = min(DSKJ);

        index = find(DSKJ == min(DSKJ(:)));
        ERR(kj, index) = ERR(kj, index) + 1;


    A(kj, 1) = ERR(kj, 1)/LS(kj);
    A(kj, 2) = ERR(kj, 2)/LS(kj);
    A(kj, 3) = ERR(kj, 3)/LS(kj);
    A(kj, 4) = ERR(kj, 4)/LS(kj);
    A(kj, 5) = ERR(kj, 5)/LS(kj);
    %A(kj, kj) = ERR(kj, kj)/LS(kj);
        i = i +1;
     end;

    kj = kj + 1;

end;



%Write confusion matrix to disk
FilenameB =
strcat(num2str(Scatter1),'_',num2str(Scatter2),'_',num2str(Scatter3),'_kn
n', num2str(knn),'.csv') ;
%FilenameB = strcat('Confusion_Matrix1.csv') ;
csvwrite(FilenameB, A);

%Write header of acoustic echo file
fid_in = fopen(FilenameB, 'r');
```

```matlab
content=fread(fid_in);
fclose(fid_in);
fid_out = fopen(FilenameB, 'w');
fprintf(fid_out, 'Asph,Asph_Sm,Grass,Gravel,Sand\n');
fwrite(fid_out, content);
fclose(fid_out);

%scatter3(S_Asph(:,1),S_Asph(:,2),S_Asph(:,3), 'k');
%hold on;
scatter3(S_Asph(:,1),S_Asph(:,2),S_Asph(:,3), 'b');
hold on;
%scatter3(S_Grav(:,1),S_Grav(:,2),S_Grav(:,3), 'r');
%hold on;
%scatter3(S_Grass(:,1),S_Grass(:,2),S_Grass(:,3), 'g');
%hold on;
scatter3(S_Sand(:,1),S_Sand(:,2),S_Sand(:,3), 'm');
%legend('Asphalt', 'Smooth Asphalt', 'Gravel', 'Grass', 'Sand');
legend('Asphalt', 'Sand');
labels2 = num2str((1:size(S_Sand,1))','%d');    %'
text(S_Sand(:,1), S_Sand(:,2), S_Sand(:,3), labels2,
'horizontal','left','vertical','bottom', 'Color','red', 'FontSize', 7);
xlabel(Scatter1);
ylabel(Scatter2);
zlabel(Scatter3);
```

Appendix G: Publications

1.   A. Bystrov, M. Abbas, E. Hoare, T-Y. Tran, N. Clarke, M. Gashinova and M. Cherniakov: "Remote road surface identification using radar and ultrasonic sensors", in *Proc. European Radar Conference,* 2014, Italy.

2.   A. Bystrov, M. Abbas, E. Hoare, T-Y. Tran, N. Clarke, M. Gashinova and M. Cherniakov: "Analysis of Classification Algorithms Applied to Road Surface Recognition", in *Proc. IEEE International Radar Conference,* 2015, USA.

3.   A. Bystrov, M. Abbas, E. Hoare, T-Y. Tran, N. Clarke, M. Gashinova and M. Cherniakov: "Road surface classification using sonar", in *Proc. European Radar Conference,* 2015, France. (waiting for acceptance)