Multiple-Sensor Based Approach for Road Terrain Classification

Shifeng Wang

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Technology, Sydney

Faculty of Engineering and Information Technology

Centre for Autonomous Systems

Intelligent Mechatronic Systems Group

Student : Shifeng Wang

Supervisor : Dr. Sarath Kodagoda

Certificate of Original Authorship

I, Shifeng Wang, certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signed: _____

Date: _____

Acknowledgment

I would like to express my heartfelt gratitude to my supervisor Dr. Sarath Kodagoda for his helpful insights, intelligent ideas and supporting. From the moment I met him when I came to UTS, he continually assisted me throughout these years resulting in this thesis. I really appreciate his patience and the directions he showed to me when I met difficulties in my research. I would like to say that it is my honour being your student.

Thanks to the people who have offered me great helps and cooperation. Firstly my colleague Dr. Stephan Sehestedt who was my neighbour and friend, he was my role model being a good researcher with a professional attitude. Dr. Zhan Wang and Dr. Rami Khushaba, they advised me a lot regarding mathematics and Matlab coding. My colleague and friend Lei Shi, he always offered me great help on programming and some suggests at anytime. Dr. Alen Alempijevic and Dr. Stephan Sehestedt who developed and set up the experiment platform CRUISE where my research work carried out. And Mr. Marian Himstedt, Dr. Shan Lynn, Dr. Jack Wang, Mr. Xiang Luo, Mr. Zulkarnain Zainudin, Mr. Xiang Ren, Miss Yan Li, and Mr. Kanzhi Wu, thanks for your jobs during the data collection experiments.

My sincere thanks go to the University of Technology Sydney who offered me scholarships all along my candidature. It would be a mission impossible without financial support for me.

Special thanks to my wife Lei Chen and lovely daughter Yifei Wang. I cannot image if I can finish this PhD without your understanding and supporting. I hope I can bring our family a better future. Thanks to my parents for their constant support in my life. Hope you are proud of me.

Contents

LIST OF ABBREVIATIONS

CRUISE	CAS Research Ute for Intelligence, Safety and Exploration
c/m	cycles/meter
DARPA	Defense Advanced Research Projects Agency
DMU	Differential Measurement Unit
FFT	Fast Fourier Transform
FPS	Frame per Second
FWT	Fast Wavelet Transform
GLCM	Grey-Level Co-occurrence Matrix
GPS	Global Positioning System
IMU	Inertial Measurement Unit
LRF	Laser Range Finder
MSE	Mean Squared Errors
MV	Majority Vote
PCA	Principal Component Analysis
PSD	Power Spectral Density
RPM	Revolutions per Minute
SVM	Support Vector Machine
USB	Universal Serial Bus

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Complete perception of the environment around a vehicle plays a crucial role in safe driving. Those include other stationary or dynamic objects, drivable regions, road signs and terrain types. In this thesis, classification of terrain types is investigated. Knowledge of the road terrain is useful to improve passengers' safety and comfort in road vehicles. It is also useful for bounding safe navigational routes for autonomous vehicles. Therefore, the aim of this thesis is to develop a methodology to identify impending road terrain types by using on-board sensors.

Two kinds of sensors are used in this research. The first kind of sensors measure parameters of the vehicle in terms of vibration and speed. Therefore, these sensors can only produce measurements while the vehicle is navigating on a particular terrain. The second type of sensors measure properties of the terrain in terms of structure and visual cues, for example cameras and Laser Range Finders (LRFs). These sensors can even produce measurements of impending terrain types. However, all those kinds of sensors have their own advantages and disadvantages. In this thesis, it is proposed to fuse them to improve the terrain type classification results.

The sensor fusion is achieved using Markov Random Field (MRF). The MRF model is designed to contain five nodes and five cliques which describe the relationships between the classification results of the accelerometer, camera, and two LRFs. The MRF model's energy function is appropriately synthesized to improve the classification accuracies of the impending terrains. Experiments carried out on a real vehicle test-bed, CRUISE (CAS Research Ute for Intelligence, Safety and Exploration) on different types of roads with

various speeds show that the MRF based fusion algorithm lead to significant improvements (approximate 30%) of the road terrain classification accuracies.

Keywords—*road terrain classification, Markov Random Field, sensor data fusion, acceleration, image, Laser Range Finder*

CHAPTER I

INTRODUCTION

1.1 Background

For many people, a car is more than a convenience as it ensures the necessary mobility. However traffic jams and accidents now occur in higher numbers as more urban areas come alive. Their growing populations, more and more of them are mobile because they are able to afford cars, and expanding human activity lead to more urban traffic. This individual transport leads to waste of space and precious natural resources. The advent of cheaper and more powerful sensors and computer units allows developing smarter transportation systems to ease or solve these issues. Smart cars are often envisioned to be a cornerstone of these smart transportation systems.

In the past decades, mobile robotic systems in particular autonomous cars [1] have been developed with significant achievements worldwide [2]. An autonomous car, also called driverless or self-driving car, is seen as a significant achievement to safely transport people in future traffic systems. Such a car is capable of sensing its environment and navigating a given road system using a variety of internal and external and powerful computers. Since the 1980s, significant advances have been made in relevant technologies. Numerous significant companies and research organizations have developed working in the field of autonomous vehicles, e.g., Mercedes Benz, Google, Continental Automotive Systems, Bosch, Nissan, Toyota, Audi, Oxford University, Massachusetts Institute of Technology. The state of Nevada in the United States of America has passed the first law

considering autonomous cars on its roads in 2011. The first license for an autonomous car was issued by the Nevada Department of Motor Vehicles in May 2012. In February 2013 Florida and California followed by passing laws considering driverless cars.

Today's autonomous cars observe the surroundings using multiple sensors, such as Laser Range Finders (LRF), cameras, radars, the Global Positioning System (GPS), Inertial Measurement Units (IMU), accelerometers, and wheel encoders. All the data from sensors is analysed and processed by computers in the car and motion commands are produced. Advanced control systems interpret sensory information to identify and decide appropriate navigation paths. The obstacles ahead and objects around, and even signage on the road are also involved for the autonomous vehicle to determine appropriate driving manoeuvres. One useful piece of information, which can in principle be extracted from sensor data, is the type of road which is currently in the field of view.

1.2 Motivation

Information about the road terrain is one of the important information an autonomous vehicle should acquire [3] [4]. Different road terrains have vastly different physical characteristics. Road terrain classification using on-board sensors can provide important, sometimes crucial information regarding safety, fuel efficiency and passenger comfort.

Different road terrain can have a significant impact on vehicle handling, ride, quality, and stability [5]. The tyre traction properties that implicate longitudinal and lateral wheel slip are dictated by the terrain's physical parameters such as the friction coefficient, soil cohesion, and internal friction angle [6] [7] [8] [9]. The normal force acting on a tyre was modulated by terrain roughness, which affects the handling characteristics of the

vehicles. For instance, when a four-wheel drive vehicle moves on mud, rocks, and snow, the corresponding driving manoeuvre could be very different [10].

Figure 1.1 A vehicle on mud terrain [11]

When a vehicle travels on muddy terrain, shown in Figure 1.1, the wheels of the vehicle tend to get stuck. To prevent the vehicle spinning and digging into the mud [11], the traction must be controlled very carefully. The driver should drive slowly and resist high acceleration. If on sand, the good operation is to steer smoothly with gear changes at high Revolutions per Minute (RPM). The driver should avoid soft sand at the base of the dunes and gullies. Once moving on them, he/she should make turns as wide as possible and do not brake rapidly. When stopping the car, decelerating slowly is better than braking down in a sudden.

Figure 1.2 A vehicle on rocky terrain [11]

When a vehicle moves on rocky terrain, shown in Figure 1.2, slip tends to occur between the tires and the top layer of rocks. The driver should avoid sharp turns and quick acceleration/ deceleration that may cause the vehicle to skid or even flip over [11].

Figure 1.3 A vehicle on snow terrain [11]

When a vehicle travels on snow or icy terrain, shown in Figure 1.3, the force of the tires to grip the road surface will significantly reduce. This makes the motions of speeding up, slowing down, and changing direction potentially dangerous. A moderate driving manner is the best way to drive in such conditions [11]. The car should move slowly while

the driver should leave extra space to slow down and stop. To avoid braking with locking the wheels, the driver is suggested to shift into a lower gear allowing the speed to fall and then gently operate the brakes. If skidding occurs, the driver should stop stepping on the acceleration pedal and turn into the skid direction immediately.

Consequently, the estimation of road terrain is very important for vehicles. The knowledge of the terrain provides crucial information for determining appropriate and safe driving manoeuvres. The aim of this thesis is to develop an approach to identify the road terrain by using the on-board sensors of a car.

1.3 Contributions

The contributions arising from this thesis are:

- Road terrain classification was tested under three types of sensor modalities. The results were compared and pros and cons of each modality were analysed. The classification problem was formulated and discussed.
- The multimodal sensor fusion strategy was proposed to achieve better road terrain classification results for impending road terrain types.
- An MRF framework was proposed to fuse accelerometer, camera, and LRF data.
- The proposed methods were tested on real data captured on real driving scenarios to validate the methodologies.

1.4 Publications

Following is the list of publications resulting from the work presented in this thesis:

S. Wang, S. Kodagoda, Rami, Khushaba, “Towards Speed-Independent Road-Type Classification”, International Conference on Control, Automation, Robotics and Vision (ICARCV), Guangzhou, China, 5-7, December, 2012.

S. Wang, S. Kodagoda, R. Ranasinghe, “Road Terrain Type Classification Based on Laser Measurement System Data”, Australasian Conference on Robotics and Automation (ACRA), Wellington, New Zealand, 3-5, December, 2012.

S. Wang, S. Kodagoda, Z. Wang, G. Dissanayake, “Multiple Sensor Based Terrain Classification”, Australasian Conference on Robotics and Automation (ACRA), Melbourne, Australia, 7-9, December, 2011.

1.5 Thesis Overview

This thesis is organized as follows:

Chapter I outlines the background of this research. The motivation and contributions are introduced in this chapter.

Chapter II summarizes related work. Researches relevant to road terrain type identification are reviewed. Their achievements and problems are presented as well.

Chapters III, IV and V present acceleration based, image based, and LRF based road terrain classification approaches respectively. These three chapters address classifying the road terrain using a single sensor. The methodologies and experiments of each approach are presented. The advantages and disadvantages of each sensor working on road terrain classification tasks are indicated in each of these three chapters.

Chapter VI presents a multiple-sensor based road terrain classification approach using MRF algorithm to fuse the classification results from different sensors. The MRF model for this case and its corresponding energy function are designed. An optimization method is presented to solve the energy function. The MRF experiment results show that this multiple-sensor method using MRF fusion method significantly improves the road terrain identification accuracy.

Chapter VII concludes the research work presented in this thesis. The methodologies of the road terrain classification are summarized. Finally, future research directions are proposed.

CHAPTER II

REVIEW OF RELATED WORK

The knowledge of road terrain types vehicles drive through plays a crucial role in vehicle driving safety. It also provides important information for the driving manoeuvre. There have been a number of researchers working on the problem of road terrain type identification. A range of sensors were used to conduct the research. The popular sensors used among them were the accelerometer, camera, and LRF. For different applications and tasks, some research teams applied only one type of sensor while others utilized multiple types of sensors. Most of the research involved machine learning theory, more specifically, the classification and prediction technique.

In this chapter, the related work of accelerometer, camera LRF and multi-sensor based approaches are reviewed.

2.1 Accelerometer Applications

The usage of the accelerometer for road terrain classification was extensive in the literature. Some were based on small sized rover platforms [12] [13] [14] [15] [16] [17] while others were based on road vehicle platforms [18] [19] [20] [21]. The small sized rover platform is very different from a road vehicle one. Firstly, the dynamic model of the vehicle has to be considered when using the accelerometer to calculate the road profile. However the small sized rover can be roughly treated as a rigid object. The second difference is the speed. The small sized rover usually moves at a low speed in a range from

0.1 *m/s* to 2 *m/s*. But the speed of a road vehicle usually ranges from 20 *km/h* to 80 *km/h*. This affects the methodology of the terrain classification.

2.1.1 Small Sized Rover Platform Using Accelerometer

A simplified experimental platform was introduced by Weiss et al [22]. An accelerometer and a computer were equipped on a small cart to collect the data. The fixed speed was manually controlled at 1 *m/s*. The target was to distinguish six different terrain types including indoor floor, asphalt, gravel, grass, paving, and boule court. The collected acceleration data was split for off-line training and testing purposes. There were four different features formed from raw acceleration data. The first two were a log-scaled Power Spectral Density (PSD) and a 128-point Fast Fourier Transform (FFT) which had been used by other researchers [10] [23]. The third feature introduced was calculated from raw acceleration vectors. This feature was simpler, more compact and faster to compute than the PSD and the FFT. The fourth feature tested was a combination of the feature representation and the PSD. The Support Vector Machine (SVM) classifier was used for the classification. The comparison experiment showed the fourth feature proposed significantly improved the classification results than that of the PSD and the FFT. Additional experiment showed that the Principal Component Analysis (PCA) did not help to improve the classification accuracy. It should be noted that although this method achieved a total classification accuracy of 95%, the experiment conditions needed to be rich.

A well-equipped experiment platform named ‘Pioneer 2-AT’ was introduced by Ojeda et al [24]. There were number of sensors available, including inertial sensors: two Coriolis gyros, one fibre-optic gyro, and three accelerometers; motor sensors: two current

and two voltage sensors; range sensors: one ultrasonic range sensor and one infrared range sensor; other sensors: one microphone and two wheel encoders. Seven types of terrain were tested to be classified employing the Neural Network classifier. The methods using different sensors for the terrain classification were investigated. Complete comparison experiments show that the X-axis gyro provided the best signal for Neural Network based terrain classification purpose. Moreover, the inertial sensors were capable of classifying gravel and pavement very well where the terrain caused distinct vibrations whereas they did not work well on dirt, sand and grass terrains. It was inferred that a multiple-sensor feature based input could help Neural Network classifier to find out the "preferred" information for the classification. As the same in [22], this research did not concentrate on the speed variety. All the logged data was based on experiments with a fixed speed of 30 *cm/s*.

More speed data was involved in the terrain classification research by Sadhukhan [10]. Similar with [22] [24], the vertical acceleration data was transferred by FFT to form the feature matrix. But the difference was the platform was operated at speeds of 0.2, 0.4, 0.6, and 0.8 *m/s*. The data sets logged at different speeds were separated for the classification experiments. Part of the data set was used for training and part for testing. The comparison experiments show that the data sets with higher speeds tended to allow higher classification accuracies. Hence, the classification accuracies depend on the platform speed using this FFT feature extraction method. The moderate speed data was then selected for the terrain classification task. The experiments showed classification accuracy of about 90% for gravel, grass, and packed dirt terrains when the probabilistic neural network classifier was used. A classifier comparison experiment was also performed. The probabilistic neural network classifier presented higher classification

accuracy than that of the back propagation network classifier. Moreover, its training time was relatively lower than that of the back propagation classifier, which was shown to be less computationally expensive.

It was not only vertical acceleration information that was considered for the terrain classification; the angular velocities of the *x* and *y* axes were also involved in the task by Dupont et al [25]. The experimental platform named robots ARTV Jr was equipped with a Differential Measurement Unit (DMU) sensor. The data of the rover's vertical acceleration and angular velocities of the *x* and *y* axes were collected when the rover moved at 0.6, 0.8, 1.0, 1.2, and 1.4 *m/s*. The logged data was then divided into segments of 10 second intervals. The magnitude frequency response was calculated for each segment using FFT to form the feature matrix. PCA was applied to the feature matrix to calculate the eigenspace value for reducing the feature matrix dimensions. The experiments show that 25 dimensional vectors of the feature matrix produced the best classification result. The training data contained speeds of 0.6, 1.0, and 1.4 *m/s* while the testing data contained speeds of 0.8 and 1.2 *m/s*. The probabilistic neural network classifier was employed to create the nonlinear boundaries within the eigenspace. The final experiments showed that this method was able to classify gravel, grass, and asphalt terrains on specific speeds, but not for the data sets with all speeds.

Vibration characteristics include vertical direction, horizontal and moving directions. Weiss et al [26] utilized all this information to solve the terrain classification problem. Referring to his previous work [22], the vibration-based terrain classification method was used, this method using a three-axis accelerometer worked well on a common outdoor rover. In [26] the experiments were conducted at three different speeds:

0.2, 0.4, and 0.6 *m/s*. Six types of terrain which were indoor, asphalt, gravel, grass, paving, and clay were trained and tested by an SVM classifier. The experiment results showed that the acceleration measurements in multiple directions were able to add more helpful information into the feature matrix to obtain higher classification accuracy. It should be noted that [10], [22], [24], and [25] all investigated this terrain classification task with specific or relatively constant speeds.

2.1.2 Road Vehicle Using Accelerometer

The experiment platforms used in [10], [22], [24], and [25] were all small sized robots which can be treated as rigid objects. There was no need to build a model for a small robot body. But for the road vehicles, it has to be a dynamic model to describe the vehicle's body while studying the vibration characteristics. A half-car model along with an accelerometer mounted on the road vehicle was applied to estimate the road condition by González et al [20]. The road profile was calculated from the acceleration data and the speed data. Then the spatial frequency was calculated from the road profile to form the PSD matrix. This approach required prior knowledge of the vehicle transform function which was relevant to vehicle parameters. Simulations of half-car models have been carried out to analyse the proposed classification method for road roughness, speed, vehicle parameters and noise. The method was shown to be robust enough for a number of road profiles, vehicle speeds and dynamic scenarios. This research involved spatial frequency term which is not usual time frequency. The car was treated as a dynamic model but not a rigid object. All simulation experiments were sound but real world results were missing.

The experiments with real-world conditions were carried out by Ward et al [21] using a road vehicle. The vehicle equipped with an accelerometer and a GPS was employed to collect data for terrain classification. The accelerometer mounted on the suspension of the vehicle measured the vertical vibration while the GPS was used to measure the vehicle's speed. The acceleration data was logged with varying speeds of the vehicle. The road profile was calculated using these two data along with the one-quarter car dynamic model. Then, the spatial frequency features were extracted from the road profile to form the feature matrix. The aim of this research was to classify asphalt, brick, gravel, and rumble strips terrains. The presented method was claimed to have superior classification accuracies at varying degrees of speed. Although this could be a valid argument in ideal scenarios where the model of the vehicle was best represented, more analysis needed to be carried out to verify the claim. In this case, the speed independency is investigated in Section 3.8.2 while the methodology of acceleration based terrain classification is presented in Chapter III.

2.2 Camera Applications

A mobile robot must traverse a variety of natural or urban terrains which could be flat, uneven, rough, or rugged. The geometrical and physical properties of the terrain raise the complexity of the navigation problem [27]. The mobile robot must be capable to assess the terrain, even determine the risk associated with the traversal of each terrain segment in some cases.

2.2.1 Small Sized Rover Using Camera

Howard et al [28] proposed a method for terrain characterization and assessment of terrain traversability for a rover. The rover was equipped with a camera capturing

images ahead of the rover. Features such as the roughness, slope, discontinuity, and hardness of the images were extracted. PCA was applied to the feature matrix for dimensionality reduction. The methodology for rough terrain assessment and classification employed the fuzzy logic and neural network, which was particularly suitable for planetary rovers.

Another planetary rover was set up by Thompson et al [29]. A method for characterizing geology during the rover traverse was presented. The discrete geologic features were detected and classified these features from the captured images. While these features profiling results in ambiguous signatures, the complex features extracted from rock detectors were accurate enough both for finding drivable region boundaries and for providing basic summaries of local geology. However, this method was encouraged to combine satellite data to provide additional information to make travel routes. But for many rover applications, no satellite is available.

Castano et al [30] focused on terrain classification using texture feature in rover navigation application. In this research, a number of images captured during autonomous navigation on cross-country terrain were used to perform texture-based classification. The purpose of this texture analysis was to form part of the closed loop that allowed a robotic system to navigate autonomously. Two representative texture-based classification methods were investigated: mixture model and histogram. The experiments show that the performances were rather satisfactory, and in the case of the geological database, the classification accuracy was much higher than that of the bound of random classification. Castano et al suggested using additional features, such as colour and shape.

The method proposed by Karlsen et al [31] for robot navigation tasks was different from [28], [29], and [30], a prior image processing was segmentation. And then the characteristic colour/luminance and texture of the terrain were extracted. The preliminary results indicated that this method had potential to segment terrain in a manner which was consistent with subjective perception. The benefit of segmentation was to provide robustness over changes in lighting, automatic camera gain and contrast adjustment. But more work on automatic gain and colour distortion was still needed. The decision tree algorithm was used in the experiment to partition the data, based on the independent variables and the dependent variables. The experiment showed the roughness and ground resistance was identified.

2.2.2 Road Vehicle Using Camera

An adaptive partition based random forest classification for a road vehicle on the outdoor terrain was proposed by Chetan et al [32]. The classifier was a combination of two random forest classifiers. One was trained by bootstrapped or off-line data while the other adapted to changes on-line. Posterior probabilities of both classifiers were fused to produce the final label. Each image was divided into small patches that each was 16 by 16 pixels. Each of the whole images was then classified into five groups to mark the drivable region. The average classification accuracy was around 80%. The experiment results show that the adaptive classifier was able to adapt to changes which occur during a traversal while holding on to what was learned at bootstrap.

Song et al [33] proposed a method that did not learn on the fly but was self-supervised. As for the car moving in urban environments, terrain classification was performed to find the drivable region ahead of the car. But before that, moving objects

were firstly detected and tracked. These objects were then classified to find out which humans were and which vehicles were. After the removal of the moving objects, the path could be bounded. Finally, colour and texture features were extracted from the path region of the images to classify the sidewalk region and driveable road. This self-supervised framework observed moving objects and learned from its environment based on the paths of moving objects. This approach did not need any model for the shape of the terrain but learned about terrains through observations.

Moving away from an urban environment, there was an application in the arctic [34]. The texture features were also essential information to assess the slope of the terrains. Williams et al [34] presented a methodology to estimate density and slope for the traversabilty. But for the case of this thesis, the texture features were extracted from the entire image to report the terrain types. This means each image only produces one classification result and not a huge number of results from the patches for one image which is specified in Chapter IV.

2.3 LRF Applications

An accelerometer senses the vibration signal for terrain classification. But LRF scans the environment to perceive the terrain structure. As for a mobile robot, an LRF sensor was also considered to be suitable to measure the height and roughness of the terrain, the objects on the road, etc.

2.3.1 Small Sized Rover Using LRF

In an application of bipedal robots, a 2D LRF sensor was used by Nishiwaki et al [35] to solve the terrain perception problem. It was achieved that the robot could navigate on

unknown rough terrains without using any assumptions or prior knowledge of the terrains. With the usage of LRF, the terrain was reconstructed in a large-scale map for the robot to plan paths.

LRF was also applied for an outdoor rover. Macedo et al [36] used a 2D LRF to discriminate grass for the autonomous navigation. A statistical analysis of the range measurements was presented to solve the problem of distinguishing grass from obstacles in a vegetated environment. The obstacle that was partially occluded by grass could be identified by this method. This research did not work on other objects but grass which was not very adaptable for other applications.

Montemerlo et al [37] concentrated on identifying the terrain by dividing view fields into patches with different sizes. This method used a terrain model which arose from the limited spatial resolution of the point cloud. The limited spatial resolution was due to the long range observation by LRF. This prevented small obstacles from being detected from long distances. Therefore a Bayes filter estimating the state of terrain segments was then proposed to deal with the problem of the difference in observation resolution. This multi-resolution method maintained multiple navigation maps, and derived rational arguments for the number of layers and the resolutions. The hierarchy of maps was devised and each map tuned to the different laser range which was a different observation resolution. A rational argument to determine the region that was covered by each map was mathematically bounded. The grid cells in each map were also provided. Hence, this made the map covered by much higher resolution. This approach presented successfully identified obstacles on the terrain of the mobile robot environment.

Andersen et al [38] did not divide the terrain map into segments. The terrain was classified to find a traversable region through scans. The paved and dirt terrains were identified by a 2D LRF. The algorithm for the terrain classification fused seven distinct characteristics: raw height, roughness, step size, curvature, slope, width and invalid data. Roughness that was the average deviation was calculated from a best-fit line over an equal distance of the laser scan. The laser scans fulfilling height criteria and adaptive roughness criteria were grouped. The roughness, step size, slope, and width were used for hypothesis testing and classification. The tilted down LRF allowed road surface scanning which reconstructed slopes and uneven terrain. The outdoor experiments showed that this approach was capable of detecting passable roads and obstacles that allowed the robot to autonomously navigate.

2.3.2 Road Vehicle Using LRF

As for the outdoor environment for a road vehicle, a 3D LRF was also applied to the navigation task. A 3D LRF provides point cloud with millions of measurement range values which contain rich information about the scanned environment. By applying 3D LRF, Vandapel et al [39] and Lalonde et al [40] classified natural terrain into three classes: clutter to capture grass and tree canopy; linear to capture thin objects, such as wires and tree branches; and surface to capture solid objects, such as ground surface, rocks and tree trunks. The saliency features were extracted to define surface, curve, and point in the local area. The scanning pattern, range resolution, noise level of the 3D LRF and variability of the scenes were captured using statistical classification techniques. A Gaussian mixture model was fit and trained to perform Bayesian classification. The experiments were implemented on an autonomous ground vehicle. The experiment results showed the

limitations of this approach. Due to the huge amount of points for each scan, the computational efficiency problem was suggested to be solved by processing only the point cloud data from the moving direction.

Urmson et al of Red Team [41] concentrated on high speed navigation for road vehicles in unrehearsed terrain. The platform was equipped with a range of sensors, including forward LRF, supplemental LRFs, stereo cameras, radar, and GPS. The forward LRF was the primary sensor to provide a terrain model with relatively long ranges. Due to the vibration of the vehicle while moving, the forward LRF was particularly affected by mechanical excitation in the pitch axis. This caused dramatic changes of the terrain scanned by LRF at a reasonably long distance. So the overall perception was degraded and convoluted with interdependencies on terrain dynamics, the speed of the vehicle, chassis, and sensor-mount characteristic responses. The gimbal was applied to reduce this effect. The terrain evaluation technique presented in [42] [43] was used. This technique applied each laser scan to return a vector of traversability scores, and each point of the scan indicated a safety coefficient. The range values of each laser scans were evaluated to find slopes, residuals, and the direction of the unsafe region. Once the current line had been evaluated, it was then clustered into the previous group with a score. Thus, the relatively safe region could be marked and found. The vehicle could pass through this traversable terrain. This vehicle named Sandstorm associated Red Team attended the 2004 Defence Advanced Research Projects Agency (DARPA) Grand Challenge.

Stavens et al [44] presented another terrain roughness estimation approach. It was indicated that the accurate detection of terrain roughness was significantly more important than that of obstacle avoidance for an autonomous vehicle when it was moving

at high speeds. The road terrain was detected by LRF. The estimation of terrain roughness was achieved by a self-supervised machine learning approach. The comparison of scanning points on the road surface was hard due to uncertainty. This was because the relatively long range points may be quite sparse so much, so that significant information about the surface was missed. Moreover, due to latency or error in pose estimation, these laser points of the higher precision project were not available. The solution was to model the sources of error as a multivariate polynomial of which the coefficients were obtained by self-supervised learning process. The drivable regions were then labelled automatically when the vehicle was driving over the target terrain. The LRF data was trained in this way, and terrain roughness was estimated by the classifier through this self-supervised approach. In this case, the target was not to find regions to drive through, but to classify the road terrain types by testing the laser scans (specified in Chapter V).

2.4 Mutiple-Sensor Applications

Each of the acceleration based, image based, and LRF based road terrain classification approaches has its own advantages and specialities. It was also considered by some researchers to apply multiple sensors in the task of terrain classification to improve traversability prediction accuracy [45].

2.4.1 Small Sized Rover Using Multiple Sensors

Halatci et al [46] employed an accelerometer mounted on the one leg of the rover and a camera capturing images ahead. Two classifications were implemented. The colour, texture, and range features extracted from images were processed based on maximum likelihood estimation and SVM. On the other hand, the vibration features extracted from acceleration data was trained and tested. Two different techniques to merge the results of

these two classification result channels were Bayesian fusion and meta-classifier fusion. The experiment results showed that the classification accuracy can be improved by classifier fusion.

Brooks et al [47] worked on the same platform but utilized a different sensor combination, that of a microphone and a camera. The microphone measured the vibration of the rover's leg while the camera captured images in front. A self-supervised classification method was presented that automatically trained an image classifier based on terrain classes that were tested out by another classifier. The experiment indicated that the self-supervised classifier worked better than a manually trained classifier based on prior image data set. This suggested that self-supervised classification did not prefer manually labelled training but a self-supervised one. This was because the self-supervised classification adapted changing illumination better which affected image quality. Additionally, the rover was set running at a fixed low speed of 6 *cm/s* which could reduce many potential adverse factors.

The experimental platform set up by Weiss et al [48] was performed in an outdoor environment set at a speed a little higher than 0.45 *m/s*. This rover was equipped with an accelerometer to measure the vibration of the rover's body and a camera looking forward. The terrain classification approach fused terrain predictions based on image data and current terrain classified by acceleration data. The SVM classifier was employed to classify fourteen types of terrain. The comparison experiments showed that the fusion sensor based classification accuracies were significantly higher than the classification of the individual approach alone.

It was different with [46] [47] [48], Gollins et al [49] applied a different sensor combination. An accelerometer and a LRF were used to perform the terrain classification. The vibration based terrain classification worked on the current terrain patch while image based terrain classification worked on the forthcoming terrain patch. Gollins et al further indicated the limitations of the methods. Visual based method could easily be fooled by fallen leaves or other superficial coverings on the ground. Vibration based method yielded a poor classification accuracy when the rover speed was significantly different with the training speeds. Although the possible solution to deal with this problem was presented, they were simulation experiments and not real-world ones.

Lu et al [50] presented a method employing a regular camera that worked at day time; infrared cameras worked on classifying mud; and a laser stripe-based sensor observed the forthcoming terrain. The features extracted from images in this approach excluded colours which were easily affected by illumination, but included spatial response and texture features which presented the spatial relationships between different grey levels. The laser stripe-based structured light sensor was employed to classify the terrains. The range values provided were used to extract spatial features. The experiments showed the effectiveness of the proposed method with an overall classification accuracy of 90%. Lu et al also indicated that the experiment performances were done at low speed which might not be applicable to other applications.

Manduchi [51] focused on cross-country environments rather than urban environments for the robot navigation using multiple sensors of a colour stereo camera and a 2D LRF. The approach contained two tasks which were an obstacle detection and terrain classification. The obstacle was detected by the stereo range measurement which

was a colour-based classification to find and label the obstacles according to a set of specific terrains. The grass terrain was distinguished from obstacles (such as tree trunks and rocks) by analysing the LRF data. The experiment results showed the viability and robustness of the approach for obstacle detection, colour-based terrain classification and obstacle labelling in the grass. However, this methodology was strictly limited in the grass vegetation environment which was not applicable to the urban environment.

2.4.2 Road Vehicle Using Multiple Sensors

In road vehicle applications, there have been very few related researches that applied multiple sensors. An urban autonomous vehicle navigation task was brought by Posner et al [52] who employed a camera and a 3D LRF. This was an appearance-based approach to identify urban environments with higher-order, semantic labels. A chain of binary classifiers labelled individual LRF data according to their origin. These LRF points were then characterised by 3D geometric data and visual cues extracted from images. The experiments showed the classification scheme was sufficient to consistently distinguish different types of terrains and walls, as well as bushes and foliage. Additionally, some common objects such as cars and vans were also recognized.

Concerning the task of this thesis, the purpose is to classify the terrain types instead of identifying the objects on the road. Multiple sensors are applied in this case as they are expected to improve the classification accuracy and reliability.

2.5 Conclusion

A number of techniques proposed in the literature have been summarized in this chapter. In terms of sensor applications, individual sensors which could be an accelerometer, camera, and LRF could be used for the terrain classification task. But just as some researchers suggested, each of the sensors had its own shortages and limitations. Hence, multiple-sensor based terrain classification techniques were introduced to overcome the problem. The combinations of accelerometer and camera; the camera and LRF were popularly used. Moreover, most of the researches in the literature were based on small sized robots working at a fixed low speed. The focus of this thesis is to devise classification strategies for road vehicles operating at a vast range of speeds. It is also intended to predict the upcoming road terrain types. This needs more specific and intense concentration and further investigations.

CHAPTER III

ACCELERATION BASED ROAD TERRAIN CLASSIFICATION

Road type classification is the process of categorizing road terrain into different types such as asphalt, concrete, grass and gravel. Intuitively, the amount of vibration that is caused by a vehicle navigating on a particular road type is a valuable source of information. Therefore, collecting a vehicle's vibration information to obtain specific characteristics of different road terrains is of high interest. For this reason, an accelerometer is mounted on the suspension to measure the vertical component of the vibration of the vehicle. Four-wheel encoders are used to measure the speed of the vehicle. Along with the one quarter model of the vehicle [53], road profiles are estimated. Six features are considered and tested for the classification to find out the best feature for the experiment. These six features tested are raw data of acceleration, raw data of road profile, time frequency features from acceleration, spatial frequency from road profile, fast wavelet transform features from acceleration, and fast wavelet transform feature from road profile.

In this chapter, the acceleration based method to classify the road terrain is presented. Firstly, it describes the road profile estimation process and the feature extraction from the road profile data. Then the normalization, PCA, K-Fold Cross Validation are introduced to form the feature matrix. Three typical different classifiers are

compared to select a most appropriate classifier. Finally, the acceleration based experiment and the results are presented.

3.1 Road Profile Estimation

There are six features tested to compare with each other in this case. Some are extracted from raw acceleration data while some from road profile data. In this section, the process to estimate the road profile from the acceleration data is presented. Figure 3.1 shows the procedure of road profile estimation by using acceleration data, speed data and the quarter vehicle model. The detailed method is described as follows:

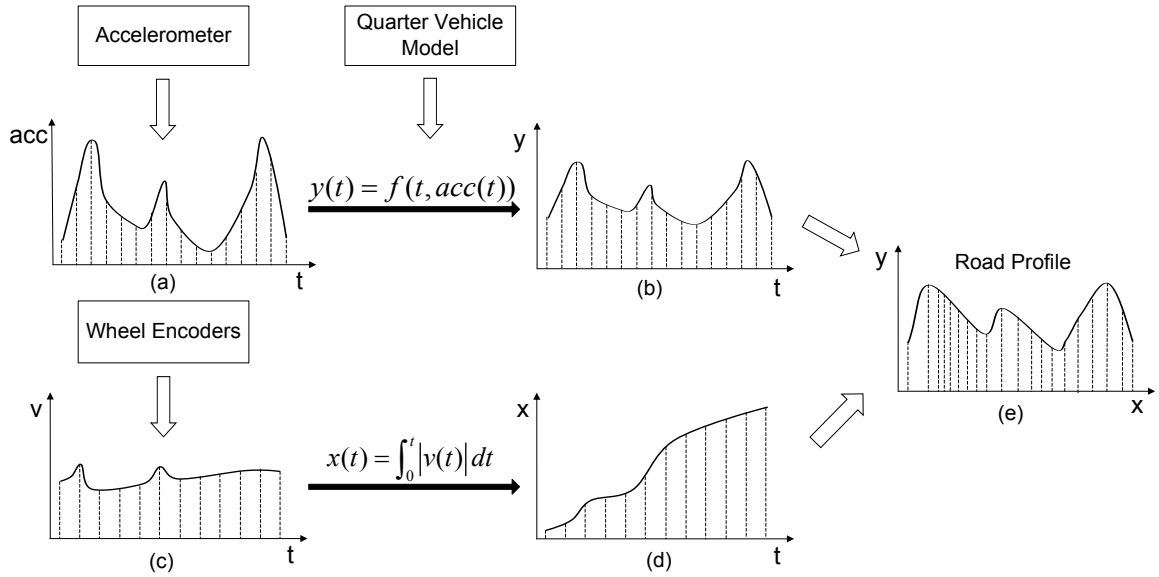


Figure 3.1 Illustration of road profile estimation

3.1.1 Acceleration (acc-t)

An accelerometer is mounted on the suspension of the vehicle (close proximity to the rear right wheel) to measure the vertical acceleration while the vehicle is in motion. The measurement range of the accelerometer is ±6 *g* (±58.8 N/s^2) which is sufficient for the vibration detection according to many prior tests.

The acceleration signal is transmitted to the computer via a Universal Serial Bus (USB) cable. The acceleration data is logged with timestamp information for synchronization. The sampling frequency is fixed at 50 *Hz*. The provided time varying raw acceleration data as shown in Figure 3.1 (a) presents vibration characteristic of the vehicle.

3.1.2 Quarter Vehicle Model (acc-t to y-t)

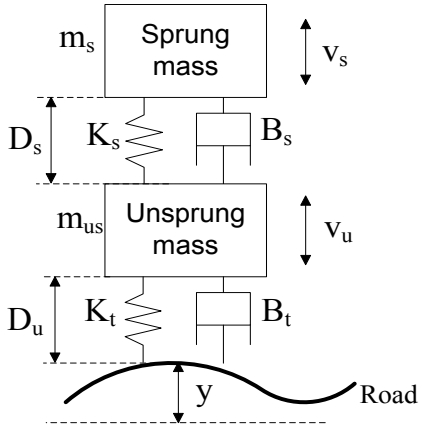


Figure 3.2 the quarter car model of the experimental vehicle

The standard quarter car model [53], shown in Figure 3.2, is used in the road surface estimation. It is comprised of the sprung mass of quarter body of the vehicle, m_s, and its vertical velocity, v_s; the suspension modelled as a spring and a damper system described by spring coefficient, K_s and damping coefficient, B_s; unsprung mass of quarter body of the vehicle, m_{us} and its vertical velocity, v_u; the pneumatic tyre modelled as another spring and damper system described by spring coefficient, K_t and damping coefficient, B_t. In addition, D_s, D_u are the compression lengths of the two springs.

Thus, the quarter vehicle model can be expressed as [21]:

$$\begin{bmatrix} v_s' \\ v_u' \\ D_s' \\ D_u' \end{bmatrix} = \begin{bmatrix} -\frac{B_s}{m_s} & \frac{B_s}{m_s} & \frac{K_s}{m_s} & 0 \\ \frac{B_s}{m_{us}} & -\frac{B_s + B_t}{m_{us}} & -\frac{K_s}{m_{us}} & \frac{K_t}{m_{us}} \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_s \\ v_u \\ D_s \\ D_u \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{B_t}{m_{us}} \\ 0 \\ 1 \end{bmatrix} \cdot y' \tag{3.1}$$

where v_s' is $\frac{dv_s}{dt}$, v_u' is $\frac{dv_u}{dt}$, D_s' is $\frac{dD_s}{dt}$, D_u' is $\frac{dD_u}{dt}$, and y' is $\frac{dy}{dt}$. Then the transfer function of unsprung mass vertical acceleration of road profiles is given by [21]:

$$\begin{aligned} \frac{v_u'(s)}{y(s)} = & \frac{m_s B_t s^5 + (m_s K_t + B_s B_t)s^4 + (B_t K_s + B_s K_t)s^3 + K_s K_t s^2}{m_s m_{us} s^4 + (m_s B_s + m_s B_t + m_{us} B_s)s^3} \\ & + (m_s K_s + m_s K_t + B_s B_t + m_{us} K_s + B_s^2 - B_s^2 m_{us} / m_s)s^2 \\ & + (B_s K_s + B_s K_t + B_t K_s - B_s K_s m_{us} / m_s)s + K_s K_t \end{aligned} \tag{3.2}$$

Table 3.1 Parameters for the quarter car model

Symbol	Quantity	Value and unit
m_s	Sprung mass of quarter body of the vehicle	500 *kg*
m_{us}	Unsprung mass of quarter body of the vehicle	65 *kg*
K_s	Spring coefficient of the suspension	80,000 *N/m*
K_t	Spring coefficient of the tyre	200,000 *N/m*
B_s	Damping coefficient of the suspension	3,500 *kg/s*
B_t	Damping coefficient of the tyre	50 *kg/s*

The experimental platform CAS Research Ute for Intelligence, Safety and Exploration (CRUISE), a Ford Courier utility vehicle (see Figure 3.15), has the parameters shown in Table 3.1. By substituting these parameters into (3.2), (3.4) can be derived which is a function of the variables acceleration and time:

$$y(t) = f(acc(t), t)$$

$$y(t) = \frac{435\cos(\frac{\sqrt{591}}{2}t)acc(t) + \frac{783557 \times \sqrt{591}\sin(\frac{\sqrt{591}}{2}t)}{1285425}}{199652 \times 10^{3.5t}} \tag{3.3}$$

$$+\frac{2560728983acc(t)}{1996520000 \times 10^{4000t}} + \frac{609acc(t)}{40000} + t \cdot acc(t)$$

That is:

$$y(t) = \frac{\left[435\cos(12.16t)acc(t) + 14.82\sin(12.16t)\right]}{/(199652 \times 10^{3.5t}) + 2560728983acc(t)}$$

$$+\frac{2560728983acc(t)}{(1996520000 \times 10^{4000t})} + 0.02acc(t) + t \cdot acc(t)$$

$$y(t) = \left[435\cos(12.16t)acc(t) + 14.82\sin(12.16t)\right] \tag{3.4}$$

$$/(199652 \times 10^{3.5t}) + 2560728983acc(t) / (1996520000 \times 10^{4000t})$$
$$+0.02acc(t) + t \cdot acc(t)$$

3.1.3 Vertical Displacement (y-t)

It should be noted that the curve *y(t)* shown in Figure 3.1 (b) is not the real road profile *y(x)*, even though in some situations the shape of it is very similar to the real road profile.

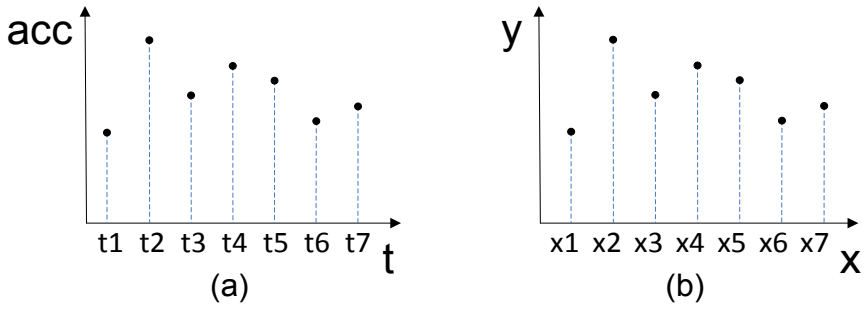


Figure 3.3 Even sampling and even profile with constant speed

For instance, Figure 3.3 (a) shows a regular sampling event, where the horizontal axis is time *t* while the vertical axis is the sampled value, which here is the measured acceleration value *acc*. The interval distances between every point are exactly the same. When transferring *acc-t* into *y-x*, as shown in Figure 3.3 (b), the curve of *y-x* looks the same as *acc-t,* given the vehicle's speed is constant. The horizontal axis is the vehicle's displacement *x* while the vertical axis is displacement *y* of the road, i.e. uneven ground, as can be seen in Figure 3.3 (b).

As can be inferred, applying FFT to *acc-t* and *y-x* respectively, would yield the same results, although applying FFT to *acc-t* happens in the time frequency domain while applying FFT to *y-x* happens in the spatial frequency domain.

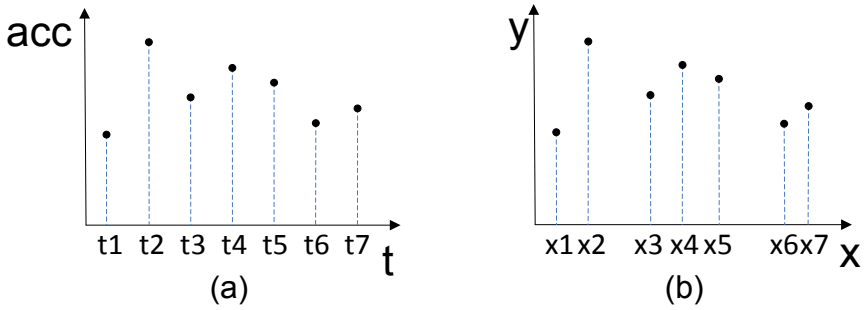


Figure 3.4 Even sampling and uneven profile with varying speed

The situation would change if the condition of the speed is changed. Even though a regular sampling to acceleration is implemented as shown in Figure 3.4 (a), the presented road profile would be similar to what is shown in Figure 3.4 (b), if the speed of the vehicle is changed. The distances between the points are not equal. Obviously, applying the FFT to such a curve would give a different frequency characteristic compared to the curve in Figure 3.3 (b). For this reason, interpolation is employed to make the points distributed evenly on a horizontal axis before the FFT implementation.

Therefore, the horizontal distance between every two sampled points is uneven in distance while it is even in time. However, if the vehicle could maintain its constant speed for a particular duration, then there is no much apparent difference with either *y(t)* or *y(x)*.

3.1.4 Speed (v-t)

Hall Effect sensors working as wheel encoders are part of the Anti-lock Braking System (ABS). They are used to obtain the vehicle's speed. The sine-like signals are isolated and amplified, then processed by a microcontroller to act as speed data. As shown in Figure 3.1 (c), the speed of the vehicle can be obtained as:

$$v(t) = \frac{\pi D_w f_v}{4p} \sum_{i=1}^{4} n_i \tag{3.5}$$

where D_w is the external diameter of the tyre; f_v is the sampling frequency, 10 *Hz*, which depends on how often the microcontroller transfers counting data to the computer; p is the number of teeth around the wheel hub, which in this case is 44 teeth; i is the index of the four wheels; and n_i is a number from each wheel, which indicates how many teeth have passed by in 0.1 *second*.

It is to be noted that the current system cannot measure a speed below 5 *km/h*. This is due to the low amplitude of the signal which could not be captured by the microcontroller.

3.1.5 Speed to Displacement (v-t to x-t)

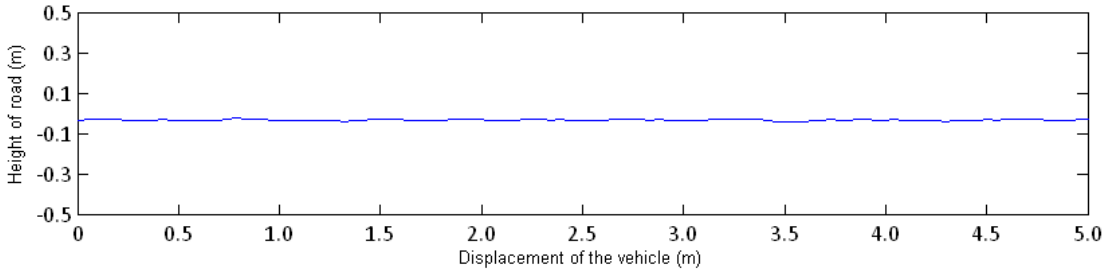
As shown in Figure 3.1 (d), the vehicle displacement is estimated as:

$$x(t) = \int_0^t \left| v(t) \right| dt \tag{3.6}$$

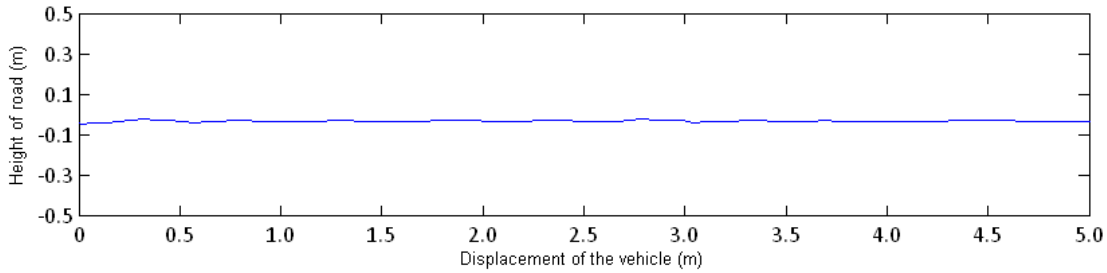
where *v(t)* is the measured vehicle speed obtained by (3.5). Since the sampling frequency is 50 *Hz* for the accelerometer and 10 *Hz* for the speed, a nearest-neighbour interpolation (also known as proximal interpolation) [54] is applied to *x(t)* data to make certain each vertical displacement y_i has a corresponding horizontal displacement data x_i.

3.1.6 Road Profile (y-x)

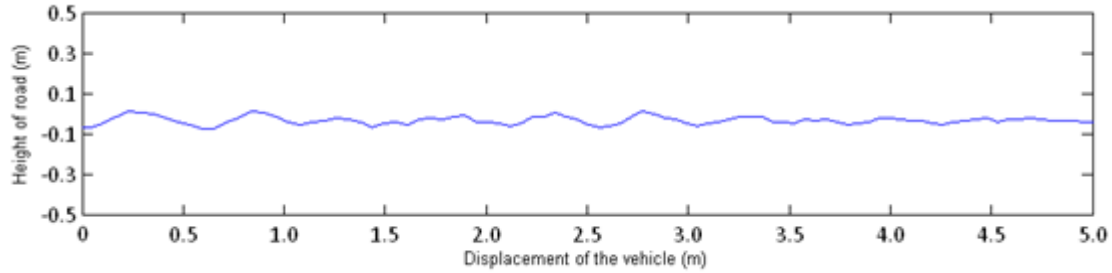
Finally, *y(x)* is estimated from *acc(t)* and *v(t)* via the quarter vehicle model. However as illustrated in Figure 3.1 (e), on the horizontal axis, the points do not scatter with equal intervals. As previously mentioned, the sampling is based on a fixed time interval, but not the vehicle's displacement position. To solve this problem, a nearest neighbour interpolation [54] is performed again, the result of which is then used for FFT feature extraction. A five metre stretch of an asphalt road, concrete road, grass road, and gravel road are shown in Figures 3.5 (a), (b), (c), and (d), respectively. They are all calculated from the acceleration data collected at the approximate speed of 20 *km/h*.



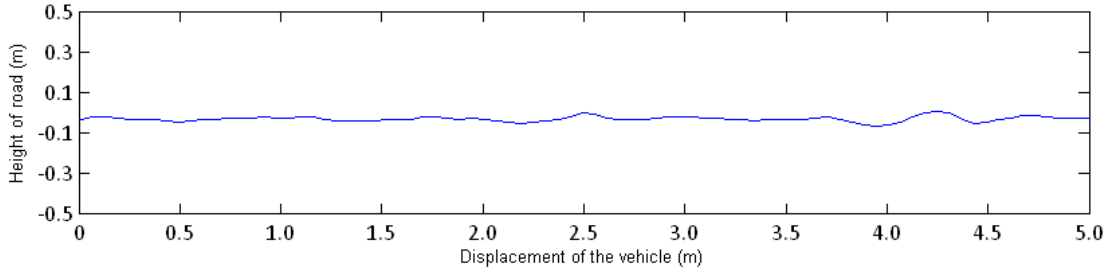
(a) A segment of asphalt road profile



(b) A segment of concrete road profile



(c) A segment of grass road profile



(d) A segment of gravel road profile

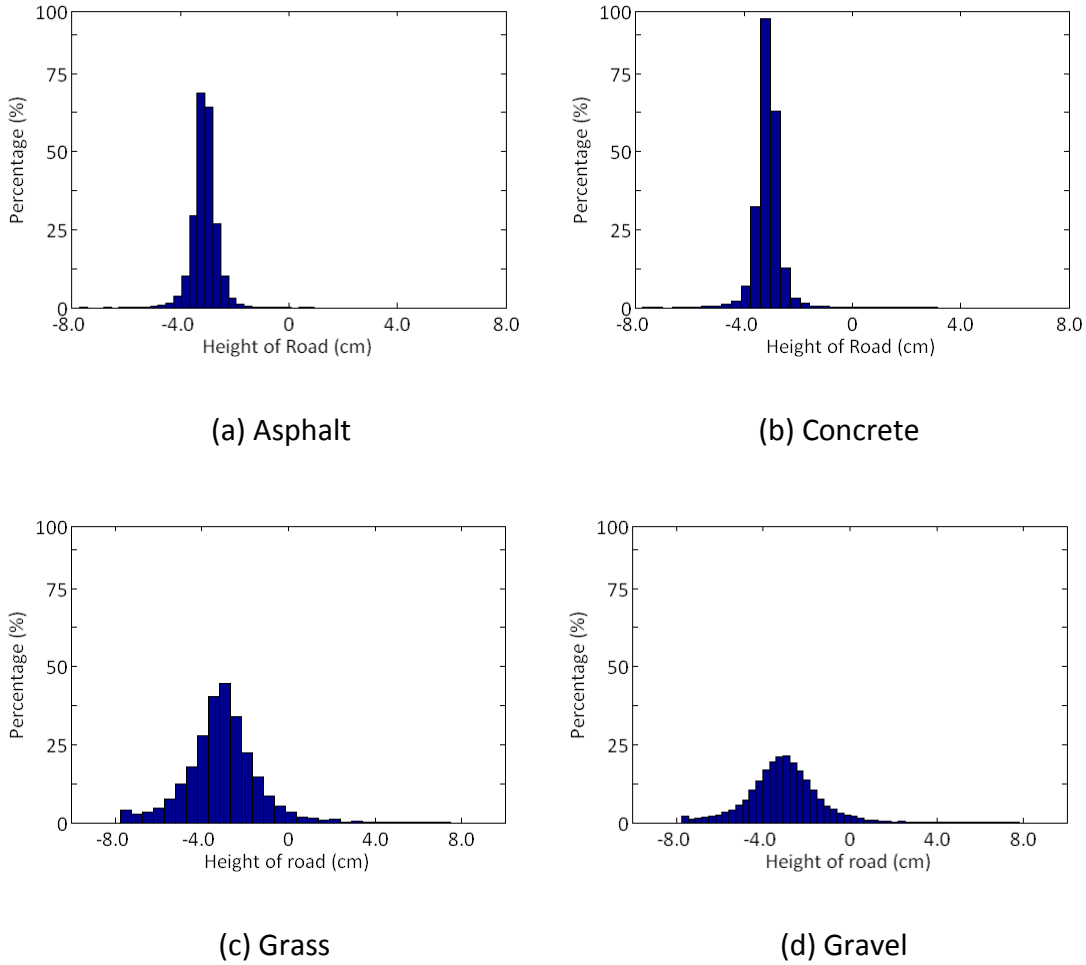
Figure 3.5 Profiles of four road terrain types

As can be seen in Figure 3.5 (a) and (b), there is not much apparent visual difference between an asphalt road and a concrete road. The highest amplitude among those four curves appears in the gravel road profile, while the amplitude of the curve in the grass road is lower than the gravel road but higher than the asphalt and concrete roads. More parameters describing the characteristics of these four road types are shown in Table 3.2. Note that data from the investigated roads were collected at the approximate speed of 20 *km/h* and cut off after travelling a distance of 500 *m.*

Table 3.2 Parameters of height of four road terrain types

Road terrain types	Min value (cm)	Max value (cm)	Amplitude (cm)	Standard Deviation (cm)	Mean value (cm)
Asphalt	-7.71	0.94	8.65	0.46	-3.11
Concrete	-7.75	3.15	10.90	0.48	-3.13
Grass	-7.76	7.49	15.25	1.67	-3.08
Gravel	-7.76	7.83	15.59	1.71	-3.08

As seen in Table 3.2, the terrain characteristics of each road type present diversities regarding many parameters in the spatial domain, such as maximum value, amplitude, and standard deviation. These indicate that the vibration condition of the vehicle is different when it is travelling on those different road terrains.



(a) Asphalt (b) Concrete

(c) Grass (d) Gravel

Figure 3.6 Histograms of four road terrain types

Further analysis is performed as follows. The start point of the vehicle is treated as zero value of the road height (on the vertical axis). With the movement of the vehicle, the road height can be calculated by (3.4) point by point. As shown in Figure 3.6, the histograms illustrate the distribution of the road height values of each road terrain. As can be seen from Figure 3.6 (a) and (b), an asphalt road terrain assembles around -3.1 *cm* scattering from -2.3 to -3.8 *cm,* whereas grass and gravel road terrains disperse more widely, as shown in Figure 3.6 (c) and (d). These indicate that vibrations of the latter two

types are more significant than the former two. Intuitively, different road terrain types should represent different spatial frequency characteristics, which make it possible to classify them by extracting spatial frequency features.

3.2 Features Extraction

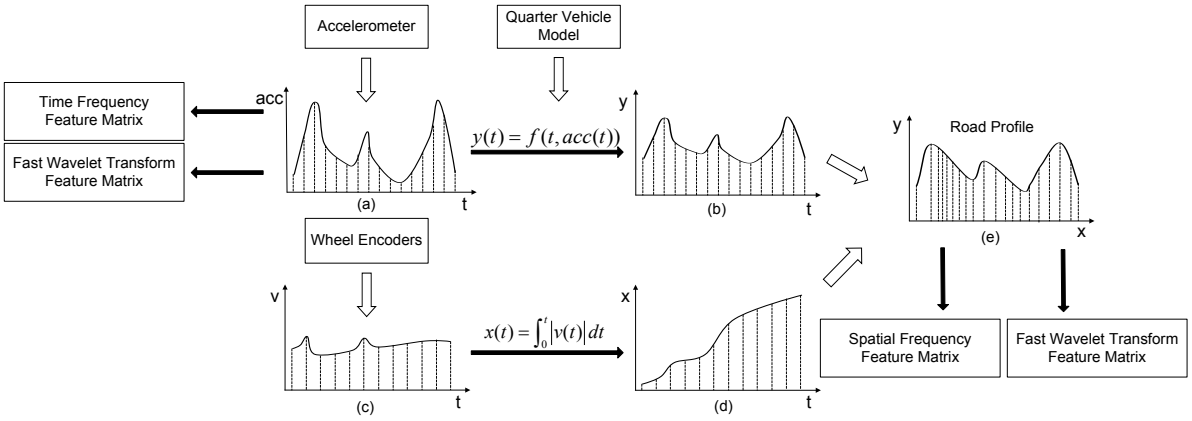


Figure 3.7 Illustration of feature extraction

Besides raw data of acc-t and y-x, as shown in Figure 3.7, another four features are also employed in tests. They are FFT features from *acc-t* (also called Time Frequency Feature, horizontal axis indicates time) and *y-x* (Spatial Frequency Feature, horizontal axis indicates distance), and Fast Wavelet Transform (FWT) features from *acc-t* and *y-x*. That is to say, there are six features are tested to find out the best feature for this case that tends to present highest classification accuracy.

3.2.1 FFT feature extracted from road profile (*y-x)*

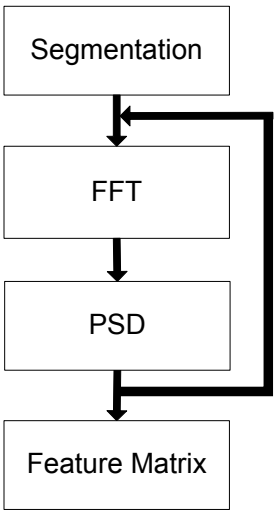


Figure 3.8 Procedure of forming the FFT feature matrix

As shown in Figure 3.8, FFT feature (spatial frequency feature) from *y-x* is extracted from the road profile data to form the feature matrix for machine learning.

a. Segmentation

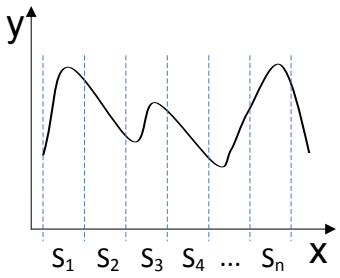


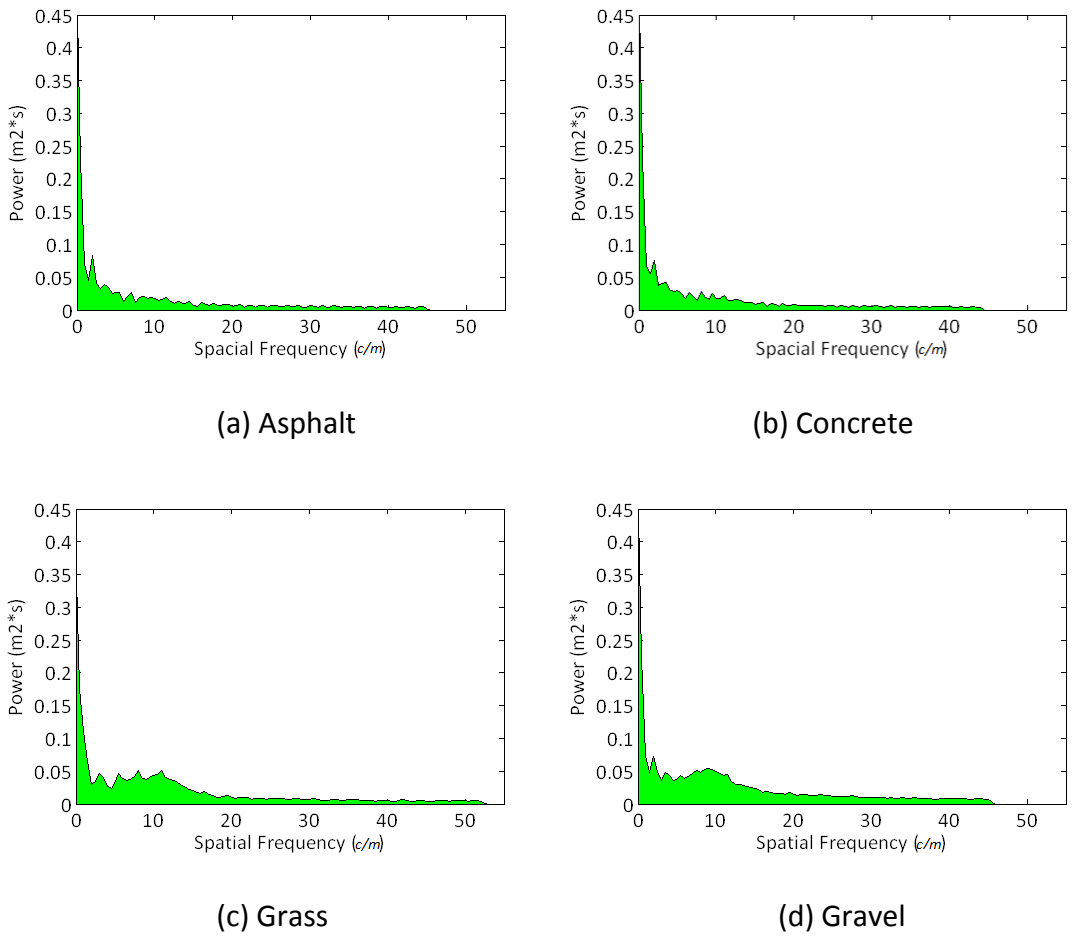
Figure 3.9 Road profile (*y-x*) – special frequency feature extraction

This road type detection system should report classification results frequently after the vehicle has travelled over a particular length of displacement. In this case it is 4 *m,* which is approximately the length of the vehicle. There is no particular crucial reason for it.

But it is considered a ppropriate for the application. As shown in Figure 3.9, the road profile is then divided into segments, marked as S_1, S_2 … S_n. In the off-line analysis stage, every 4 *m* long profile is processed as a sample, for either the training or testing phase.

The road profile data is segmented as the fixed length. Due to the variability in speed, although the *y-x* data is divided with different numbers, this does not affect FFT and PSD calculation results. And the number of samples only depends the distance that the vehicle has travelled.

b. Fast Fourier Transform



(a) Asphalt (b) Concrete

(c) Grass (d) Gravel

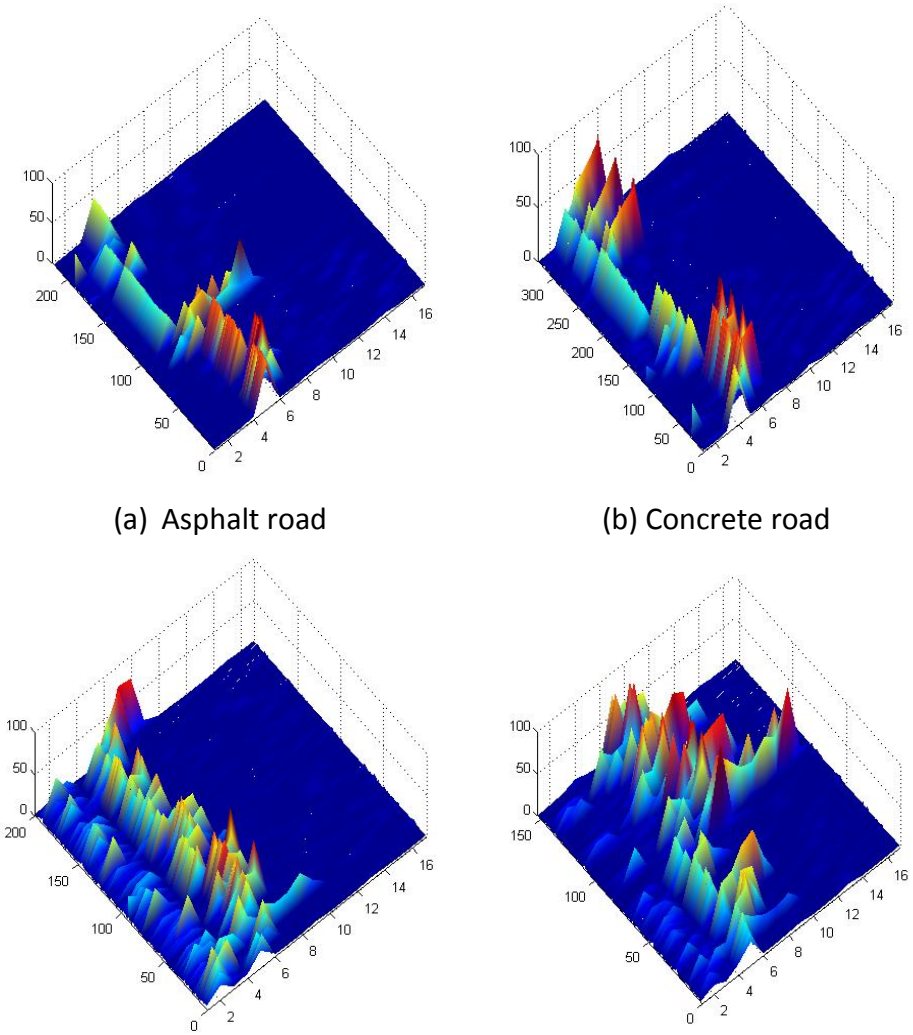
Figure 3.10 Spatial frequency distributions of four road types

Each segment is subsequently processed using FFT to form a feature vector one by one. For instance, a piece of each class of road profile (at approximately 20 *km/h*) is transformed into the spatial frequency domain. As can be seen in the figure 3.10, asphalt and concrete road profiles are very similar, even in the spatial frequency domain, which raises difficulties during road classification. As for grass and gravel road profile types, they have more distinguished characteristics, making them easier to differentiate from other road profile types. However, all four types contain low spatial frequency components which are around 0-2 *cycles/meter* (*c/m) (*caused by slight upslopes and/or downslopes of the road). On the other hand, high frequency components beyond 20 *c/m* (caused by the system noise) exit in each of the road types.

c. Power Spectral Density

To establish the spatial frequency characteristics, the PSD of each segment is calculated using Welch's method [55]. This requires the start, end and step of the frequency to be determined. It is not necessary to investigate very high frequency components that have little useful information for the classifier. The involved frequency components should be the ones that can be distinguished from each other but not up to infinite frequency. The step frequency chosen should not be immoderate small as this increases the computational burden. A value that results in just sufficient detail for classification is desired. After many prior tests, those are empirically determined to be 0 *c/m* to 150 *c/m* with 0.5 *c/m* interval. The smaller and larger intervals are tested very attentively, when fix other parameters. The tests show that immoderate small intervals increase the computational burden while immoderate large intervals reduce the classification accuracy. This appropriate value of interval leads to better classification

accuracy. Samples with the same number of features are thereby obtained one by one by implementing the loop as illustrated in Figure 3.8. The data acquired in four different road types with previously given speeds are processed to generate road profile features. FFT is used for this purpose.



(a) Asphalt road (b) Concrete road

(c) Grass road (d) Gravel road

Figure 3.11 Spatial frequency domain of four road types

As shown in Figure 3.11, *x* axis is the spatial frequency, *y* axis is the index number of samples, and *z* axis is the power between every 0.5 *c/m* of spatial frequency. Since there are not many non-zero value on *the z axis*, only 0-16 *c/m* of each class is presented

in Figure 3.11, though 0-150 *c/m* spatial frequency is analysed. Moreover, there are approximately 220 samples of asphalt road, 330 samples of concrete road, 200 samples of grass road and 150 samples of gravel road. Half of each of them is used in the training process while the other half is used for the testing process. As for *z* axis, different road type scatters around different range values on the flat of *x*-*y*. Referring to Figure 3.11 (a) and (b), high *z* axis values amass between 2 *c/m* to 6 *c/m* for both asphalt and concrete road values causing it is difficult to segregate. As shown in Figure 3.11 (c), gravel terrain has a completely different power distribution, which makes it easier to separate from the other terrain types. Through sheer intuition distinguishing asphalt roads from concrete roads is much harder than distinguishing the other two classes using the classifier.

d. Feature Matrix

$$F_{pro} = \begin{bmatrix} F_{s_1,f_1} & \cdots & F_{s_n,f_1} \\ \vdots & \ddots & \vdots \\ F_{s_1,f_m} & \cdots & F_{s_n,f_m} \end{bmatrix} \tag{3.7}$$

The PSD of each sample, which defines the feature vector, is then arranged column by column to form the feature matrix F_{pro}, as seen in (3.7). In this matrix, every column refers to a feature vector (a segment of road profile) while every row represents the features (PSD) extracted using the above procedure. In this case, F_{s_1,f_1} presents the PSD from 0 *c/m* to 0.5 *c/m* of the 1st segment, while F_{s_n,f_1} presents the PSD from 0 *c/m* to 0.5 *c/m* of the nth segment, and F_{s_n,f_m} presents the PSD from 149.5 *c/m* to the end 150 *c/m* of the nth segment. Overall, F_{pro} (3.7) contains *n* samples while each sample has *m* features.

Each value of the elements in the matrix (3.7) indicates a power at a particular investigated frequency.

3.2.2 FFT feature extracted from acceleration (*acc-t)*

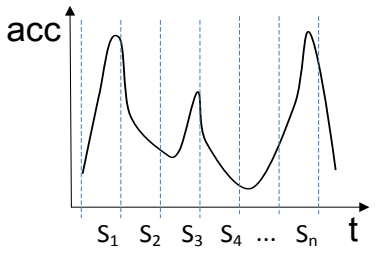


Figure 3.12 Acceleration (*acc-t*) – time frequency feature extraction

The segmentation and feature extraction of acceleration data are similar to that of road profile data, as shown in Figure 3.12. In order to compare the extracted features from *y-x* and *acc-x* equitably, the segments should be of the same 'length'. This is achieved by segmenting the data based on the time index.

$$T_{acc} = \begin{bmatrix} T_{s_1,f_1} & \cdots & T_{s_n,f_1} \\ \vdots & \ddots & \vdots \\ T_{s_1,f_m} & \cdots & T_{s_n,f_m} \end{bmatrix} \tag{3.8}$$

Thus, the feature matrix of *acc-t*, T_{acc} (3.8) has the same number of samples *n* and features *m*. In terms of the PSD calculation, the start frequency, end frequency and step are also the same, however, different powers in different frequencies are produced.

3.2.3 Fast wavelet transform feature extracted from acceleration (*acc-t)* and road profile (*y-x)*

Due to the importance of the power distribution of the different frequencies in this problem, there is then a need for a method that can localize important information from a time-frequency point of view. For this the wavelet transform is selected [56]. A wavelet function is a fundamental building block function, analogous to the trigonometric sine and cosine functions used in FFT. It simply divides the signal of interest into different frequency components, whereby each component can be studied at a resolution matched to its scale [57]. The wavelet transform has the important and useful ability of detecting and recognizing stationary, nonstationary, or transitory characteristics of signals including abrupt changes, spikes, drifts, and trends. In its implementation, the FWT leads to dyadic pyramidal implementation using filter banks and the corresponding Mallat algorithm [58] [59] is utilized.

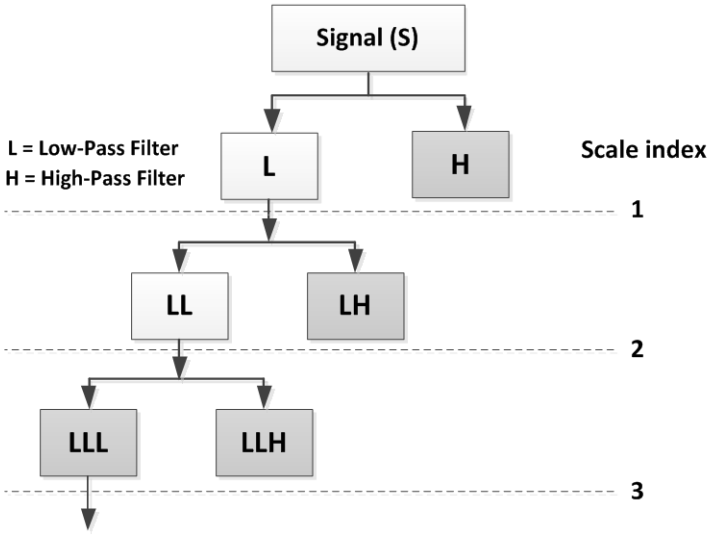


Figure 3.13 Acceleration/road profile data – FWT feature extraction

As shown in Figure 3.13, FWT develops the two channel filter banks through which the signal is split into two subspaces, *L* and *H*, which are orthonormally complementary to each other, with *L* being the space that includes the low frequency information about the original signal and *H* including the high frequency information. The decomposition of the low frequency subspace *L* as shown in Figure 3.13 is repeated with three levels of decomposition. Feature vectors are computed in two simple steps. Firstly, each record of the collected data is decomposed using FWT (ninth decomposition level and Daubechies (db4) family of wavelets are chosen). The FWT algorithm can extract features from different frequency components. The features extracted from ninth decomposition level and Daubechies family of wavelets describe the details of the signal in many frequency levels appropriately. Secondly, features are then constructed by simply determining the variance of the wavelet coefficients at each of the terminal nodes of the wavelet tree (marked with dark grey).

3.3 Normalization

Normalizing the feature matrix is very important in machine learning. Normalization has two main advantages for classification. It avoids attributes in greater numeric ranges dominating those in smaller numeric ranges. It also avoids numerical difficulties during the calculation. According to Hsu et al's instruction [60], all feature matrices are subsequently scaled to the range of [0, 1]. It should be noted that it is obligatory for each feature matrix to be normalized into the same range before training and/or testing, particularly when fusing or comparing feature matrices extracted from different categories.

3.4 Principal Component Analysis

For the machine learning procedure, the number of features used could vary from several, dozens or even hundreds. However, not every feature vector contributes to forming a training model using the classifier. Some of the feature vectors could only confuse the classifier while others could present differences for the classifier. Therefore, it is an important step to choose the best features, which contributes to dimensionality reduction as well.

PCA [61] uses an orthogonal transformation to convert a set of observations of possible correlated variables into a set of values of linear uncorrelated variables. These are called principal components. PCA is an optimal dimensionality reduction technique in terms of capturing the variance of a feature matrix. It converts each chromatographic vector into a single point in a principal component space, essentially projecting the data onto a new set of orthogonal axes that are sorted according to the amount of variance captured. The number of principal components, which depends on the demand of the specific application, can be less than or equal to the number of original variables.

3.5 K-Fold Cross Validation

In a machine learning procedure, the selection of training data among all off-line data is very important. The model trained by the classifier has a great deal to do with the training data which contains the feature information that is to be processed. The model is learned and formed from the training data provided. Although the model can fit the training data extremely well, it can probably fail to perform well with other data [62]. That is to say, if the selected training data does not present all feature information that the testing data has, this will certainly adversely affect the classifying results. Therefore, the

manner of selecting the appropriate training data set and leaving the rest for testing would significantly affect the classification accuracy.

To get more robust and reliable classification results, a method for selecting training and testing data set, the K-fold cross validation, is then employed [63]. This method divides available data into K folds. The number of K here is decided by the user in different conditions. Next, the data of one fold is then processed as testing data while the other K-fold is processed as training data. This procedure then will implement K times unless all the data of each fold is processed as testing data. An average value of these K classification results is then calculated as the final, accuracy of the classifier.

As illustrated in [63], set K as 5 is an optimized value that always leads to stabilized result. In this case, the parameter is set as five-fold cross validation to assess the mean performance of whole methodology using available data.

3.6 Alternative Classifiers

A freely available, highly accepted machine learning tool kit, WEKA, [64] is used to implement classification. WEKA has data pre-processing, classification, clustering, attribute selection, and data visualization. WEKA also includes support for association rule mining, comparing classifiers, data set generation, facilities for annotated documentation generation for source code, distribution estimation, and data conversion. The main reasons to apply WEKA are its Matlab interface and that it is very convenient to change classifiers for implementation effect evaluation.

A number of classifiers for the purposes of this thesis are chosen: Neural Network classifier, Naïve Bayes classifier, and SVM. Each of the classifiers is tested in off-line experiments to find out the most suitable one.

3.6.1 Naïve Bayes Classifier

With known feature vectors of samples, a Naïve Bayes Classifier is trained to estimate the probability of each class. It provides the fundamental probability model for well-known classification procedures, such as the statistical discriminate analysis. The probability is estimated by Bayes' Theorem [66]:

$$P(C = c_k \mid \vec{x}) = \frac{P(\vec{x} \mid C = c_k)P(C = c_k)}{P(\vec{x})} \tag{3.9}$$

where C is a random variable whose values are the classes, and $\vec{x}$ is vector of features. $P(C = c_k \mid \vec{x})$ is the conditional probability which is unable to be calculated without simplifying assumptions. According to Good's instruction [65], it is assumed that the features $x_1, x_2, ..., x_n$ are conditionally independent with known class label variable *C*. This leads to a simplified calculation:

$$P(\vec{x} \mid C = c_k) = \prod_i P(x_i \mid C = c_k) \tag{3.10}$$

The advantage of the assumption above is that $P(x_i \mid C = c_k)$ it is able to be modelled with relatively few parameters. Then based on this assumption, (3.9) can be transformed [66] as (3.11) in order to estimate $P(C = c_k \mid \vec{x})$.

$$P(c_k \mid x) = \frac{P(c_k) \times \prod_{j=1}^{d} P(x_j \mid c_k)}{P(x)} \tag{3.11}$$

This expression then becomes the probability estimation for classification. The purpose of minimizing the number of errors is to obtain as high value of $P(c_k \mid x)$ as

possible. Normally, finding maximum value of the numerator of (3.11) is performed to make a classification decision instead of calculating the value of $P(c_k|x)$.

3.6.2 Neural Network Classifier

Neural Network Classifier is a network of units. It provides a direct estimation of the posterior probabilities. Its input units present terms while the output units present the category or categories of interest. The weights on the edges which connect units present dependence relations. Part of training a Neural Network classifier is a process of estimating the parameters of the network while eliminating or minimizing errors.

A Neural Network for a classification task can be treated as a mapping function *F*:

$$R^d \rightarrow R^M \tag{3.12}$$

where *d*-dimensional input *x* is submitted into the network and *M*-vectored network output *y* is obtained to make the classification decision. Typically, the network is established to minimize the Mean Squared Errors (MSE) of the overall error measurement. In least squares estimation theory, the mapping function $F: x \rightarrow y$ that minimizes the expected squared error [67] is:

$$E\left[y - F(x)\right]^2 \tag{3.13}$$

That is the conditional expectation of *y* given *x*:

$$F(x) = E\left[y \mid x\right] \tag{3.14}$$

In the classification task, the output *y* is a binary value vector and it is the *j*th basis vector $e_j = (0,...,0,1,0,...,0)^t$ if $x \in j$. Then the *j*th element of *F(x)* is obtained*:*

$$
\begin{aligned}
F_j(x) &= E\left[y_j \mid x \right] \\
&= 1 \cdot P(y_i = 1 \mid x) + 0 \cdot P(y_i = 0 \mid x) \\
&= P(y_i = 1 \mid x) \\
&= P(\omega_j \mid x)
\end{aligned}
\tag{3.15}
$$

That is to say, the least squares estimate of the mapping function for a classification task is just the posterior probability.

Then the mean squared error function (3.13) can be derived [68]:

$$
MSE = \sum_{j=1}^{M} \int_{R^d} \left[F_j(x) - P(\omega_j \mid x) \right]^2 f(x) dx + \sum_{j=1}^{M} \int_{R^d} P(\omega_j \mid x)(1 - P(\omega_j \mid x)) f(x) dx \tag{3.16}
$$

Theoretically speaking, Neural Network is able to approximate any function arbitrarily and precisely. But because of the local minima problem, suboptimal network architecture and the finite sample data, the mapping function presented by the network does not perform perfectly. Therefore, Neural Network can provide estimates of posterior probabilities. The output of Neural Network is the least square estimate of the Bayesian a *posteriori* probability. Due to the binary output characteristic, the cross entropy function can be the more appropriate criterion than the squared error cost function in training Neural Network classifier for classification. To overcome this problem, improved and optimal Bayes rule can be performed by training Neural Networks with a generalized mean-squared error objective function [69].

3.6.3 Support Vector Machines Classifier

SVM was originally proposed by Vapnik [70] and has become popular since then. SVM has strong theoretical foundations and excellent empirical successes. SVM shows remarkable robust performance with successively extended. It has been improved by a number of other researchers, especially by Cristianini and Shawe-Taylor [71].

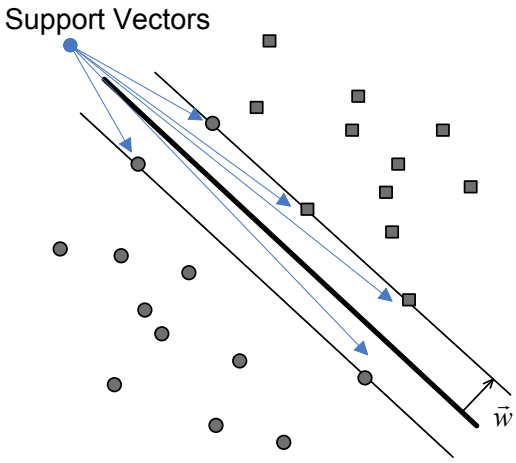


Figure 3.14 Linear SVM

Consider a set of binary labelled training data with a hyper-plane which separate the training data by a maximal margin, as shown in Figure 3.14. In cases that there is no possible linear separation, a technique called 'kernels' is introduced to automatically realize a non-linear mapping to a feature space. The hyper-plane decided by the SVM in feature space corresponds to a non-linear decision boundary in the input space.

Let the j^{th} input $x^j = (x_1^j, ..., x_n^j)$ be vectors in an *n*-dimensional space. Each vector x^j is set either of two classes identified by the label $Y^j \in \{-1, +1\}$. Let Φ: $I \subseteq R^n \rightarrow F \subseteq R^n$ be a mapping from the input space $I \subseteq R^n$ to a feature space *F*. It is assumed that a

sample of *S* of *m* labelled data: $S = \left\{(x^1, y^1), ..., (x^m, y^m)\right\}$. The SVM then finds a hyper-plane (*w, b*) that is evaluated by:

$$\gamma = \min_i y^i \left\{\left\langle w, \Phi(x^i)\right\rangle - b\right\} \tag{3.17}$$

where *w* has the same dimension with *F*, $\|w\|$ is a constant, and b is a real number. When (3.17) is maximized, *γ* is defined as *margin*. The distance between x^j and the decision boundary can be calculated by $\left(\left\langle w, \Phi(x^i)\right\rangle - b\right)$. As been set, the value of y^j is either (*-1*) or (*+1*), $\left(\left\langle w, \Phi(x^i)\right\rangle - b\right)$ times y^j gives a positive value for all correct classifications and a negative value for the incorrect ones. If the data is linearly x^j separable, the minimum of this quantity over all x^j will be positive, which is called margin. When input a new *x* to classify, a label will be decided according to its relationship with the decision boundary. The related function is:

$$f(x) = sign\left(\left\langle w, \Phi(x^i)\right\rangle - b\right) \tag{3.18}$$

It is proved [71] that for the maximal margin hyper-plane:

$$w = \sum_{i=1}^{m} \alpha_i y^i \Phi(x^i) \tag{3.19}$$

where α^i are positive real numbers which maximize:

$$\sum_{i=1}^{m} \alpha_i - \sum_{i,j=1}^{m} \alpha_i \alpha_j y^i y^j \left\langle \Phi(x^i), \Phi(x^j)\right\rangle \tag{3.20}$$

subject to:

$$\sum_{i=1}^{m} \alpha_i y^i = 0, \quad \alpha > 0 \tag{3.21}$$

then (3.18) can be transferred as:

$$f(x) = sign\left(\sum_{i=1}^{m} \alpha_i y_i \left\langle \Phi(x^i), \Phi(x^j) \right\rangle - b \right) \tag{3.22}$$

It can be seen in (3.22) that α_i is associated with the training x^j. A remarkable property of this alternative representation is that only a subset of x^j will be associated with non-zero α_i. These points are defined as *support vectors*, shown in Figure 3.14. Those support vectors lie closest to the hyper-plane.

After training the SVMs classifier, an optimal hyper-plane that has a maximal margin to separate two classes can then be fitted. By calculating the spatial distance between the vectors and the hyper-plane, it is able to produce posterior probabilities that are directly comparable between classes.

3.7 Experiment

3.7.1 Experiment Platform

The experimental test bed CRUISE (Figure 3.15), developed in-house, is employed for experimentation. It has capabilities of sensing and computing that are used in data collection. CRUISE is equipped with a range of sensors including an accelerometer, LRFs, a camera, a GPS unit, an Inertial IMU, a millimetre wave radar, and four wheel encoders, among others. Two PC104 computers mounted in the back tray and a laptop, all

connected via an Ethernet switch, are used for data collection, logging, and computing. A separate battery bank provides the required power.

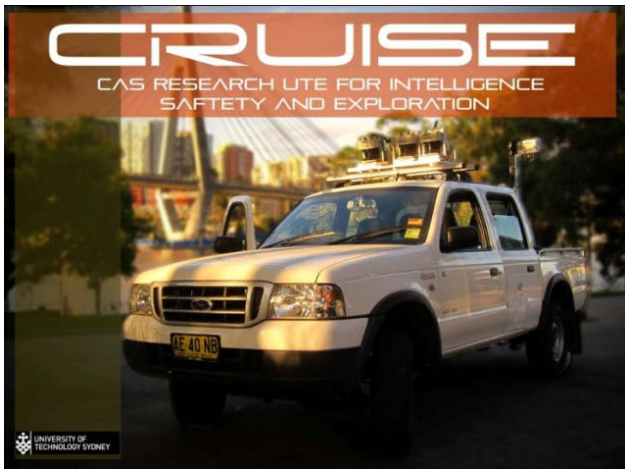


Figure 3.15 Research platform: CRUISE

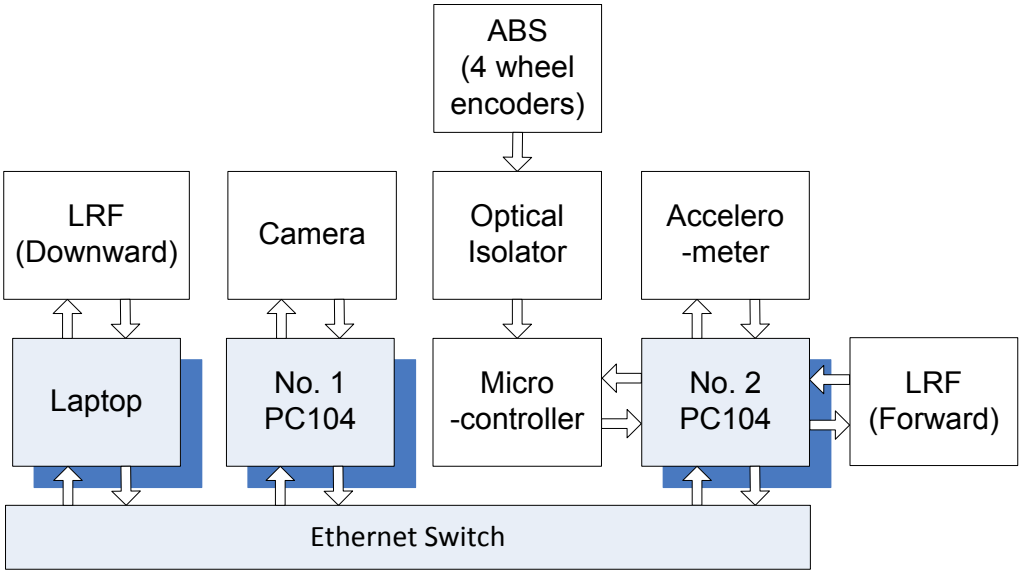


Figure 3.16 The hardware structure of the system

As shown in Figure 3.16, three computers communicate with each other via Ethernet. The laptop located in the driving cab is used to operate the other two PC104 computers remotely for logging data and computation. The laptop is also connected with the downward LRF transmits out configuration commands and receives measurement data.

Since magnanimity image data coming from the camera while working, a slow computer, No. 1 computer has been distributed to log the image data only but no other sensors (seen from Figure 3.16). The camera is configured as 30 Frame per Second *(FPS)* with 800×600 resolution. Additionally, both of the PC104 computer hard drive disks were taken place by solid state drives to avoid damage by extreme intense vibration.

The rest of sensors that wheel encoders, accelerometer, and the forward LRF are all connected with No. 2 computer (seen from Figure 3.16). The digital signal of the accelerometer is transmitted via USB cable to the No. 2 computer. And the sampling frequency of it is set at 50 *Hz*. As indicated in Figure 3.17, the accelerometer is mounted on the right rear suspension, closed to the right rear wheel.

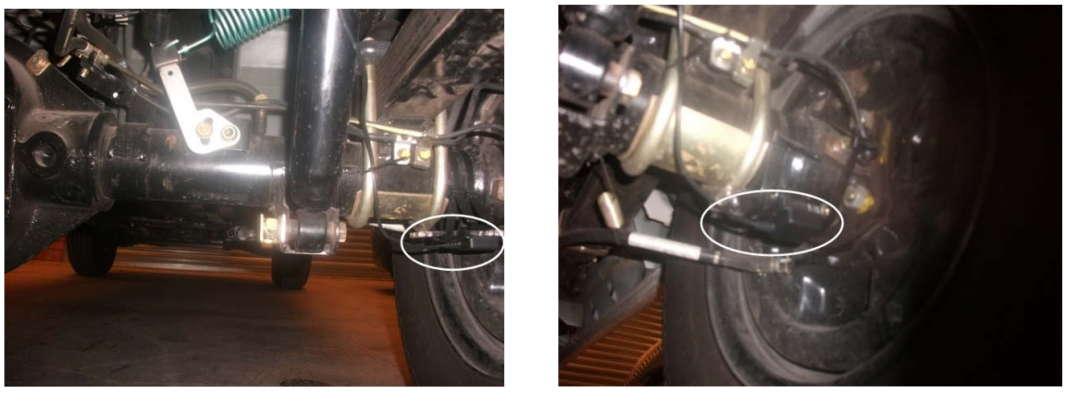


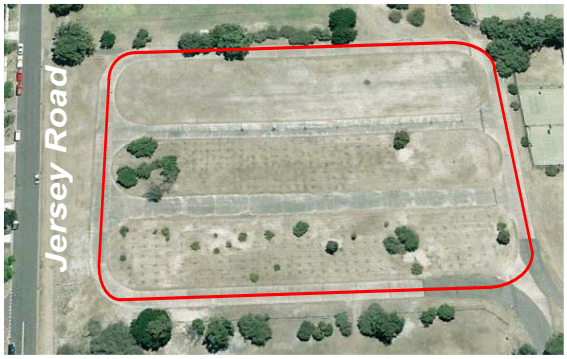
Figure 3.17 The mounted accelerometer on the suspension of CRUISE

The four wheel encoders in ABS are tapped through optical isolators and amplifiers, interfaced with a PIC microcontroller for digitizing. The microcontroller type chosen is a *Microchip*® PIC18F4431 product. Three external interruptions and an interrupt-on-change interruption are employed to detect four sine-like signals from four-wheel encoders. Each interrupt program works as a counter while the main program sends out counted numbers regularly (per 0.1 *second*). The microcontroller is connected to Computer No. 2 via RS-232 serial port with the sampling frequency of 10 *Hz* so that the data synchronization can be achieved effectively.

3.7.2 Acceleration Based Experiments

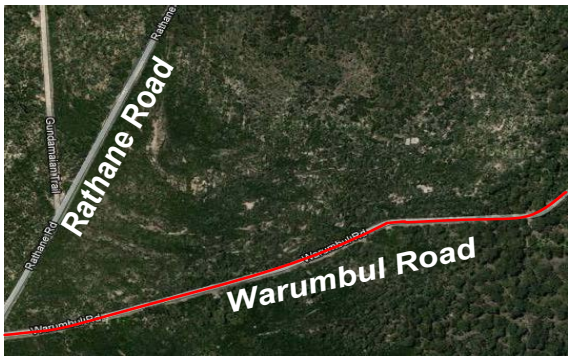
The experiments were carried out during fine days with average summer temperature and humidity in the urban area of Sydney, Australia. The data from each of the sensors is logged and recorded properly with timestamp for off-line analysis and processing. Four road terrain types were chosen for the project, with each of the four roads owning typical terrain characteristics. At least two operators were required for the experiment, one to drive while the other to attend to data logging.

(a) Asphalt route



(b) Concrete route

(c) Grass route



(d) Gravel route

Figure 3.18 Data Collection Routines of Four Road Types

CRUISE's driving routes in the four road types during the data collection experiment are shown in Figure 3.18. Continuous data logging ensued in elliptical roads with few other running vehicles. The grass road is more challenging to find. Having found only a one hundred metre stretch of a grass road, CRUISE then had to travel on this short road over several times with intensive vibration. Difficulties were also encountered in the search for a gravel road. A one kilometre long gravel road is eventually found in the Royal National Park. The location of each data collection place is shown in Table 3.3.

Table 3.3 Locations of the data collection experiments

Road Type	Location
Asphalt Road	Park Parade, Pagewood, NSW 2035, Australia
Concrete Road	Jersey Road, Maroubra, NSW 2036, Australia
Grass Road	Anzac Parade, Maroubra, NSW 2035, Australia
Gravel Road	Warumbul Road, Royal National Park, NSW 2230, Australia

For the purpose of investigating speed independency, the CRUISE is driven on four types of roads at different speeds. Considering driving safety and practical requirements, the data is collected on asphalt and concrete roads at speeds of 20, 30, and 40 *km/h*; grass roads at speeds of 20 and 30 *km/h*; and gravel roads at a speed of 20 *km/h*, approximately. Critical vibration is felt by the passengers while the CRUISE is running on grass roads over 20*km/h*, something neither comfortable nor safe for the driver and the vehicle with all the attached equipment.

Apart from the places described above, the experiment also involved more public roads. Since the speed and routines were not restricted, the driving circumstance contained many turns, starts, and stops. Moreover, the driving is over a variety of asphalt road surfaces including highways, tunnels, and suburb distributor roads. Finally, the experiments were implemented using the accelerometer, camera, and LRFs and a combination of using all the aforementioned tools for the MRF algorithm described in Chapter VI.

3.8 Experiment Results

Before the acceleration based road terrain classification experiment, a number of data sets were tested to investigate the optimized feature and classifier, the effect of employing PCA, and speed independency.

3.8.1 Feature Selection

The features considered in this study are the raw data, FWT features and FFT features. In order to know which features can produce better classification accuracies in different scenarios, representational data set combinations are tested and investigated.

The five-fold cross validation is also employed to get more robust results. The SVM is then selected to be the classifier for this experiment. Three data sets are chosen to be tested. Each of the data sets has different numbers of road type and speed. As shown in Table 3.4, the first row presents the classification results of 2 road types (asphalt and concrete) at the speed of 40 *km/h*; the second row presents the results of three road types (asphalt, concrete and grass road) at 30 *km/h;* while the third row presents the results of all four road types (asphalt, concrete, grass and gravel road) at 40 *km/h*. In addition, the results in the first three rows are all garnered by using acceleration data (*acc*-*t*) while the last three row results are garnered using road profile data (*y*-*x*).

Table 3.4 Classification with different feature selection

	Training and Testing Terrain Types @ Speed (*km/h*)	***Raw data***	***FWT feature***	***FFT feature***
acc-t	2 types @ 40	83.4%	94.2%	71.9%
	3 types @ 30	81.2%	80.5%	85.2%
	4 types @ 20	79.1%	70.6%	95.6%
y-x	2 types @ 40	75.8%	95.5%	98.5%
	3 types @ 30	62.7%	90.4%	97.3%
	4 types @ 20	58.8%	71.8%	99.3%

Using raw data as features provide a baseline for comparisons. While using raw acceleration data, the classification results are comparable with those of the FWT and FFT features. However, road profile raw data does not seem to provide results comparable with those of the FWT and FFT. This could be due to the irregularity of the data points in the *y-x* space, which is caused by variations in the vehicle speed. The raw data is used without interpolation between points. The FWT features did not show a significant difference with using acceleration raw data or profile data. On the other hand FFT features extracted from profile data shows higher accuracies than that of those extracted from acceleration data.

Consequently, the FFT feature extracted from road profile data (spatial frequency feature) presents the best performance. This test result responds to our hypothesis that

the road profile of each road type has unique and exclusive characteristics. And among those exclusive characteristics, it is the optimized spatial frequency feature that always presents better classification results in different scenarios.

3.8.2 Speed Dependency

In the following experiments, the FWT features and FFT features are applied to make comparisons. Like in the previous test, the five-fold cross validation and SVM are employed in this experiment. The intention of this experiment is to test the speed dependence in road type classification. It is to be noted that on the asphalt road (Type 1 in Table 3.5) and concrete road (Type 2), data is collected in three speeds, whereas data is collected in two speeds on the grass road (Type 3). These different speeds on the particular road types are treated as belonging to separate classes. The classification accuracies were analysed to verify the speed independency.

Table 3.5 Classification of same road type with different speeds as different classes

	Training and Testing Terrain Types @ Speed (*km/h*)	***FWT feature***	***FFT feature***
acc-t	type 1 @ 20 30 40	80.1%	63.2%
	type 2 @ 20 30 40	90.1%	93.1%
	type 3 @ 20 30	88.6%	82.6%
y-x	type 1 @ 20 30 40	72.4%	48.6%
	type 2 @ 20 30 40	90.9%	86.9%
	type 3 @ 20 30	85.0%	65.3%

In this case higher classification accuracy refers to higher discrimination between speeds (of the same road type) and consequently higher speed dependence. On the other hand lower classification accuracies mean lower speed dependence. It can be observed from Table 3.5 that on average FWT features produce slightly better classification results with acceleration data compared to road profile data. However, the differences are not that significant, which means using road profile data does not significantly contribute to speed independency while using FWT features. On the other hand, FFT features produce lower accuracies than those of FWT features leading to better speed independency. Furthermore, the accuracies with FFT on acceleration data is on average higher than that of road profile data. This means that the road profile data provides better speed independency with FFT features.

Table 3.6 FFT feature on acceleration and road profile based classification accuracy training and testing with different speeds

Training Terrain Types @ Speed (*km/h*)	**Testing Terrain Types @ Speed (*km/h*)**	***FFT Feature on acc-t***	***FFT Feature on y-x***
2 types @ 20	2 types @ 30	89.4%	98.3%
2 types @ 40	2 types @ 20	66.2%	84.1%
2 types @ 40	2 types @ 30	49.7%	97.8%
2 types @ 30	2 types @ 40	28.2%	85.4%
2 types @ 20 30	2 types @ 40	29.6%	85.4%
2 types @ 20 40	2 types @ 30	94.4%	98.3%
2 types @30 40	2 types @20	99.2%	99.2%

The speed independency is tested again with a different concept. As can be seen in Table 3.6, in this scenario, the classifier is trained at one speed and tested at a different speed. Higher classification results mean higher speed independency and lower classification results mean higher speed dependency. In general FFT features extracted on profile data provided better classification accuracies than those extracted from acceleration data and hence exhibited more speed independency.

To further clarify this speed independency issue, anther experiment is designed to test the effect of speed. Since the spatial frequency feature presents better test results than other features, as shown in Table 3.7, the data from asphalt and gravel roads are used to train and test at different speeds employing spatial frequency feature, five-fold cross validation, and SVM classifier.

Table 3.7 FFT feature on acceleration and road profile based classification results training and testing with different speeds

Training Terrain Types @ Speed (*km/h*)	**Testing Terrain Types @ Speed (*km/h*)**	**FFT Feature on *y-x***
2 types @10	2 types @10	100.00%
2 types @20	2 types @20	98.33%
2 types @30	2 types @30	100.00%
2 types @10 20 30	2 types @10	100.00%
2 types @10 20 30	2 types @20	96.01%
2 types @10 20 30	2 types @30	100.00%
2 types @10 20 30	2 types @10 20	97.19%
2 types @10 20 30	2 types @10 30	100.00%
2 types @10 20 30	2 types @20 30	96.23%
2 types @10 20 30	2 types @10 20 30	97.40%

Firstly, the speeds of 10, 20, 30 *km/h* are used respectively in the training process and then did the testing at the same speeds. The results in the first three rows in Table 3.7 show extremely appealing classification accuracy. Furthermore, the data is trained at the speeds of 10, 20 and 30 *km/h* and tested in different combinations of 10, 20 and 30 *km/h* speeds. As shown in the last seven rows of Table 3.7, the outcomes still show perfect results. One of the reasons for these incredibly good classification accuracies is the application of five-fold cross validation that helps overcoming the lack of insufficient training data which does not contain all the testing data features. These results indicate that it is very helpful and necessary for the training process to include the specific speed(s) if the testing process has these on hand. Consequently, it gives out a hint that it would be better to train as more data including all types and speeds as can be collected before the testing experiment to produce more robust and reliable results.

3.8.3 Classifiers Selection

In order to select an appropriate classifier, the WEKA [64] tool kit is used to implement the experiments. WEKA is a very popular machine learning workbench which contains the implementation of algorithms for classification, clustering, and association rule mining, along with graphical user interfaces and visualization utilities for data exploration and algorithm evaluation. WEKA has good interfaces to other popular software, in this case Matlab is used, to do the off-line process and analysis.

One of the main features of WEKA is its classification methods have over 100 classifiers to select from. Those classifiers are divided into Bayesian methods, lazy methods, rule-based methods, tree learners, function-based learners, and miscellaneous methods. For this task, three classifiers: Naïve Bayes, Neural Network, and SVM, are

selected to make the comparison to find out the most appropriate one in terms of implementation time and accuracy rate.

For this experiment, the same Personal Computer is used:

CPU: Intel Pentium III Xeon processor

Memory: 4 GiB

System: Microsoft Windows XP Professional

with Matlab 2010a and WEKA installed.

The same data set is chosen which contains asphalt road at 20, 30 *km/h* speeds, gravel road at 20, 30 *km/h* speeds, grass road at 10, 20 *km/h* speeds, and gravel road at 10, 20 *km/h* speeds to implement the road type classification. The five-fold cross validation is employed as well. The feature matrices are extracted from road profiles to be the spatial frequency features.

Table 3.8 Comparison of different classifiers usage

Classifiers	Naïve Bayes	Neural Network	SVM
Implement Time	109 seconds	2335 seconds	120 seconds
Accuracy Rate	91.3%	99.4%	99.3%

The computing results in Table 3.8 shows Naïve Bayes and SVM classifiers present similar computing speeds that are much faster than the Neural Network classifier. The result of the Naïve Bayes classifier shows the lowest classification accuracy when comparing with SVM and Neural Network classifiers. Therefore, considering these experimental results, SVM is selected for the following experiments because of its relatively faster computing speed and higher accuracy.

3.8.4 Acceleration Based Experiment Result

All the parameters for this algorithm have thus far been identified. The acceleration based road terrain classification experiment can now be performed.

The data set employed is the collected acceleration data, which is on a variety of road terrain all around urban and rural areas with asphalt, concrete, grass and gravel roads. Besides the selected routes are shown in Figure 3.18, asphalt roads on the way back to the University of Technology Sydney (UTS): Sir Bertram Stevens Drive, Princes Highway, Broadway, etc. are also involved. All the data set used is collected as the vehicle is operated in normal circumstances, with the speed on each road gradually changing during acceleration, de-acceleration, and stopping.

Table 3.9 Acceleration data for training and testing

Road Terrain	Training			Testing		
	Speed (km/h)	Displacement (m)	Number of Samples	Speed (km/h)	Displacement (m)	Number of Samples
Asphalt	20, 30, 40, 50	328	82	0 ~ 70	3948	987
Concrete	20, 30	284	71	0 ~ 45	952	238
Grass	10	304	76	0 ~ 25	876	219
Gravel	20, 30	312	78	0 ~ 35	2712	678
Total		1228	307		8488	2122

In order to analyse on-line experimental conditions, as shown in Table 3.9, neither all of the data set nor five-fold cross validation is used to do the classification experiment. The data with relatively constant and definite speed is selected to be the training data. All the other data with continuous speed is then used in the testing stage. As it is difficult to find roads with all four transitions, the data is concatenated to reflect a combination road with the transitions, for example type 1 and type 2 in Figure 3.19 (a).

For balanced and appropriate training, as shown in Table 3.9, the numbers of samples per road type are approximately equal. Each class has approximately 70 to 80 samples to implement the training process. Due to the data collection experiment procedure, the data of each class is selected at particular and different speeds. They are expected to reflect all characteristics that other speed data might have. Nevertheless, the extracted feature is in the spatial frequency domain, which is least affected by speed. For the testing stage, there is no any speed selection. It is simulated to act as though the CRUISE is driven under regular circumstances all the way. So the speeds are continuous and various at valid ranges. As previously mentioned, the data of concrete and grass roads are hard to collect while asphalt road data is more easily gathered. All in all, the CRUISE travelled around 8.5 *km* to collect all the experimental data.

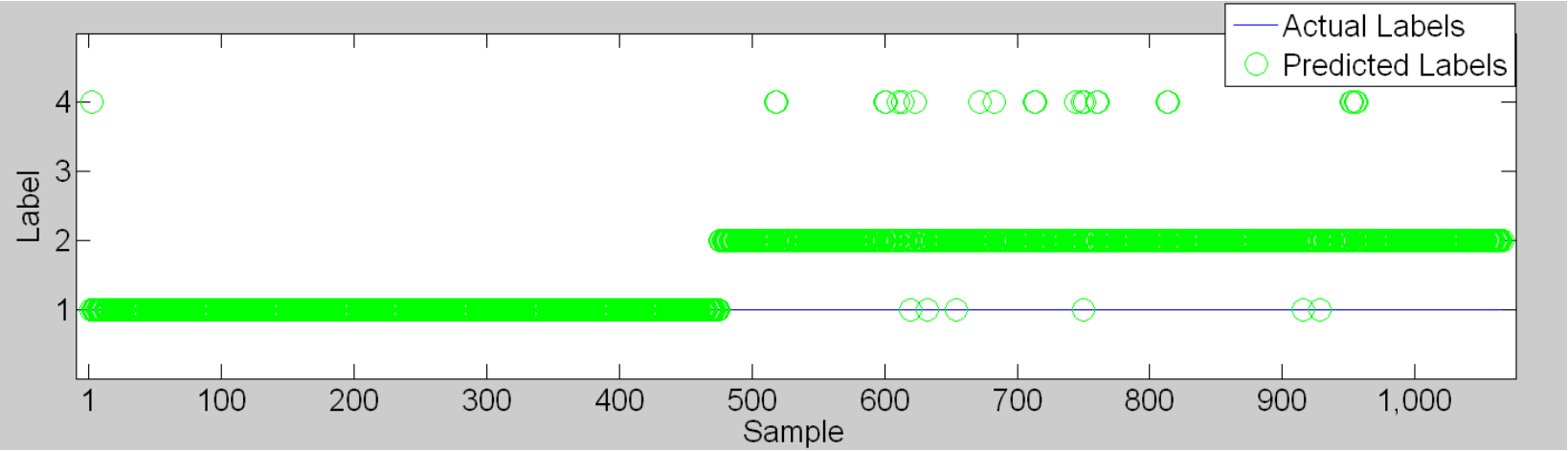


Figure 3.19 (a) Acceleration based experiment label sequence: segment 1

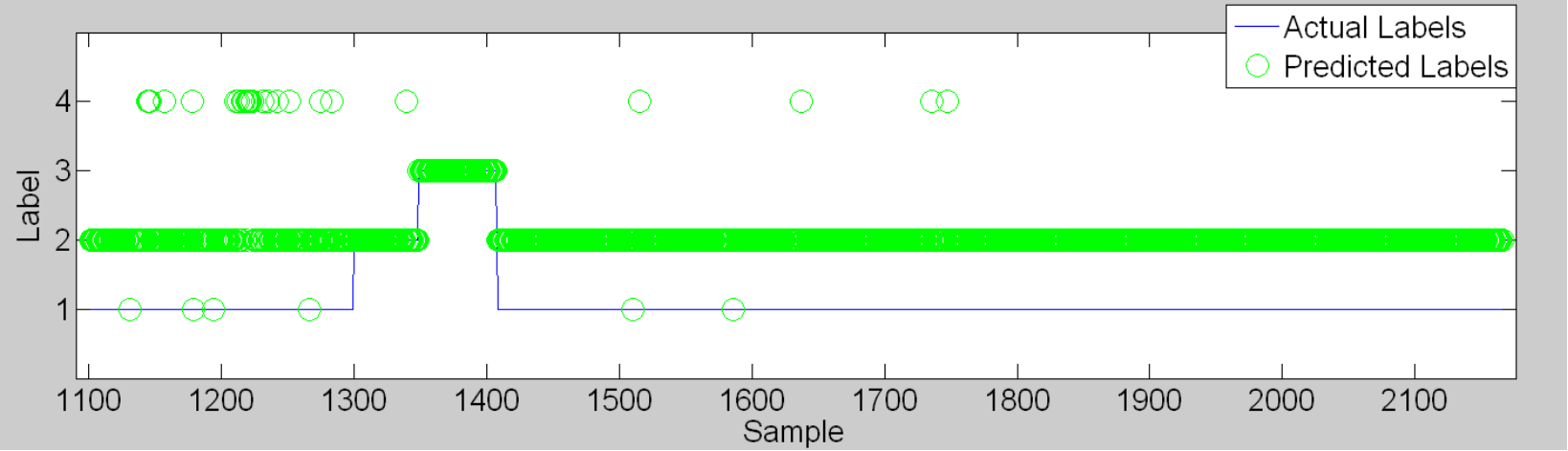


Figure 3.19 (b) Acceleration based experiment label sequence: segment 2

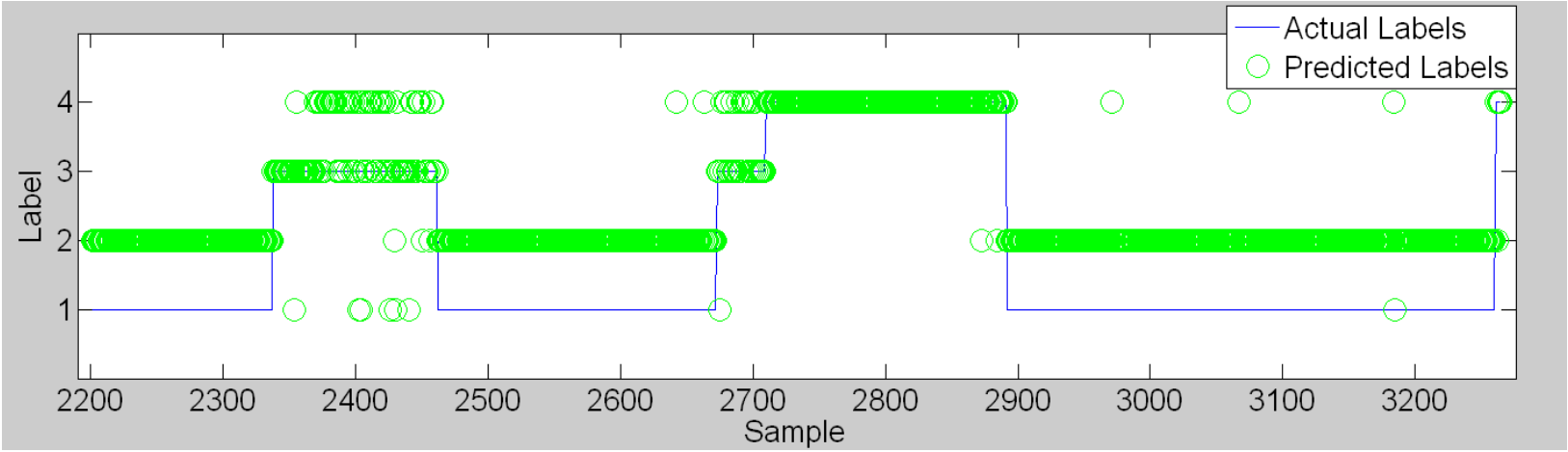


Figure 3.19 (c) Acceleration based experiment label sequence: segment 3

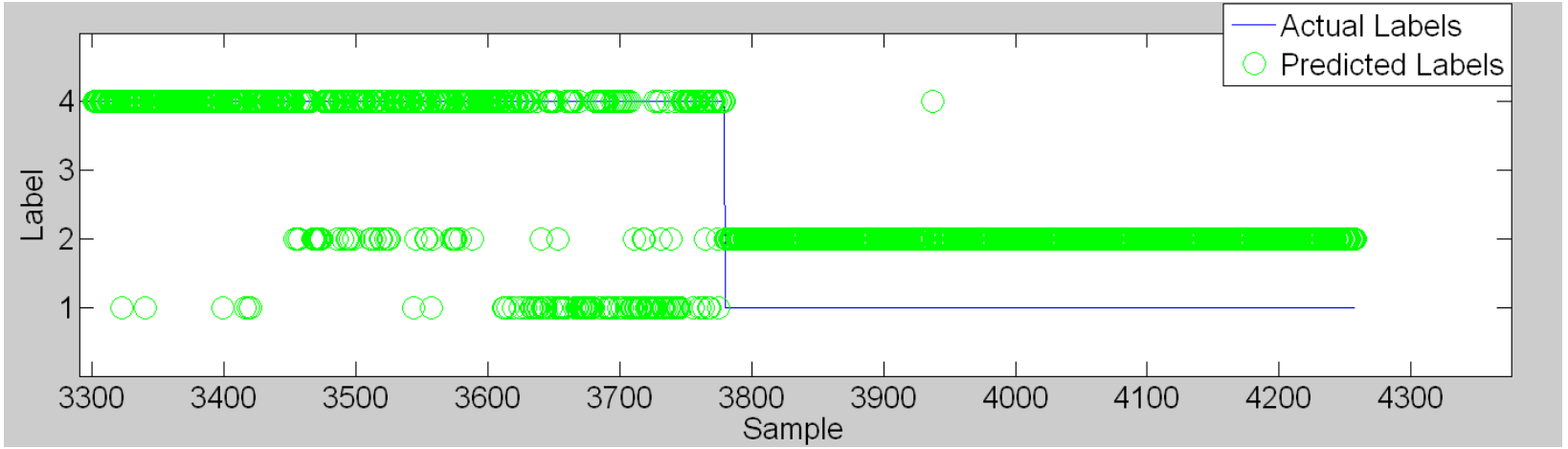


Figure 3.19 (d) Acceleration based experiment label sequence: segment 4

Figures 3.19 (a) to (d) show the acceleration based experiment label sequences. The horizontal axis is the index of the samples (road segments). There are four points (1, 2, 3, and 4) on the vertical axis that show the class labels. The predicted labels classified by SVM classifier are marked by green circles while the actual labels are represented by blue lines. All the testing acceleration data is arranged to act as the vehicle runs from one terrain type to another by random road segments. The classifier confuses most of the asphalt terrain with the concrete terrain starting at time index 480. This is mainly because from this point on, even though the vehicle runs on asphalt, the surface of this new asphalt road presents very similar vibration characteristics to that of a concrete road. The accelerometer data cannot be reliably to distinguish between asphalt and concrete terrain.

Table 3.10 The experimental results using acceleration data without PCA process

Training / Testing	Asphalt @20,30,40,50 *(km/h)*	Concrete @20,30 *(km/h)*	Grass @10 *(km/h)*	Gravel @20,30 *(km/h)*
Asphalt @0~70	150/987 15.2 %	821/987 83.2%	0/987 0%	16/987 1.6%
Concrete @0~45	0/238 0%	235/238 98.7%	0/238 0%	3/238 1.3%
Grass @0~25	7/219 3.2%	3/219 1.4%	154/219 70.3%	55/219 25.1%
Gravel @0~35	91/678 13.4%	48/678 7.1%	0/678 0%	539/678 79.5%

Average Classification Accuracy: 65.9%

Table 3.10 shows the statistics table of Figure 3.19 performance. Each of the rows shows how many samples are classified to be the specific class and the total number of that class. Then the classification accuracy can be calculated. Taking the first class of Asphalt road for instance, as can be seen in Table 3.10, there are 987 samples (each sample refers to 4 meters travel displacement for CRUISE) are tested. Only 150 samples which are 15.2% of the actual class were classified correctly. 83.2% of the Asphalt road is classified as Concrete road. In this way, Table 3.10 presents all classification accuracies. The Asphalt, Concrete, Grass and Gravel roads are classified correctly as 15.2%, 98.7%, 70.3% and 79.5%. The average accuracy of this method is 65.9%. Furthermore, it can be seen that the Asphalt road often confuse the classifier to be Concrete road. As explained above, this is due to the road profiles of these two types of terrain are similar to each other which puzzles the classifier during the training process. For the Grass and Gravel roads, there is 25.1% misclassification from Grass road to Gravel road, lower than 83.2% from Asphalt to Concrete road. This still reflects some similarities between Grass and Gravel roads.

Besides the classification results using above methodology, an extra method is tested as well. The testing feature vectors are computed using PCA process firstly before input into the classifier. As mentioned in section 3.4, PCA converts each chromatographic vector into a single point in principal component space, reduces the dimensionality of the feature matrix, and get the most useful variables. Dimensionality reduction is then carried out using PCA. Empirical evaluations lead to the selection of 28 dimensions. It is expected that the PCA would help to improve the classification accuracy.

Table 3.11 The experimental results using acceleration data with PCA process

Training / Testing	Asphalt @20,30,40,50 *(km/h)*	Concrete @20,30 *(km/h)*	Grass @10 *(km/h)*	Gravel @20,30 *(km/h)*
Asphalt @0~70	174/987 = 17.6%	680/987 = 68.9%	0/987 = 0%	133/987 = 13.5%
Concrete @0~45	0/238 = 0%	237/238 = 99.6%	1/238 = 0.4%	0/238 = 0%
Grass @0~25	5/219 = 2.3%	4/219 = 1.8%	164/219 = 74.9%	46/219 = 21.0%
Gravel @0~35	40/678 = 5.9%	43/678 = 6.3%	17/678 = 2.5%	578/678 = 85.3%

Average Classification Accuracy: 69.4%

As shown in Table 3.11, with the use of PCA, the classification accuracies of grass and gravel roads show an increase. But it does not make any significant contributions to the asphalt road classification accuracy. It can be seen that some of the former incorrect labels (shown in Table 3.10) on the concrete road has shifted to the grass and gravel roads. This indicates that the PCA process does not always help in specific scenarios during road terrain classification. Sometimes it leads to unreasonable results which are relatively obviously wrong. As for the grass and gravel terrains, the accuracy increase slightly.

3.9 Conclusion

In this chapter, a method to classify different road terrains has been presented. Due to the variety of road profiles that represent different but not exclusive spatial characteristics, the road types could be recognized to a certain extent using machine learning techniques.

The road profile was calculated by using the acceleration data, the vehicle's speed data, and the one quarter model of the vehicle. In the training stage, the spatial frequency features were extracted from the road profile to form the feature matrix for the SVM classifier. The prior off-line training and testing experiments were performed before the final acceleration based experiment to acquire the optimal parameters of the algorithm.

The acceleration based experimental results showed that the method worked fairly well on concrete, grass, and gravel roads but not asphalt road. Since the asphalt road is the most common road type in urban environments, it would be better if a method to classify asphalt road with better accuracy could be found. Furthermore, it should be noted that the classification accuracy also depended on whether there was sufficient data that was with different speeds available in the training stage. A limited range of speeds in the training data tended to result in lower classification accuracy.

In order to address the above issues, use of the camera image is investigated in the next chapter.

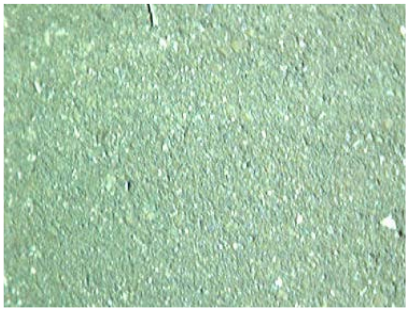
CHAPTER IV

IMAGE BASED ROAD TERRAIN CLASSIFICATION

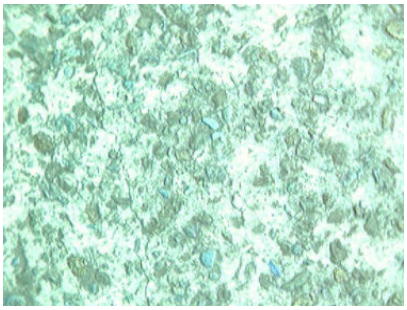
The images of different road terrain types present abundant information and significantly different characteristics. Hence, it is considered to use a camera to capture images of road surfaces in an attempt to overcome the issues found using accelerometer data only. By processing and extracting features from those images, road terrain classification is expected to be improved.

In this chapter, image based method to classify the road terrain types is presented. The texture features from image are firstly introduced. Then the imaged based experiment and the results are presented.

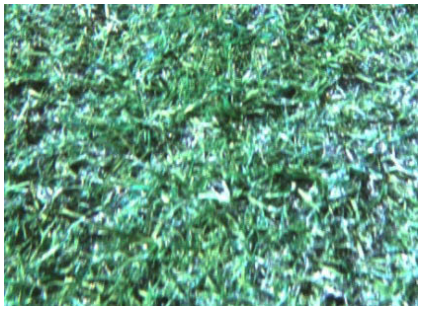
A camera is mounted on the top of a frame above the back tray of the CRUISE, looking downward to the ground. As indicated in Figure 3.16, the image data is transmitted to Computer No.1 (PC104) via an IEEE 1394 interface. The logged image data are stored in the computer for processing and analysis. The view field of the camera is adjusted to avoid capturing other objects besides the road surface. The images of the four road terrain surfaces are shown in Figure 4.1.



(a) Asphalt



(b) Concrete



(c) Grass

(d) Gravel

Figure 4.1 Four types of road terrain surfaces

4.1 Texture Features from Image

A digital image contains numbers of pixels that represent objects in the view field. Those pixels occur as a two-dimensional array with each pixel having a corresponding coordinate and grey level value. In recent years, a tremendous amount of processing methods, such as coding, restoration, enhancement, and classification have been developed and presented. In this case, a method to extract texture features from images to realize the purpose of road terrain classification is applied.

As is known, texture and grey level value of an image have an inextricable relationship to each other. When a small area of an image has a slight variation, such as variation of features of discrete grey level, the dominant property of this is the grey level. If the variation of discrete grey level is wide, then the dominant property of the area is texture. This texture also depends on the size of the area, the relative sizes of the discrete

features, and the number of features. More specifically, if the small area occurs only in one pixel because there is only one discrete feature, then it presents a grey level as the only property. When the number of features of the discrete grey level increases within the small area, it is presented as a texture property.

One of the important properties of grey level and texture is the spatial pattern of the pixels that compose grey level features [72]. If there is no spatial pattern and the grey level variation is wide, a fine texture is achieved. Meanwhile, it becomes a relatively coarse texture if the spatial pattern is more definite and involves more pixels.

Moreover, the discrete grey level features are extremely fuzzy in the preceding description of texture. In the task of analysing the texture, it is suggested to consider more general or macroscopic concepts than discrete grey level features.

Therefore, to obtain the texture features of an image, it is supposed that the texture information on the image is contained in the overall spatial relationship. Furthermore, if the specific angular relationships and distance information between neighbouring pairs of pixels of the image can be described or calculated, then the computed results would present as the texture features. In other words, the formed feature matrix is comprised of spatial dependent grey-level information of nearest neighbours with specified angular intervals. These feature matrices make road terrain type classification possible.

4.2 Image Feature Matrix Establishment

4.2.1 Gray-Level Co-occurrence Matrix

The Grey-Level Co-occurrence Matrix (GLCM) [73] is applied to form the texture feature matrix for the task of road terrain classification.

Suppose a rectangular image has a resolution of $x \times y$ which means N_x columns and N_y rows. Also suppose that the grey level appearing at each pixel is quantized to N_g levels.

$$L_x = \{1, 2, \cdots, N_x\} \tag{4.1}$$

$$L_y = \{1, 2, \cdots, N_y\} \tag{4.2}$$

$$G_x = \{0, 1, \cdots, N_g - 1\} \tag{4.3}$$

Let (4.1) be the columns, (4.2) be the rows, and (4.3) be the set of N_g quantized grey levels. The set $L_y \times L_x$ is the set of pixels of the image ordered by their row-column designations. The image I can be represented as a function thatintos some grey level in G to each pixel or pair of coordinates in $L_y \times L_x$; $I : L_y \times L_x \to G$.

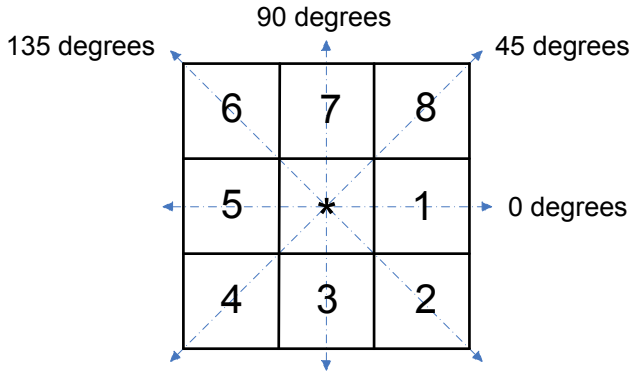


Figure 4.2 Angular relationships with nearest neighbours

The essential idea of this method is a measure, more precisely, four closely related measures that lead to the texture features [73]. These measures form angular nearest-neighbour grey-level spatial-dependence into matrices. To describe these, it is considered that a pixel that excluding the pixels on the edge of the image, to have eight nearest-

neighbour pixels. As shown in Figure 4.2, pixel * has 8 spatial nearest neighbours around it. They are pixels 1 and 5 at 0 degree nearest neighbours to pixel *; pixels 8 and 4 at 45 degrees nearest neighbours; pixels 7 and 3 as 90 degrees nearest neighbours; and pixels 6 and 2 as 135 degrees nearest neighbours to pixel *.

The texture-context information is adequately specified by the matrix of relative frequencies P_{ij} with two neighbouring pixels separatediny distance d occurring in the image I, one with grey level co-occurrence frequencies as a function of the angular relationship and distance between the neighbouring pixels and a function of the distance between them. Formally, for angles quantized to 45 degree intervals, the unnormalized frequencies are defined as (4.4) [73], where $\#$ denotes the number of elements in the set.

$$
\begin{aligned}
P(i,j,d,0^\circ) &= \#\{((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \\
&\quad \left| k-m=0, |\, l-n \,| = d, I\left(k,l\right)=i, I\left(m,n\right)=j \right\} \\
P(i,j,d,45^\circ) &= \#\{((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \\
&\quad \left| (k-m=d, l-n=-d,) \text{ or} \right. \\
&\quad (k-m=-d, l-n=d), I(k,l)=i, I(m,n)=j\} \qquad (4.4) \\
P(i,j,d,90^\circ) &= \#\{((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \\
&\quad \left| \, |\, k-m \,| = d, l-n=0, I\left(k,l\right)=i, I\left(m,n\right)=j \right\} \\
P(i,j,d,135^\circ) &= \#\{((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \\
&\quad \left| (k-m=d, l-n=d,) \text{ or} \right. \\
&\quad (k-m=-d, l-n=-d), I(k,l)=i, I(m,n)=j\}
\end{aligned}
$$

It should be noted that these matrices are symmetric:

$$P(i,j,d,a) = P(j,i,d,a) \tag{4.5}$$

The distance matrix ρ implicit in the preceding equations can be explicitly defined by:

$$\rho((k,l),(m,n)) = \max\{|k-m|,|l-n|\} \tag{4.6}$$

4.2.2 Feature Extraction and Feature Matrix Formation

There are 20 texture features applied in this case. The following equations define these features. Let $p(i,j)$ be the $(i,j)th$ entry in a normalized GLCM. The mean and standard deviations for the rows and columns of the matrix are [74]:

$$\mu_x = \sum_i \sum_j i \cdot p(i,j) \tag{4.7}$$

$$\mu_y = \sum_i \sum_j j \cdot p(i,j) \tag{4.8}$$

$$\sigma_x = \sum_i \sum_j (i-\mu_x)^2 \cdot p(i,j) \tag{4.9}$$

$$\sigma_y = \sum_i \sum_j (j-\mu_y)^2 \cdot p(i,j) \tag{4.10}$$

The features applied are calculated as follows [73] [74] [75]:

1) Autocorrelation:

$$f_1 = \sum_i \sum_j (i,j)\, p(i,j) \tag{4.11}$$

2) Contrast:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \| i-j | = n \right\} \tag{4.12}$$

3) Correlation:

$$f_3 = \frac{\sum_i \sum_j (i,j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{4.13}$$

4) Cluster Prominence:

$$f_4 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i,j) \tag{4.14}$$

5) Cluster Shade:

$$f_5 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i,j) \tag{4.15}$$

6) Dissimilarity:

$$f_6 = \sum_i \sum_j | i-j | \, p(i,j) \tag{4.16}$$

7) Energy:

$$f_7 = \sum_i \sum_j p(i,j)^2 \tag{4.17}$$

8) Entropy:

$$f_8 = -\sum_i \sum_j p(i,j) \log(p(i,j)) \tag{4.18}$$

9) Homogeneity:

$$f_9 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \tag{4.19}$$

10) Maximum Probability:

$$f_{10} = \underset{i,j}{MAX}\ p(i,j) \tag{4.20}$$

11) Sum of Squares: Variance

$$f_{11} = \sum_i \sum_j (i-\mu)^2\, p(i,j) \tag{4.21}$$

12) Sum Average:

$$f_{12} = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{4.22}$$

13) Sum Variance:

$$f_{13} = \sum_{i=2}^{2N_g} (i-f_{16})^2\, p_{x+y}(i) \tag{4.23}$$

14) Sum Entropy:

$$f_{14} = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log\left\{p_{x+y}(i)\right\} \tag{4.24}$$

15) Difference Variance:

$$f_{15} = \text{variance of } p_{x-y} \tag{4.25}$$

16) Difference Entropy:

$$f_{16} = -\sum_{i=0}^{N_{g-1}} p_{x-y}(i) \log\left\{p_{x-y}(i)\right\} \tag{4.26}$$

17) 18) Information Measures of Correlation:

$$f_{17} = \frac{HXY - HXY1}{\max\left\{HX, HY\right\}} \tag{4.27}$$

$$f_{18} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2} \tag{4.28}$$

$$HXY = -\sum_{i}\sum_{j} p(i,j) \log(p(i,j))$$

$$HXY1 = -\sum_{i}\sum_{j} p(i,j) \log\left\{p_x(i) p_y(j)\right\}$$

$$HXY2 = -\sum_{i}\sum_{j} p_x(i) p_y(j) \log\left\{p_x(i) p_y(j)\right\}$$

19) Inverse Difference Normalized:

$$f_{19} = \sum_{i,j=1}^{G} \frac{p(i,j) / \sum_{i,j=1}^{G} p(i,j)}{1 + |i-j|^2 / G^2} \tag{4.29}$$

where G is the quantized number of grey levels of the image.

20) Inverse Difference Moment Normalized:

$$f_{20} = \sum_{i,j=1}^{G} \frac{p(i,j) / \sum_{i,j=1}^{G} p(i,j)}{1 + (i-j)^2 / G^2} \tag{4.30}$$

where G is the quantized number of grey levels of the image.

The texture features of (4.11) to (4.20) are first introduced by Soh et al in [74]. The features of (4.21) to (4.28) are used by Haralick et al in [73]. The last two features of (4.29) and (4.30) are applied by Clausi in [75]. All of these texture feature values are then employed to form a feature matrix of the road terrain images as (4.31):

$$F_{im} = \begin{bmatrix} F_{s_1,t_1} & \cdots & F_{s_n,t_1} \\ \vdots & \ddots & \vdots \\ F_{s_1,t_m} & \cdots & F_{s_n,t_m} \end{bmatrix} \tag{4.31}$$

As seen from (4.31), the feature vector of each image is arranged column by column to form the feature matrix F_{im}. In this matrix, every column refers to a texture feature vector extracted from an image while every row represents the particular features extracted from all images.

4.3 Experiment

4.3.1 Experimental Platform

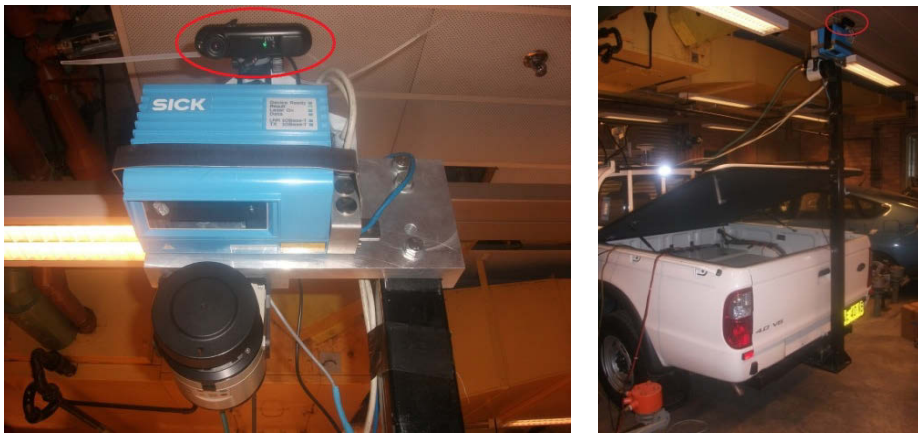


Figure 4.3 The mounted camera on top of the frame

The CRUISE is employed as the experimental platform once again. As shown in Figure 4.3, the camera is mounted on top of the frame above the vehicle's back tray,

connecting with Computer No. 1 (PC104) via a Firewire interface. A laptop transmits the controlling commands to Computer No. 1 (PC104) to start logging image data, saving them, and so on. The camera is configured at 30 *FPS* with 800×600 pixels resolution. The other devices are configured exactly as in the previous accelerometer experiment.

4.3.2 Image Based Experiments

Since all the sensor data are logged at the same time with timestamps, all the images are captured in the same travelling routes of the CRUISE, as shown in Figure 3.18.

The collected image data is then processed for the training to generate the classification model for the consequent testing. It is designed so that each of the images produces a road terrain result. These results are also considered (to be introduced in Chapter VI) for other enhanced methods for the final target.

4.4 Experiment Results

As described in Section 3.8.4, the image data is logged when the CRUISE is driven on asphalt, concrete, grass and gravel roads during daytime with fine sunlight. The drives on different road terrains factor in various circumstances of the different environments and natural driving conditions.

Table 4.1 Image data for training and testing

Road Terrain	Training			Testing		
	Speed (km/h)	Displacement (m)	Number of Samples	Speed (km/h)	Displacement (m)	Number of Samples
Asphalt	20, 30, 40, 50	328	305	0 ~ 70	3948	4481
Concrete	20, 30	284	282	0 ~ 45	952	828
Grass	10	304	384	0 ~ 25	876	813
Gravel	20, 30	312	308	0 ~ 35	2712	2491
Total		1228	1279		8488	8613

Similar to the acceleration based experiment, an on-line image based simulated experiment is conducted. As shown in Table 4.1, data with particular speeds for training are selected and then other data with all possible ranges are tested.

As the previous tests, the classifier is empirically selected to be SVM. The normalization is carried out even as the five-fold cross validation is not applied to the image feature matrix for classification. The actual labels are manually configured to compare with the predicted labels according to the visible road recognized by the operator.

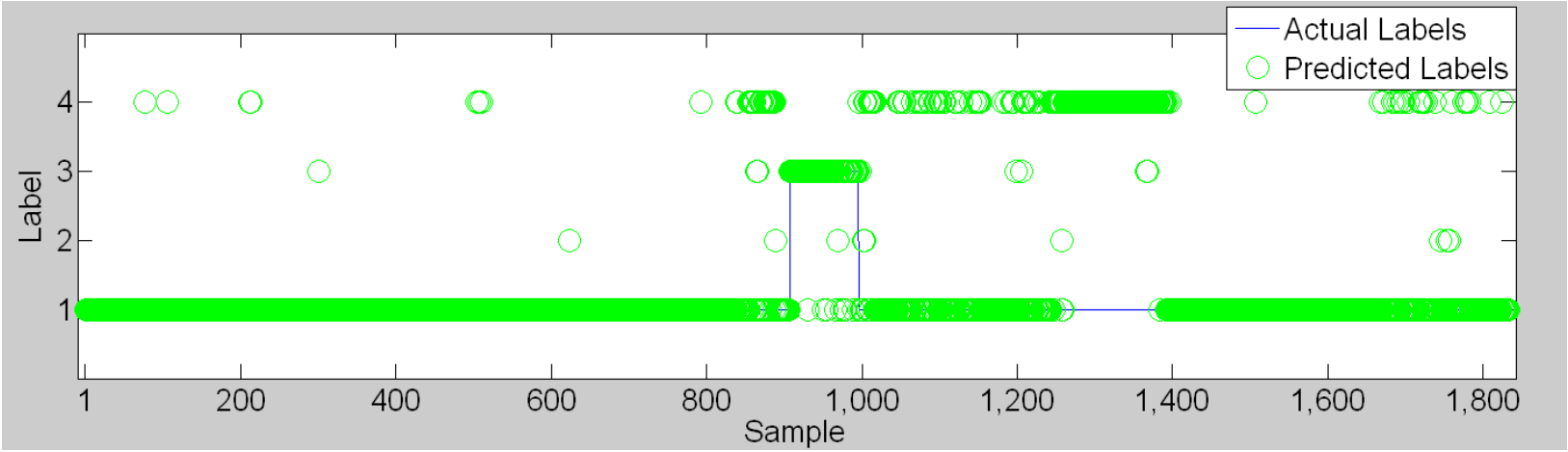


Figure 4.4 (a) Image based experiment label sequence: segment 1

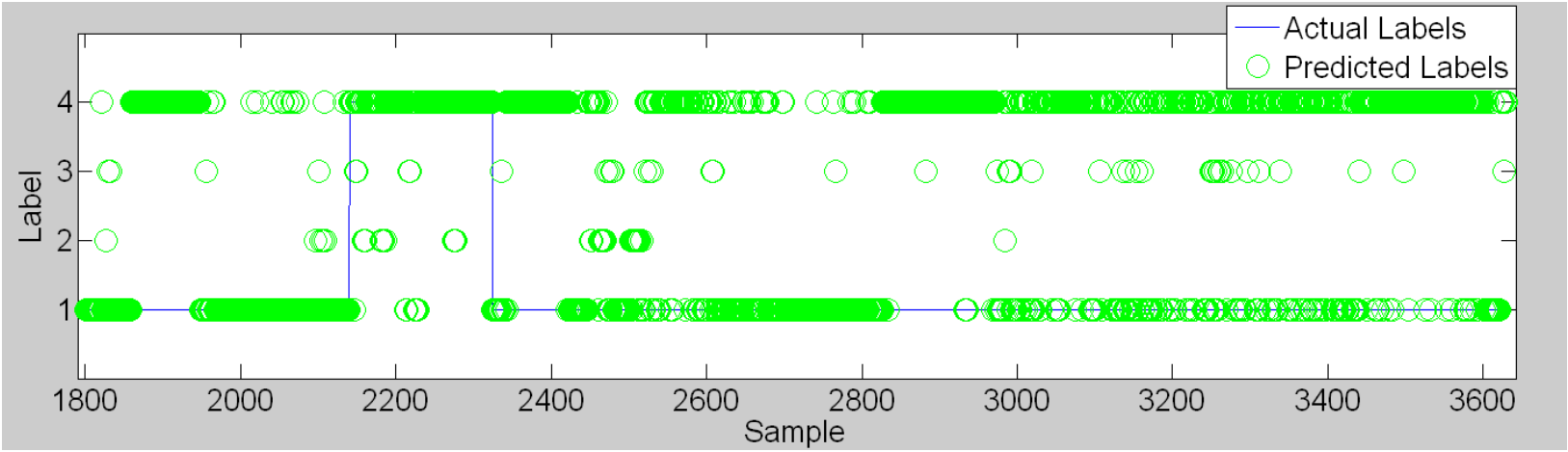


Figure 4.4 (b) Image based experiment label sequence: segment 2

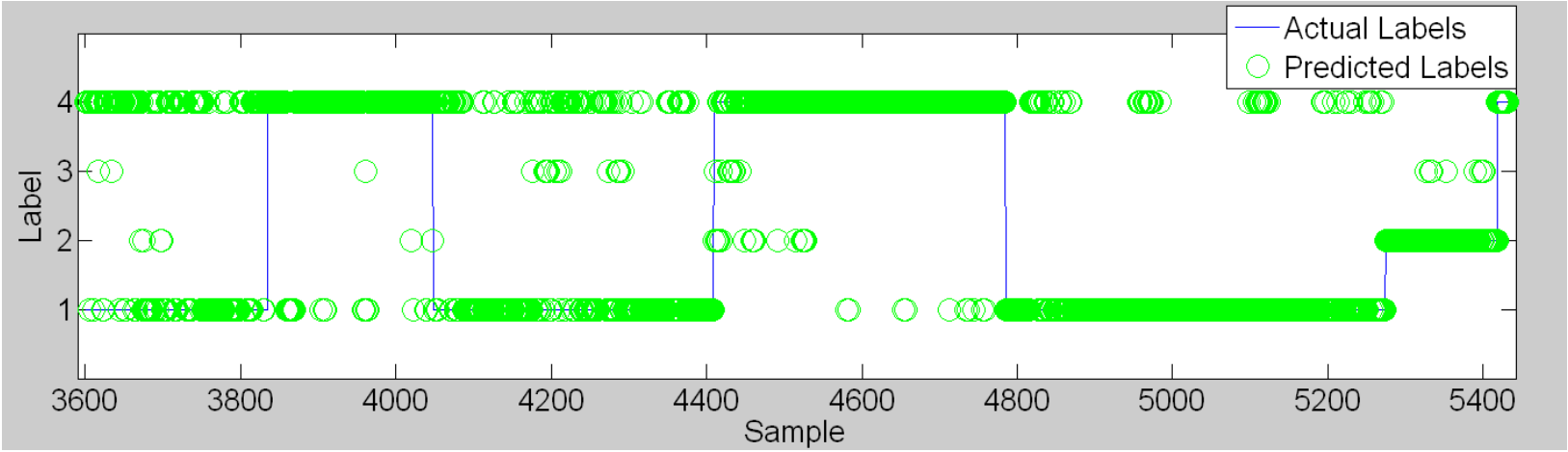


Figure 4.4 (c) Image based experiment label sequence: segment 3

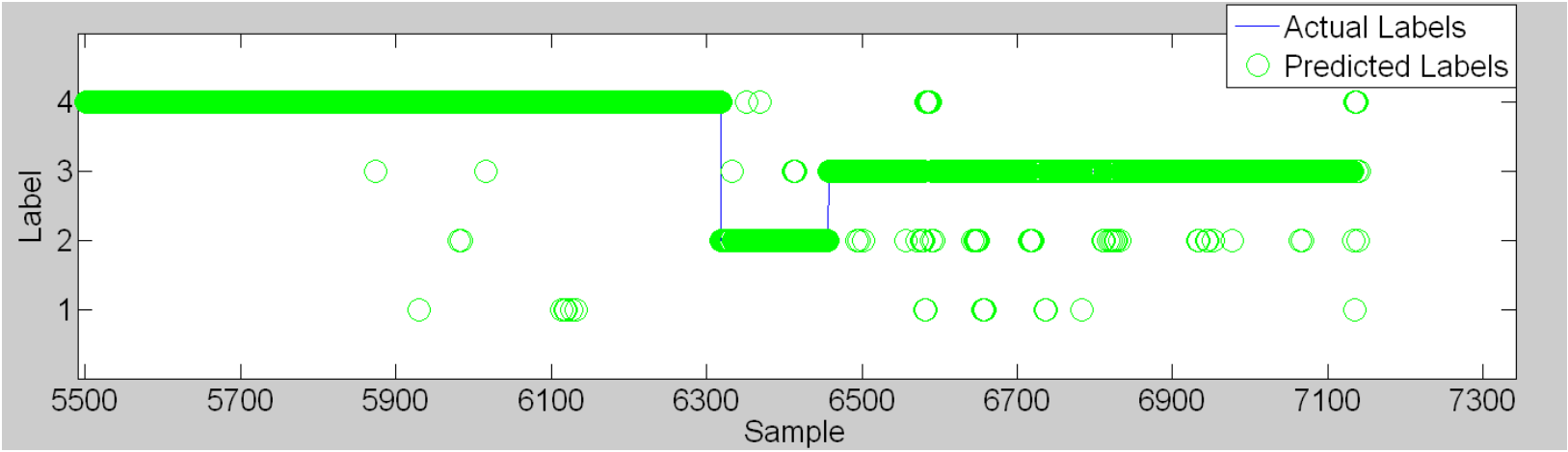


Figure 4.4 (d) Image based experiment label sequence: segment 4

Figures 4.4 (a) to (d) show the image based experiment label sequences. The horizontal axis is the index of the samples (images). There are four points (1, 2, 3, and 4) on the vertical axis that show the class labels. The predicted labels classified are marked by green circles while the actual labels are represented by blue lines. The entire testing image data is actually synchronized with acceleration data. Therefore the arranged data acts as the vehicle runs from one type of terrain to another by random segments of a road, as in Figure 3.18. As illustrated in the Figure 4.4 (a), from an index of 1200, many asphalt images are classified to be gravel roads. Even in some road segments, asphalt terrain is continuously identified as gravel terrain. Some of those image sequences (odd index images only) are found and shown in Figures 4.5 (a) and (b).

Figure 4.5 (a) shows a particular asphalt road segment while Figure 4.5 (b) shows a gravel road segment. These captured images have changing shadows, motion blur, and over-exposure damages. A regular camera can calibrate automatic exposure time and this exposure time depends on the average luminance of a whole image and it takes time to set a new value of the exposure time. This means the camera is not able to capture every picture with perfect parameters. Changing shadows caused by the moving vehicle affect the quality of images. So does the high speed of the vehicle. Due to the tendency to drive at higher speeds on asphalt roads, poor quality images are more often gathered on such a terrain. As can be seen in Figure 4.5 (a), the over-exposed asphalt images hardly show any difference from the over-exposed concrete images (Figure 4.5 (b)), which confuses the classifier and causes it to identify the asphalt terrain to be gravel.

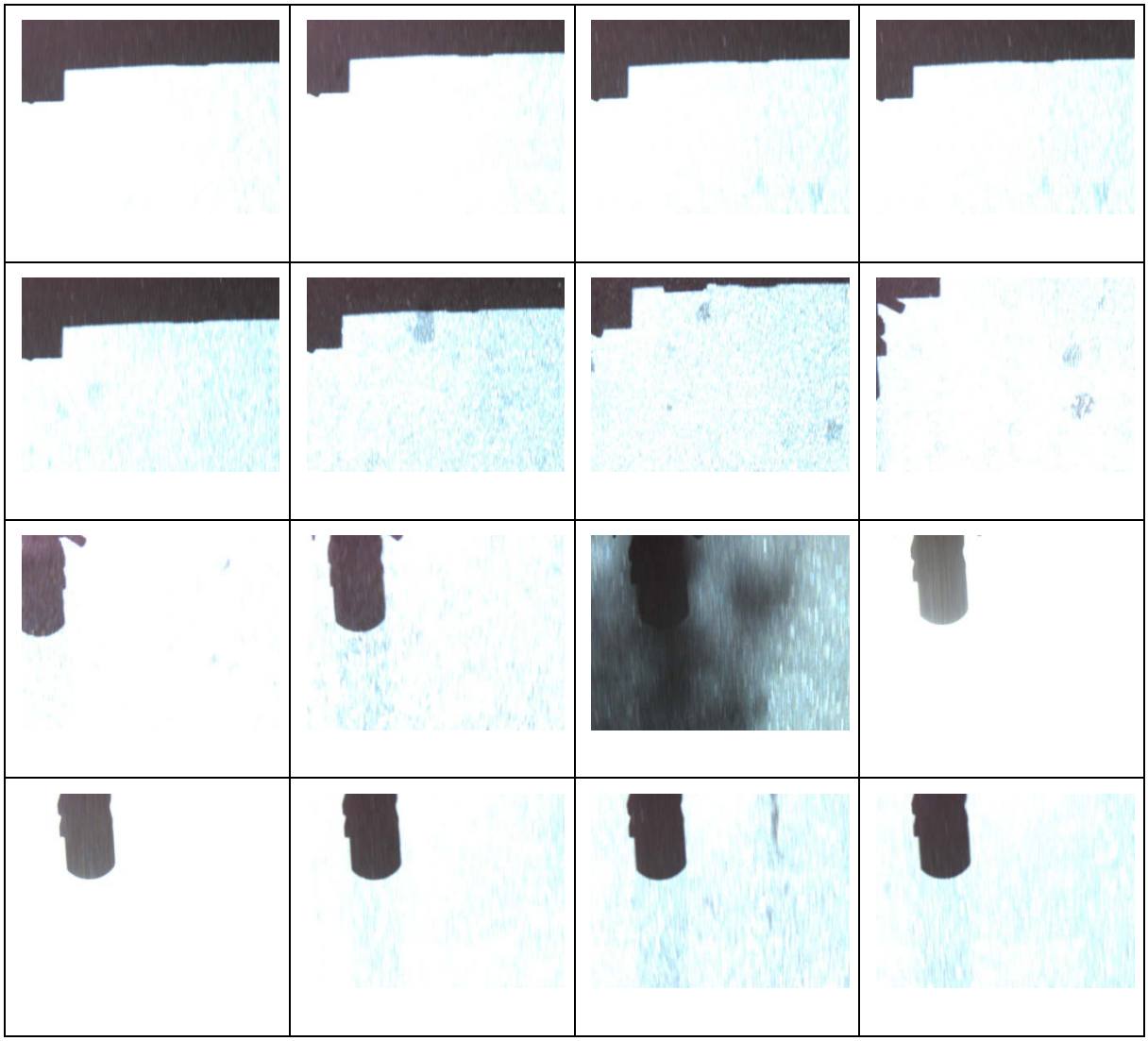


Figure 4.5 (a) Image sequence of an asphalt terrain segment

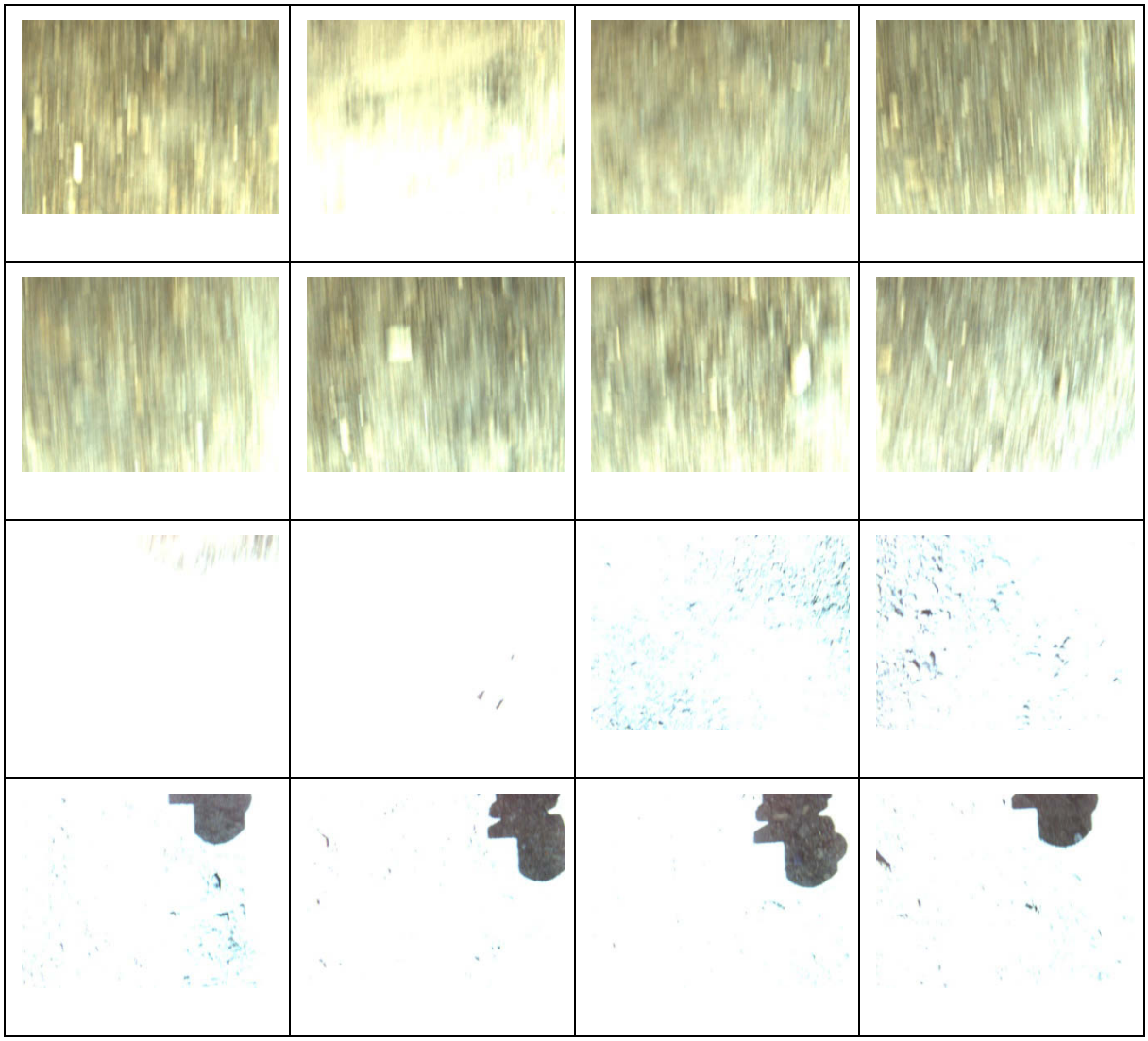


Figure 4.5 (b) Image sequence of a grave terrain segment

Table 4.2 The experimental results using image data without PCA process

Training / Testing	Asphalt @20,30,40,50 (*km/h*)	Concrete @20,30 (*km/h*)	Grass @10 (*km/h*)	Gravel @20,30 (*km/h*)
Asphalt @0~70	2939/4481 65.6%	41/4481 0.9%	72/4481 1.6%	1429/4481 31.9%
Concrete @0~45	0/828 0%	778/828 94.0%	44/828 5.3%	6/828 0.7%
Grass @0~25	19/813 2.3%	42/813 5.2%	743/813 91.4%	9/813 1.1%
Gravel @0~35	87/2491 3.5%	45/2491 1.8%	25/2491 1.0%	2334/2491 93.7%

Average Classification Accuracy: 86.2%

The statistic of Figure 4.4 is shown in Table 4.2. Similar with Table 3.9, each of the rows shows the numbers of samples that are classified to be a specific class and the total number of that class, along with its classification accuracy for each.

As indicated in Table 4.2, with the exception of the asphalt road, the concrete (94.0%), grass (91.4%), and gravel roads (93.7%) are classified reliably. The classification accuracies are all above 90% except asphalt road. The asphalt roads are still the most difficult to classify, presenting the lowest rate of 65.6%. In this experiment, the classifier confuses asphalt roads with gravel roads in approximately one third of the cases. The average accuracy using image data is 86.2% which is significantly higher than that of using acceleration data which achieves an accuracy of only 65.0%.

Table 4.3 The experimental results using image data with PCA process

Training / Testing	Asphalt @20,30,40,50 (*km/h*)	Concrete @20,30 (*km/h*)	Grass @10 (*km/h*)	Gravel @20,30 (*km/h*)
Asphalt @0~70	3110/4481 69.4%	41/4481 0.9%	22/4481 0.5%	1308/4481 29.2%
Concrete @0~45	0/828 0%	820/828 99.0%	4/828 0.5%	4/828 0.5%
Grass @0~25	19/813 0.9%	42/813 5.4%	743/813 92.5%	9/813 1.2%
Gravel @0~35	87/2491 4.3%	45/2491 0.7%	25/2491 0.7%	2334/2491 94.2%

Average Classification Accuracy: 88.8%

The input feature vector extract from each image is also processed through the PCA algorithm. By the empirical evaluation tests, the number of reduced dimensions is selected to be 75. Table 4.3 shows that all the individual classification accuracies slightly increase as expected, however, without major improvement. The 1% to 5% higher rates for these four classes lead to an average classification accuracy of 88.8% which is slightly better than without using the PCA method. However, it is still recommended to employ the PCA process during the process to achieve classification results as accurate as possible based on image data only.

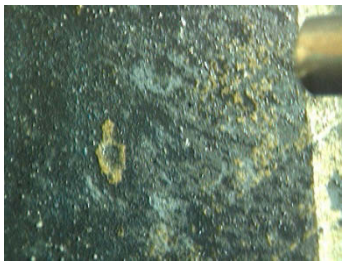
		
(a) Asphalt – motion blur	(b) Asphalt – different texture	(c) Concrete – over exposure
		
(d) Grass – motion blur	(e) Grass – over exposure	(f) Gravel – motion blur

Figure 4.6 Poor quality images

Some samples are wrongly classified due to the poor quality of images. As can be seen in Figure 4.6, images can be corrupted by over/under exposure or motion blur. Specifically, Figure 4.6 (b) shows an asphalt road surface but with an irregular asphalt texture which is not contained in the training data set. This confuses the classifier during testing. Figures 4.6 (a), (d), and (f) present asphalt, grass, and gravel road surfaces with motion blur. The vehicle's fast speed negatively impacts the sharpness of the images. Depending on the performance of the camera used, the automatic exposure time sometimes does not adjust fast enough. This leads to over exposure of the images, as shown in Figure 4.6 (b) and (e), with direct strong sunlight.

4.5 Conclusion

In this chapter, a method to classify road terrain types based on road surface image data has been proposed. The experiment results indicated the feasibility of identifying road terrain types using road surface images.

The images captured by the camera are processed to extract texture features for the classification. The feature matrix extracted from and formed by the image data was trained using an SVM classifier. The generated model was then used in the testing stage. The experiment results showed that the proposed method worked in many scenarios.

However, the method's main weakness was its reliance on camera images, where the quality of the images heavily depends on illumination. Apart from that, vehicle speeding led to motion blur in images as well. A camera with a faster shutter would solve this problem, but this brought another problem of overexposure. In the tests, the problem due to blurred images seems to be significant when the speed was over 50 km/h.

LRF is known to be unaffected by ambient light conditions in most situations. Furthermore, high-speed LRFs are available and able to produce usable sensor data at normal vehicle speeds. Therefore, the next chapter investigates the use of laser scanners for road type classification.

CHAPTER V

LRF BASED ROAD TERRAIN CLASSIFICATION

The utilization of an accelerometer is supposed to estimate the road profile from which the spatial features are extracted. However, the estimated road profile only represents the curvature of the terrain on the vehicle's moving orientation. Here it is proposed to employ the LRF sensor to represent the road surface which would contain substantial information. Moreover, because it works with higher scanning speed, this classification method is expected to be minimally affected by the change of the vehicle's speed.

In this chapter, the LRF based method to classify the road terrain types is presented. At the beginning, the geometric arrangement of the LRF is introduced. Then a description how to reconstruct the 3D road surface where the features are extracted is given. Then LRF based experiment and the results are presented.

5.1 Geometric Arrangement of the LRF

A downward-looking SICK LMS111 is mounted on the CRUISE as shown in Figure 5.1. It scans the road surface vertically in a two-dimensional plane at a 50 *Hz* sampling rate. The LRF has a 270° field of view with 0.5° angular resolution providing 541 range values per scan. While the vehicle is moving forward, it leaves a trace of three dimensional point cloud of the surface.

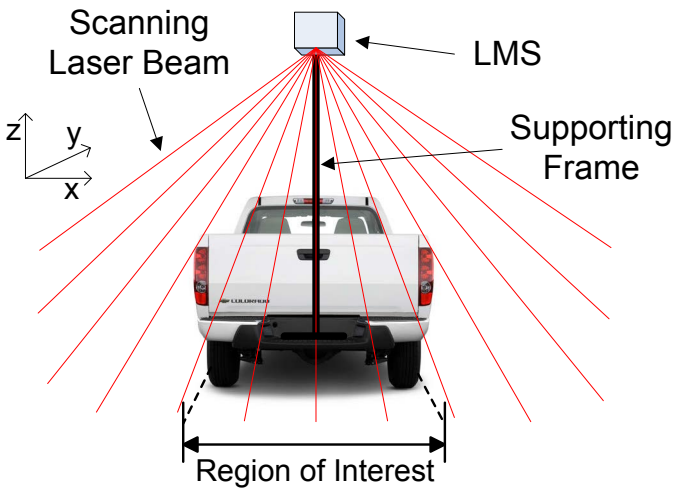


Figure 5.1 The geometric arrangement of the LRF

5.2 Reconstruction of the Road Surface

It is well known that the laser beam emitted in an LRF is deflected using a rotating mirror so that the surroundings are scanned in a circular manner. In general, once a laser beam hits a surface, the light is reflected. This reflected energy can be partly received by the photodiode in the LRF calculating the range to an object based on the time of flight measurements.

As shown in Figure 5.2, the surface is estimated from the laser range data and the speed data coming from the wheel encoders.

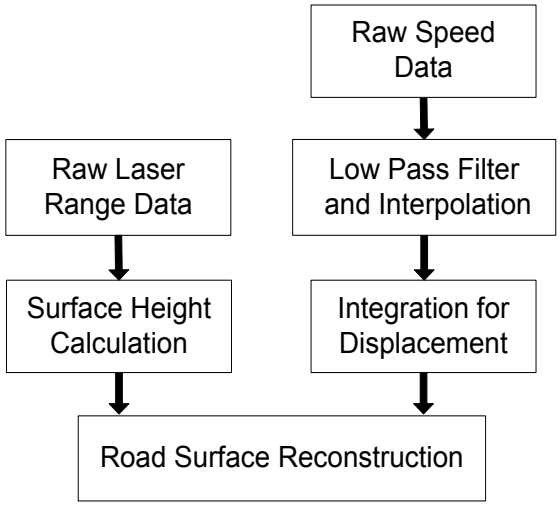


Figure 5.2 Reconstruction of the road surface

5.2.1 Range Data Processing

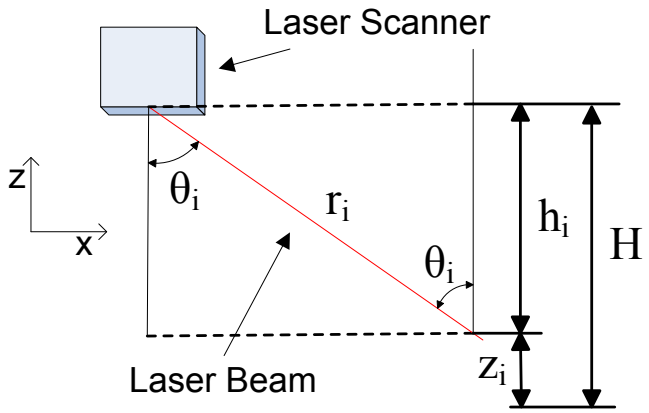


Figure 5.3 Mounting geometry of the LRF

The vertical coordinate, z_i, of each range measurement can be easily reconstructed by (see Figure 5. 3):

$$z_i = H - r_i \cos(\theta_i) \tag{5.1}$$

where, r_i is laser range value, ϑ_i is included angle between the current laser beam and *z* axis, and *H* is the reference height from a relatively flat floor to the height of the LRF. In a similar way, the x-axis coordinate can be calculated as:

$$x_i = r_i \sin(\theta_i) \tag{5.2}$$

The 270° scanning field of view contains road surface as well as other nearby objects. Therefore, as illustrated in Figure 5.1, a 1.3 *m* wide region of interest is defined for purposes of this experiment. This brings the distance between two sampling points of a particular scan to approximately two centimetres on a road surface. This LRF is mounted 2.2 meters high above the road surface.

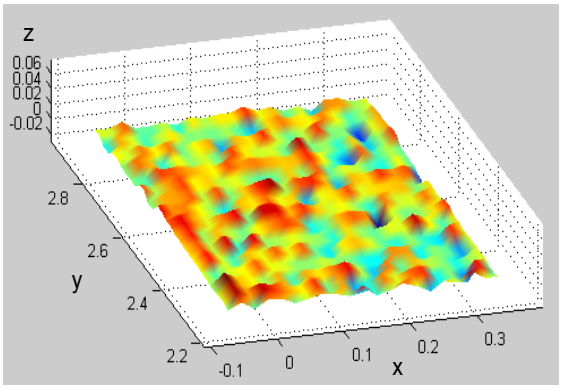
5.2.2 Speed Data Processing

As mentioned in Section 3.4.1, the speed of the vehicle is also measured by the wheel encoders. Since this speed sampling frequency is immoderate low to match the LRF data, it is interpolated to 50 *Hz* using the proximal interpolation method (also known as nearest-neighbour interpolation) [60]. Then, the longitudinal vehicle displacement, *y(t)*, is estimated as:

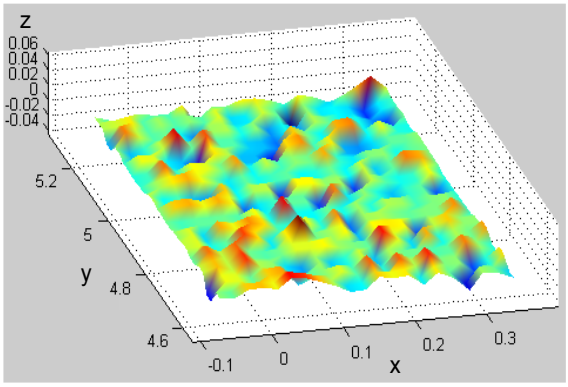
$$y(t) = \int_0^t \left| v(t) \right| dt \tag{5.3}$$

where *v(t)* is the estimated vehicle speed.

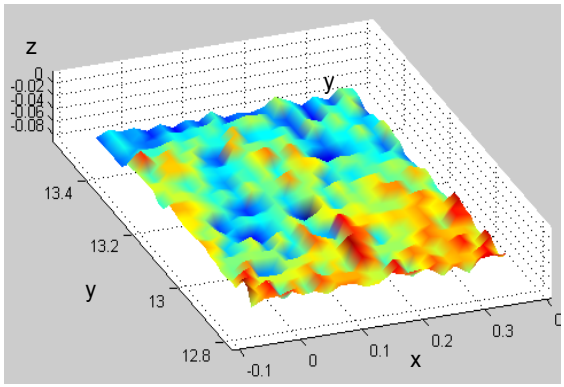
5.2.3 Road Surface



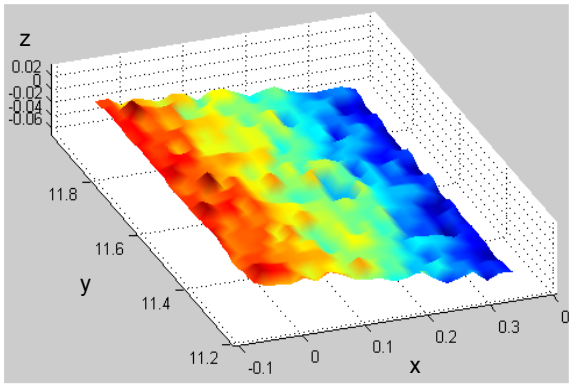
(a) Asphalt



(b) Concrete



(c) Grass



(d) Gravel

Figure 5.4 Three-dimensional surface data of four different road types

Once the data is processed to estimate the longitudinal and lateral displacements of the laser data point clouds, a three-dimensional view of the displacements can be generated. Such displacement profiles of the road types are shown in Figure 5.4. It appears possible to make a qualitative assessment of the plots and their differences purely by visual inspections.

5.3 Feature Matrix

The estimated 3D surface has more dissimilarity in the resolution along the vehicle moving direction due to the variability of the speed. However, the vehicle speed has minimal effect on the lateral data. Therefore, in this work, only the lateral components are considered. Rather than relying on each new lateral scan of data, this will concatenate a group of scans captured over the vehicle length, which is 4 metres. This is done for future comparisons with other sensor modalities. The number of scans depends on the vehicle's speed. In general, however, between 35 to 145 scans are captured at speeds of 20 to 80 *km/h*.

The feature matrix is formed by carrying out the FFT on each scan. The PSD is then calculated using Welch's method [55]. Similar with the processing in section 3.2.1.c, but the different parameters are set. The start frequency, end frequency and frequency step are decided as 0 *c/m* to 35 *c/m* with 0.1 *c/m* interval.

$$F_{ls} = \begin{bmatrix} F_{s_1,f_1} & \cdots & F_{s_n,f_1} \\ \vdots & \ddots & \vdots \\ F_{s_1,f_m} & \cdots & F_{s_n,f_m} \end{bmatrix} \tag{5.4}$$

The PSD of each group of scans, which defines the feature vector, is then arranged column by column to form the feature matrix *F*, as in (5.4). In this matrix, every column refers to a feature vector (a group of scanning lines) while every row represents the features (PSD) extracted using the above procedure. For instance, in this case, F_{s_1,f_1} presents the PSD from 0 *c/m* to 0.1 *c/m* of the first group of lines, while F_{s_n,f_1} presents the PSD from 0 *c/m* to 0.1 *c/m* of the nth group, and F_{s_n,f_m} presents the PSD from 34.9 *c/m* to the end 35 *c/m* of the *n*th segment. Overall, the matrix *F* contains *n* samples while each

sample has *m* features. Each value of the elements in the matrix (5.4) indicates a power at a particular investigated frequency forming the final feature matrix. It is normalized to reflect each row of *F* mapped to [0, 1].

5.4 Experiment

5.4.1 Experimental Platform

Figure 5.5 The mounted downward-looking LRF

The CRUISE is again used as the experimental test bed. As shown on the right side of Figure 5.5, a SICK LMS111 is mounted on the vehicle looking down for scanning the road surface. It is aligned with the central axis of the vehicle and the scanning plane is perpendicular to the ground. When the vehicle moves forward, the LRF scans the surface of the road and then leaves a structured point cloud corresponding to the road surface. A laptop computer is used for laser data logging where a PC104 computer logs the speed data. Both computers are synchronized via Ethernet using the Network Time Protocol.

5.4.2 LRF Based Experiments

As mentioned previously, the downward-looking LRF data is logged at the same time with other sensors. It is then used to train the classifier. The testing part is then designed to report the road terrain labels through the scanning of the laser beam.

5.5 Experiment Results

As for the LRF based experiments, the selected data with particular speeds is used to investigate the speed independency of the LRF sensor, and used trained classification model to do the testing for the road terrain classification.

Table 5.1 LRF data for training and testing

Road Terrain	Training			Testing		
	Speed (km/h)	Displacement (m)	Number of Samples	Speed (km/h)	Displacement (m)	Number of Samples
Asphalt	20, 30, 40, 50	328	4005	0 ~ 70	3948	61194
Concrete	20, 30	284	3950	0 ~ 45	952	11836
Grass	10	304	5383	0 ~ 25	876	11386
Gravel	20, 30	312	4225	0 ~ 35	2712	24208
Total		1228	17563		8488	108624

As shown in Table 5.1, similar to the previous two sensors, LRF data with limited speeds are selected to train while other LRF data are used to test. The classifier is set as SVM. The actual labels are manually set to compare with the testing results. Apparently,

the amount of the data that a LRF produces is quite large. The amounts of the training samples of all four classes are approximately equal to conduct balanced and appropriate training. On the other hand, the amount of the testing data is much more. It could be noted that the asphalt road is the most widely available road types whereas concrete and grass roads are rare to find in Sydney.

5.5.1 Speed Independency

The speed independency is investigated to prove the assumption that the LRF data collected are minimally affected by the operating speed of the vehicle. This is reasonable as the LRF has a fast sampler which can capture range data in a few microseconds. To get more robust and reliable classification results in this case, the K-fold cross validation is then employed, where K is set as five as well.

This hypothesis is tested with data captured on Asphalt and Gravel roads at a range of different speeds. While the classifier is trained at a particular speed, it is tested at a different speed on the same road type. The data set is divided into training and testing parts with particular speeds for the speed independency tests.

Table 5.2 shows the classification results as the CRUISE is running at different speeds. The first row shows the classification accuracies of asphalt and gravel tested at 40 *km/h* (marked as A40) and 10km/h (marked as G10) while trained both at the speed of 20 *km/h* (A20&G20). The other rows are some other different combinations of speeds of the training and testing data sets. The results based on different training and testing data sets show very high accuracies, leading to the conclusion that the LRF data is speed independent.

Table 5.2 Classification at different speeds

Training Speed (km/h)	Testing Speed (km/h)	Asphalt Accuracy	Gravel Accuracy
A20 & G20	A40 & G10	100.0%	99.2%
A20 & G20	A30 & G30	100.0%	100.0%
A20 & G20	A30 & G10	100.0%	99.2%
A30 & G30	A20 & G10	100.0%	100.0%
A30 & G30	A40 & G20	100.0%	100.0%

As mentioned in Section 3.8.2, the acceleration data based classification has the problem of speed dependency. It requires sufficiently large training data covering all speeds for better classification results. However, as observed through this experiment, when using LRF data, it is not necessary to use all speed data in the training phase but doing so still provides reasonable classification results.

5.5.2 LRF Based Experiment

According to the prior tests, the classifier is selected to be SVM as well. The normalization is also applied to the LRF feature matrix for the classifier. The actual labels are manually configured to compare with the results according to the visible road recognized by the operator during the experiment.

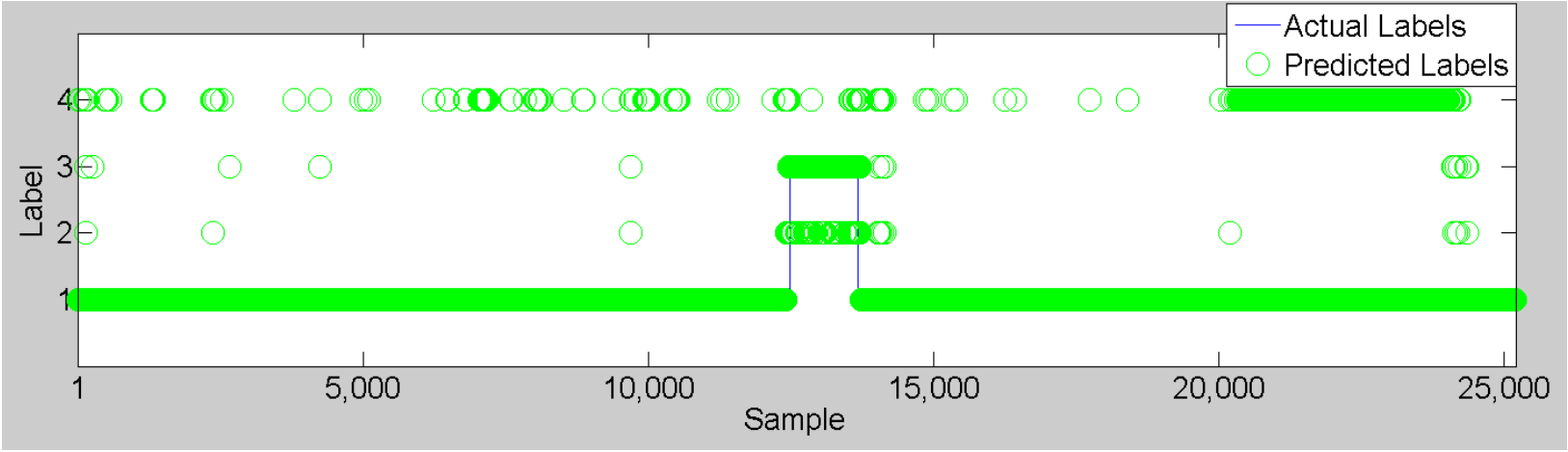


Figure 5.6 (a) LRF based experiment label sequence: segment 1

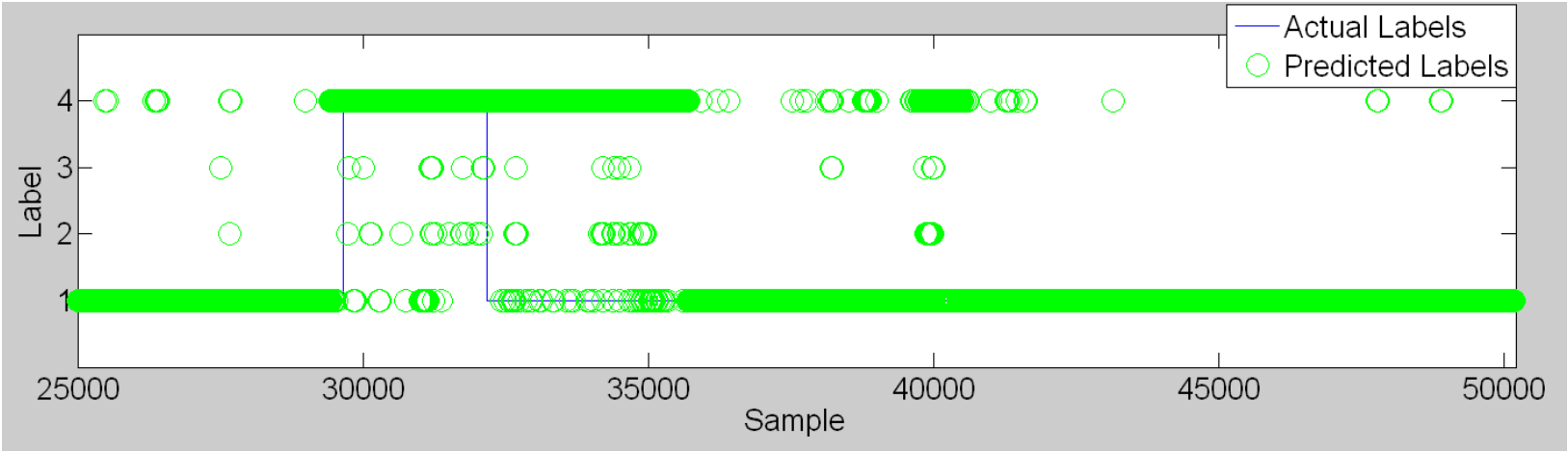


Figure 5.6 (b) LRF based experiment label sequence: segment 2

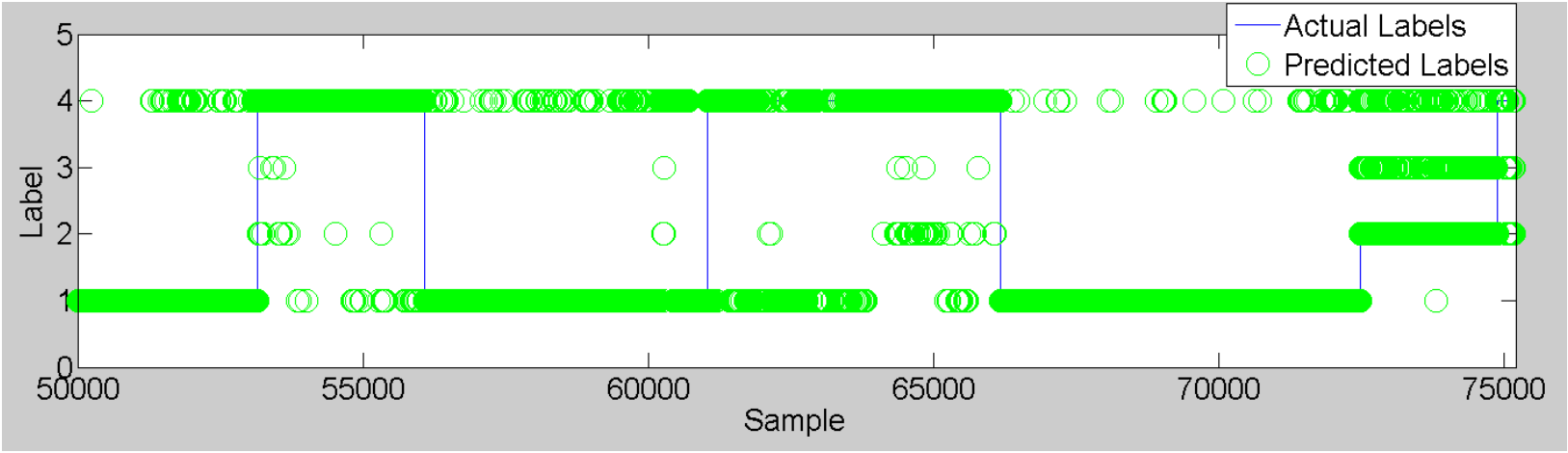


Figure 5.6 (c) LRF based experiment label sequence: segment 3

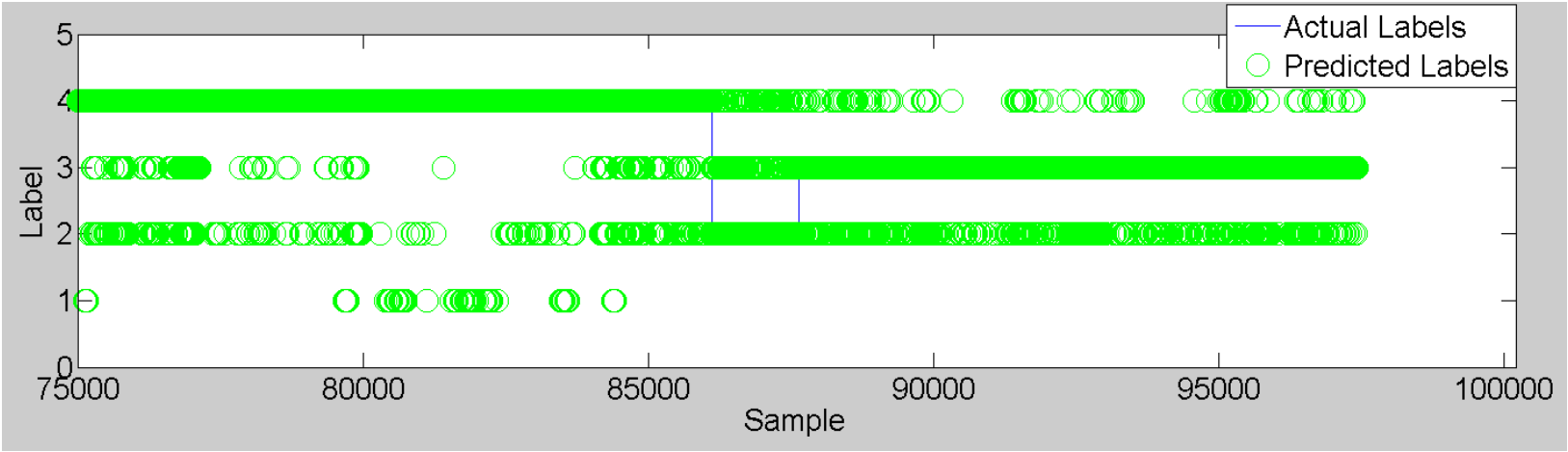


Figure 5.6 (d) LRF based experiment label sequence: segment 4

Figures 5.6 (a) to (d) show the LRF based experiment label sequences. The horizontal axis is the index of the samples (laser beam scans). There are four points (1, 2, 3, and 4) on the vertical axis that show the class labels. The predicted labels classified are marked by green circles while the actual labels are represented by blue lines. The entire testing LRF data is actually synchronized with acceleration and image data. So the details arranged to simulate that the vehicle runs from one type of terrain to another by random segments of road shown Figure 3.18. Taking a qualitative look at Figures 5.6 (a) -(d), it can be seen that each of the classes is classified fairly accurately, save for a few exceptions, where concrete terrain is confused to be grass terrain and grass terrain is wrongly labelled as concrete terrain. In addition, in comparison to the acceleration based experiment (Table 3.10), this LRF based approach is more capable of distinguishing between asphalt and concrete terrains.

Table 5.3 The experimental results using downward LRF data without PCA process

Training / Testing	Asphalt @20,30,40,50 (*km/h*)	Concrete @20,30 (*km/h*)	Grass @10 (*km/h*)	Gravel @20,30 (*km/h*)
Asphalt @0~70	53790/61194 87.9%	306/61194 0.5%	61/61194 0.1%	7037/61194 11.5%
Concrete @0~45	0/11836 0.0%	8616/11836 72.8%	2107/11836 17.8%	1113/11836 9.4%
Grass @0~25	0/11386 0.0%	1617/11386 14.2%	9507/11386 83.5%	262/11386 2.3%
Gravel @0~35	2058/24208 8.5%	1089/24208 4.5%	460/24208 1.9%	20601/24208 85.1%

Average Classification Accuracy: 82.3%

As can be seen in Table 5.3, in the experiment, the accuracies of asphalt, concrete, grass, and gravel road terrain types are 87.9%, 72.8, 83.5%, and 85.1%, respectively. Generally, even though the accuracies are not as high as the acceleration and image based experiments, the asphalt and concrete terrain are distinguished reasonably well. Moreover, it can be seen that some small segments of concrete road are recognized as grass roads 17.8% of the time. This is mainly due to the structures of concrete and grass roads, which are somehow so similar that they confuse the classifier. Overall, the LRF data based road terrain classification method is feasible, however, with average classification accuracy is 82.3%.

Table 5.4 The experimental results using downward LRF data with PCA process

Training / Testing	Asphalt @20,30,40,50 (*km/h*)	Concrete @20,30 (*km/h*)	Grass @10 (*km/h*)	Gravel @20,30 (*km/h*)
Asphalt @0~70	53790/61194 21.2%	306/61194 38.3%	61/61194 16.5%	7037/61194 24.0%
Concrete @0~45	0/11836 0.0%	8616/11836 79.1%	2107/11836 20.9%	1113/11836 0.0%
Grass @0~25	0/11386 0.0%	1617/11386 19.5%	9507/11386 76.7%	262/11386 3.8%
Gravel @0~35	2058/24208 0.0%	1089/24208 3.1%	460/24208 7.6%	20601/24208 89.3%

Average Classification Accuracy: 69.8%

Table 5.4 shows the classification results with the PCA process with relatively poor results when comparing with Table 5.3. Although the accuracies of concrete road and

gravel roads are higher, the average classification accuracy decreases significantly, especially for the asphalt road.

5.6 Conclusion

In this chapter, the method to classify road terrain types based on LRF data was presented. The experimental results showed that it was a feasible method to classify road terrain types using LRF sensor.

The employed LRF sensor scanned the road surface that the vehicle travelled on. With the given speed of the vehicle, the road surface could be reconstructed using the 3D point cloud data of the LRF sensor. The spatial frequency features were extracted again to form the LRF feature matrix for the classification. The speed independency was investigated to verify the assumption that LRF had much less speed effect than the acceleration based one. The generated model was then applied in the testing experiment.

The testing results showed that the LRF based method had success in identifying the asphalt road which was very difficult to achieve using acceleration or image data. Unlike the camera, the LRF sensor had the ability of working at night environment or lack of background light. Although using LRF data has overcome the shortages of using acceleration and image data, it is still expected to improve the entire system. For these three methods, it is a must for the system that it only reports the types of the road patch which the vehicle has passed. It would be more significant to know the road patch types that the vehicle has not passed but is forthcoming.

Therefore, it is then considered to put another LRF sensor on the roof of the CRUISE that looks forward onto the forthcoming road. It is expected to predict the road terrain types that the vehicle is about to run onto.

CHAPTER VI MULTIPLE-SENSOR BASED ROAD TERRAIN CLASSIFICATION

In the previous chapters, road terrain classification using an accelerometer, a camera and a downward facing LRF have been investigated. However, in the used configurations classification would only report information about the terrain after it was passed. Obviously, driver assistance or autonomous driving, it would be more useful if road terrain type information could be provided before the vehicle moves onto a given road patch. Thus, a forward LRF is now mounted on the roof of the vehicle to achieve the purpose of predicting the upcoming road terrain in advance. THe classification results for the road terrain ahead of the vehicle could be used e.g., for decision making in autonomous driving.

In this chapter, the multiple-sensor based methodology to classify the road terrain types is presented. At first, use a forward looking LRF to predict the forthcoming road terrain is presented. Next, the multiple-sensor based method using MRF algorithm is proposed to improve the classification accuracy. The MRF model and the corresponding energy function are designed and solved. Then multiple-sensor fusion based experiment and the results are presented.

6.1 Predicting LRF Based Probe

Figure 6.1 The mounted forward-looking LRFs

The CRUISE is equipped with three SICK LMS291 LRFs on the roof looking forward with different tilt angles to the ground. As illustrated in Figure 6.1, the middle one with appropriate detect distance is selected for the task. It scans the road surface in a two dimensional plane at a 75 *Hz* sampling rate. This LRF has a 100° field of view with 0.5° angular resolution providing 201 range values per scan. But only the measurements that laser beam scans on the ground are used in this procedure. The scan data is restricted by the approximate width of the vehicle.

The only difference between the LRF on the roof and the one on the frame looking downward (the one used in the Chapter V) is that the former one has a scanning tilt angle to the ground while the latter one scans vertically down to the ground. Except a simple calculation to change the range data into the vertical distance via the known tilt angle, the other data processing works are the same as the LRF data processing procedure in Chapter V. The spatial frequency features is extracted from the data of LMS291 to form the feature matrix which is then trained and tested using the SVM classifier.

As can be seen in Table 6.1, the forward-looking LRF performs well on Gravel road only with a classification accuracy of 94.1%, whereas the other terrain types cannot be classified correctly in many cases. Although the classification accuracy for concrete roads could be seen as acceptable at 78.2%, the performance is very poor for asphalt and grass. The main reason for this is the measurement angle at which the laser beams hit the terrain. The result is that often measurements are of poor quality affecting the classifier's performance.

Table 6.1 The experimental results using forward LRF data

Training / Testing	Asphalt @20,30,40,50	Concrete @20,30	Grass @10	Gravel @20,30
Asphalt @0~70	41214/91791 44.9%	8537/91791 9.3%	8444/91791 9.2%	33596/91791 36.6%
Concrete @0~45	2308/17754 13.0%	13883/17754 78.2%	1563/17754 8.8%	0/17754 0.0%
Grass @0~25	4509/17079 26.4%	2733/17079 16.0%	9308/17079 54.5%	529/17079 3.1%
Gravel @0~35	1561/36312 4.3%	108/36312 0.3%	473/36312 1.3%	34170/36312 94.1%

Average Classification Accuracy: 67.9%

According to the previous experiments using different sensors, it can be seen that each sensor has its own advantages and disadvantages. Therefore, it is intended to find an approach to fuse these multiple sensors to improve the road terrain identification accuracy, especially the accuracy of the LRF that works as a predicting sensor.

Concerning such circumstance, there are four channels that report road terrain results regularly. And those reported labels from different sensors have a common tendency to produce actual labels. Furthermore, for each of the sensors, the current label has a tight relationship with the former one. In other words, the label sequence from a particular sensor tends to keep consistency on the label, because a road terrain the vehicle runs on usually does not change very frequently. All of these enlighten that there must be some kind of relationship between those sensors and between the labels of a particular sensor. In order to fuse these classification results from multiple sensors, the Markov Random Field (MRF) algorithm is considered to express these relationships.

6.2 Markov Random Field

Probabilities play a crucial role in pattern recognition. All the probabilistic inference and learning uses the repeated application of the sum rule and the product rule. It is then proceeded to formulate and solve complicated probabilistic models purely by algebraic manipulation. However, it is highly advantageous to augment the analysis using diagrammatic representations of probability distributions, called probabilistic graphical models [76].

A graph comprises nodes which are connected by links. In a probabilistic graphical model, each node represents one or a group of random variables. The relationships between any of those two variables are expressed by these links. The graph then captures the way in which the joint distribution over all of those random variables. These variables can be decomposed into a product of factors. Each of these factors depends only on a subset of the variables.

Bayesian networks are also known as directed graphical model. The links of the graphs in the model have a particular directionality in these networks. On the other hand,

the other major class of graphical models is MRF that is also known as undirected graphical models. The links in these models do not have directional significance. It is useful for directed graphs to express causal relationships between random variables, while it is better suited for undirected graphs to express soft constraints between random variables.

A factorization of the joint distribution is specified by directed graphical models over numbers of variables into a product of local conditional distributions. These directed graphical models also define numbers of conditional independence properties which should be satisfied by a distribution which factorizes according to the graph. The undirected graphs specify both a factorization and a set of conditional independence relations.

A Markov Random Field [76] has a set of nodes, where each of the nodes corresponds to one or a group of variables. This is similar to a set of links that each connects a pair of nodes. These links are undirected and they do not carry any arrows. To describe the undirected graphs, it is convenient to start with conditional independence properties.

6.2.1 Conditional Independence Properties

It is possible to test if a particular conditional independence property holds by applying a graphical test called d-separation [77]. It involves testing whether or not the paths connecting two sets of nodes are ‘blocked’. The definition of blocked, however, is somewhat subtle due to the presence of paths that have head-to-head nodes. It might be asked if it is possible to define an alternative graphical semantic for probability distributions, so that conditional independence can be determined by the simple graph separation. This is the case and corresponds to undirected graphical models. The

subtleties associated with head-to-head nodes no longer arise if the directionality is removed from the links of the graph. That is to say, the asymmetry between parent and child nodes is removed.

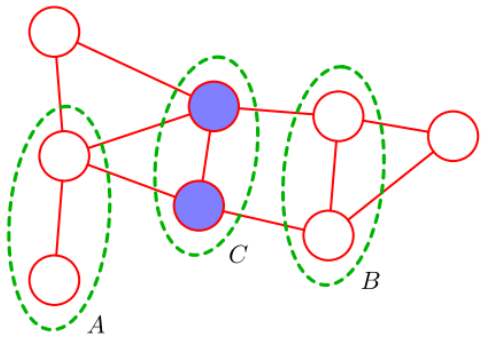


Figure 6.2 An example of an undirected graph

Figure 6.2 shows an example of an undirected graph. Every path from any node in the set *A* to any node in set *B* passes through at least one node in set *C*. Therefore, the conditional independence property that holds for any probability distribution described by this graph is [77]:

$$A \perp\!\!\!\perp B \mid C \tag{6.1}$$

All possible paths connecting nodes in the set *A* to nodes in set *B*, a probability distribution defined by a graph, are tested to check if they satisfy the property (6.1). All such paths will be blocked if all of them pass through one or more nodes in set *C*, then the conditional independence property holds. Nevertheless, the property does not necessarily hold if there is at least one such path that is blocked. More specifically, there will be some distributions that correspond to the graph which do not satisfy this conditional independence relation.

Another way to do the conditional independence test is to remove all nodes along with any of their links in set *C* from the graph. Then if there does not exist at least one path connecting any node in the set A to any node in set B, the conditional independence property must hold.

It is a particularly simple form of the Markov blanket for an undirected graph, as indicated in Figure 6.2, because a node is only dependent on its neighbouring nodes.

6.2.2 Factorization Properties

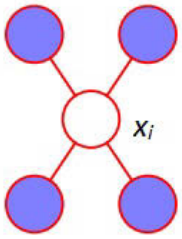


Figure 6.3 A Markov blanket of a node in an undirected graph

Define the joint distribution $p(x)$ as a product of functions [77]. Figure 6.3 shows an undirected graph, the Markov blanket of x_i that consists of the set of neighbouring nodes. The conditional distribution of x_i, which is conditioned on all the remaining variables in the graph, is dependent only on the variables in the Markov blanket.

If these two nodes x_i and x_j which are not connected by a link are considered, then these variables must be conditionally independent given all other nodes in the graph. This conditional independence property can be expressed as [77]:

$$p(x_i, x_j \mid X \setminus \{i, j\}) = p(x_i \mid X \setminus \{i, j\}) p(x_j \mid X \setminus \{i, j\}) \tag{6.2}$$

where $X \setminus \{i, j\}$ denotes the set X of all variables with x_i and x_j removed.

A clique is then introduced, which is defined as a subset of the nodes in a graph, so that there is a link between all pairs of nodes in this subset. Moreover, in a maximal clique it is not possible to include any other nodes from the graph in the set if it is ceasing to be a clique. As shown in Figure 6.4, a four-variable undirected graph presents those concepts. The clique is outlined in green while the maximal clique is marked in blue. This graph then has five cliques and two maximal cliques.

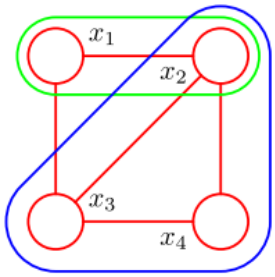


Figure 6.4 A four-node undirected graph

Thus, the factors can be defined in the decomposition of the joint distribution to be functions of the variables (nodes) in the cliques. Then denote a clique by *C* and the set of variables in that clique by X_C. The joint distribution is written as a product of potential functions $\Psi_C(X_C)$ over the maximal cliques of the graph [77]:

$$p(x) = \frac{1}{Z}\prod_C \Psi_C(X_C) \tag{6.3}$$

The quantity Z, also called the partition function, is a normalization constant. It is given by [77]:

$$Z = \sum_X \prod_C \Psi_C(X_C) \tag{6.4}$$

Equation (6.4) ensures that the distribution $p(x)$ given by (6.3) can be normalized properly. It is ensured that $p(x) \geq 0$ by considering only potential functions which satisfy $\Psi_C(X_C) \geq 0$. It should be noted that in order to avoid making a formal connection between conditional independence and factorization for undirected graphs, the potential function $\Psi_C(X_C)$ must be positive. Given this restriction, a precise relationship between factorization and conditional independence can be obtained. Furthermore, since the potential functions must necessarily be positive, to express them as exponentials is more convenient [77]:

$$\Psi_C(X_C) = \exp\{-E(X_C)\} \tag{6.5}$$

Here $E(X_C)$ is called an energy function. The joint distribution is defined as the product of potentials; then the total energy can be obtained by adding the energies of each of the maximal cliques.

On the other hand, the potentials in an undirected graph do not have a specific probabilistic interpretation, compared to the factors in the joint distribution for a directed graph. Although it offers greater flexibility in choosing the potential functions, it does raise the problem of motivating a choice of potential function for a particular application. The reason for that is there is no normalization constraint.

However, it can be accomplished by using the potential function to express which configurations of the local variables are preferred to others. Global configurations which have a relatively high probability can keep a good balance to satisfy the influences of the clique potentials.

6.3 Establishment of MRF Application

An MRF method is applied to fuse a variety of sensor classification results. It is expected that the road terrain classification will improve its accuracy and reliability.

In this case, there are four sensors: a forward looking LRF (LRF1) to the ahead ground to predict the forthcoming road type; an accelerometer (Acc) mounted on the suspension to detect the road type the vehicle is moving on; a camera (Cam); and another LRF (LRF2) downward looking to the road surface that the vehicle has driven through. LRF1 works as predicting sensor to report the upcoming road type while Acc, Cam and LRF2 work as posterior detectors. Although each sensor reports a classification result regularly, the reported labels from them are not certain to be the same. That is because the detection capabilities of the sensors are different. Furthermore, it is expected that the system eventually reports the forthcoming road types rather than reporting road types after the vehicle has moved on.

6.3.1 Nodes in MRF

An undirected graph model is proposed to the observed classification labels from each sensor of the whole system. As shown in Figure 6.5, the MRF is composed of 5 node types: the purple, green, yellow, and blue nodes are variables *y, u, w, and v* which denote the classification results from LRF1, Cam, Acc, and LRF2, respectively. The red nodes are unobservable variables *x* that needs to be 'reset' for the final classification results of LRF1. This reset value is the new value that needs to be assigned using this MRF algorithm. It actually indicates the final terrain types produced by the whole system. Index $i = 1,\ldots, n$ runs over all classification results. Specifically, Index *i* is the current observed node that needs to be reset while *i-1* is the former neighbour.

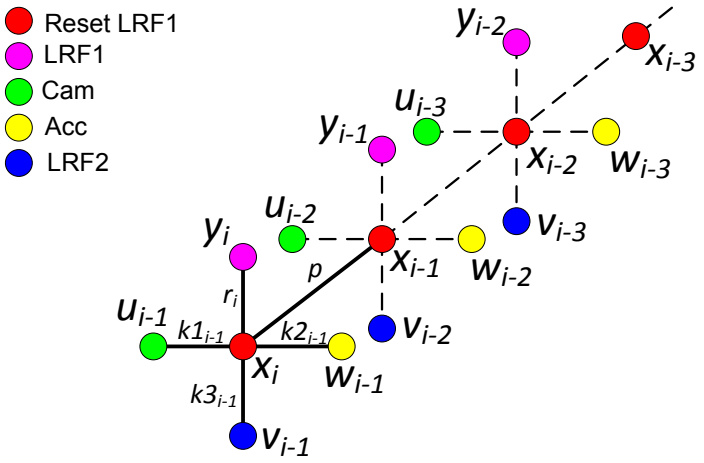


Figure 6.5 MRF Model for road terrain identification

6.3.2 Variable Values of Nodes in MRF

The variable values of the nodes shall be defined for purposes of energy function computation. They should reflect the characteristics of different classes and also have the ability to compute energy function. As the road surface spatial frequency analysis in Chapters III and V reveal, there are gradual changes from asphalt road to a gravel road. Furthermore, if these four road types are arranged in order of asphalt, concrete, grass, and gravel, each has similar characteristics as its neighbour(s). Therefore, the value of each variable in the MRF algorithm is set as in Table 6.2.

Table 6.2 The values of variables

Class	Value
Asphalt road type	1
Concrete road type	2
Grass road type	3
Gravel road type	4

As in many prior tests, the values do not work well if they are non-integers, because it is hard to achieve 'equal power' for each class in the energy function. It does not work as well if some of them are negative while others are positive. The reason for this is the number of the classes is even but not odd which would present 'balance power' in the energy function. Nevertheless, instinctively, the sequential positive integer numbers at least would not lead to an MRF result which is immoderate far from the real value if it is wrong.

6.3.3 Clique Potentials in MRF

The mode of the probability distribution defined by the MRF is designed to produce a final road type detection result by fusing the classification results of multiple sensors. The intuition behind the MRF is that the reset label x_i is strongly correlated with its corresponding observed label y_i; three other reported sensor labels; and correlated between its prior reset label x_{i-1}. As can be seen from Figure 6.5, the five clique potentials are defined as:

- r_i describes the strength correlation between the reset label x_i and its corresponding observed label y_i;
- $k1_{i-1}$, $k2_{i-1}$, and $k3_{i-1}$ describe the correlation between x_i and u_{i-1}, x_i and w_{i-1}, x_i and v_{i-1};
- p describes the correlation between x_i and x_{i-1}.

It should be noted that the clique potentials of $k1_{i-1}$, $k2_{i-1}$, and $k3_{i-1}$ of the cliques are between x_i and u_{i-1}, between x_i and w_{i-1}, and between x_i and v_{i-1}, but not u_i, w_i, and v_i. This is because to detect the same patch of road surface type, Com, Acc and LRF2 act as posterior detectors that the index *i-1* of theirs are the index *i* of LRF1 observations.

Therefore, to aim at the same patch of the road, the links between x_i and u_{i-1}, between x_i and w_{i-1}, and between x_i and v_{i-1} are built.

6.3.4 Values of Clique Potentials in MRF

The values of the five clique potentials p, r_i, $k1_{i-1}$, $k2_{i-1}$, and $k3_{i-1}$ indicate the correlations between x_i that is going to be reset and other labels that help to correct the wrong x_i labels. These weighting factors provide the links between central node x_i to other nodes. However, those links are not supposed to devote equal weights to the central node. Each link should devote information depending on how 'confident' it is.

The probability of each label resulted from SVM classifier provides the confidence needed. As is known, each classified label produced by the SVM classifier has a probability value. It indicates the confidence in its classification result. Therefore, the classification result probability is utilized as the weighting factor to each clique potential in the MRF energy function. It should be noted that the index of each clique potential needs to be the same as the corresponding node's index. Additionally, besides r_i, $k1_{i-1}$, $k2_{i-1}$, and $k3_{i-1}$ which have particular indices to set the values, p that refers to the link between current reset nodes and its former one, is set as constant 1. The reason is that the current node is strongly relevant to the former one that has already reset by the MRF algorithm.

6.3.5 Energy Function

An energy function (potential function) is supposed to be an arbitrary, nonnegative function over a maximal clique. It can be multiplied by any nonnegative functions of subsets of the clique, or equivalently which the corresponding energies can be added in. In this case, this allows setting the particular probability to multiply its corresponding label

difference measurement. Specifically, referring to Figure 6.5, the MRF is defined through the following potentials:

1) The clique between the reset classification label (Reset LRF1) and its former label:

$$\Psi = \rho \sum_i \sqrt{\left| x_i^2 - x_{i-1}^2 \right|} \tag{6.6}$$

2) The clique between the reset classification label (Reset LRF1) and its observed classification label (LRF1):

$$\Phi = \sum_i r_i \sqrt{\left| x_i^2 - y_i^2 \right|} \tag{6.7}$$

3) The cliques between the reset classification label (Reset LRF1) and its corresponding other three classification labels (Cam, Acc, LRF2), respectively:

$$\Lambda = \sum_i k1_{i-1} \sqrt{\left| x_i^2 - u_{i-1}^2 \right|} + \sum_i k2_{i-1} \sqrt{\left| x_i^2 - w_{i-1}^2 \right|} + \sum_i k3_{i-1} \sqrt{\left| x_i^2 - v_{i-1}^2 \right|} \tag{6.8}$$

The energy function of of the model in this case takes the form:

$$E(x, y, u, w, v) = \Psi + \Phi + \Lambda \tag{6.9}$$

which defines a joint distribution over x, y, u, w, v given by:

$$p(x, y, u, w, v) = \frac{1}{Z} \exp(-E(x, y, u, w, v)) \tag{6.10}$$

where Z is a normalization constant (partition function):

$$Z = \sum_{x, y, u, w, v} \exp(-E(x, y, u, w, v)) \tag{6.11}$$

6.3.6 Optimization

For the purpose of correctness for the observed classification labels of the forward LRF, the reset labels x having a high probability should be found, ideally the maximum probability. This requires the energy function $E(x, y, u, w, v)$ has the lowest value. To accomplish this in on-line computation, an iterative technique called iterated conditional modes [78] is used, which is an application of coordinate-wise gradient ascent.

This method first initializes the variables $\{x_i\}$, which are set to $x_i = y_i$. Then the current observed node x_i is selected and the total energy for the four possible states $x_i = 1$, $x_i = 2$, $x_i = 3$, and $x_i = 4$ is evaluated, keeping all other node variables fixed. Then x_i is set to the state with the lowest energy. This will either leave the probability unchanged, if x_i is unchanged, or to lower or increased energy. Because only one variable is changed, this is a computationally cheap local computation. By implementing this MRF algorithm, the classification results of the forward LRF are expected to be improved.

6.4 Experiment

6.4.1 Experimental Platform

As before, the experimental platform CRUISE is equipped with a number of sensors which are four wheel encoders, an accelerometer, a camera, and two LRFs (one downward-looking and the another forward-looking for prediction). As shown in Figure 3.16, two PC104 computers and a laptop are employed for data collection and computing. In the multi-sensor fusion experiment, all of these sensors' real time data are fused using the MRF algorithm for road terrain classification.

6.4.2 Multiple-Sensor Fusion Based Experiment

The multiple-sensor fusion experiment is then performed using all the labels produced by the accelerometer, the camera, and the downward LRF, and the forward LRF. The MRF algorithm and its relevant computation are implemented. The actual road types are marked manually during the experiment as a baseline.

6.5 Experiment Results

The system is designed to report road types at regular distance intervals (4 meters, approximate length of the vehicle) past. Because different sensors have very different reporting frequencies, the classification label series are divided into groups by the fixed travelling length (4 meters). Then a simple Majority Vote (MV) [79] algorithm is applied to reduce them into one resulting label for each group. MV algorithm is a decision calculation that selects alternatives which have a majority that occupies the most quantity among the alternatives. In this case, every group of labels has numbers of classification results marked as “1”, “2”, “3”, and “4”. For example, MV will produce a “1” if a group has most quantity of “1” than “2”, “3”, or “4”. As a result, each label group can be turned into only one label instead of plenty of them in order to be ready for the MRF algorithm calculation.

Table 6.3 The numbers of samples before and after the MV

	Accelero -meter	Forward LRF	Camera	Downward LRF	After MV
Asphalt	3004	4481	94488	61194	3004
Concrete	149	282	5924	3950	149
Grass	220	813	17074	11386	220
Gravel	666	1749	36916	24208	666
Total	4039	7325	154402	100738	4039

As can be seen from Table 6.3, the classification labels from two LRFs and Camera are MV processed to reduce the numbers of each class. As shown in the last column of Table 6.3, the numbers of the labels from these three sensors are matched to those of the accelerometer. Thus, all the classification results can be matched to each other, so that each sensor only provides one classification result for each road patch (4 meters long). Eventually, this allows the MRF algorithm to process those classification label series to correct and improve the results of one of the sensors. Hereby, as mentioned, it is set as the forward-looking LRF to predict the road terrain types of forthcoming road.

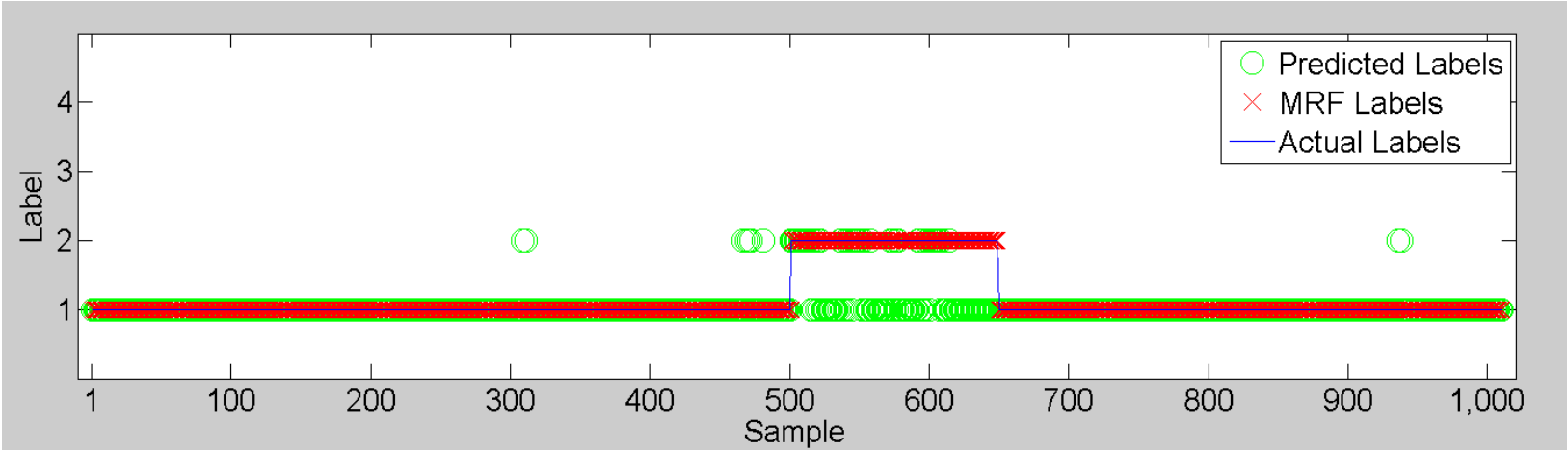


Figure 6.6 (a) MRF experiment label sequence: segment 1

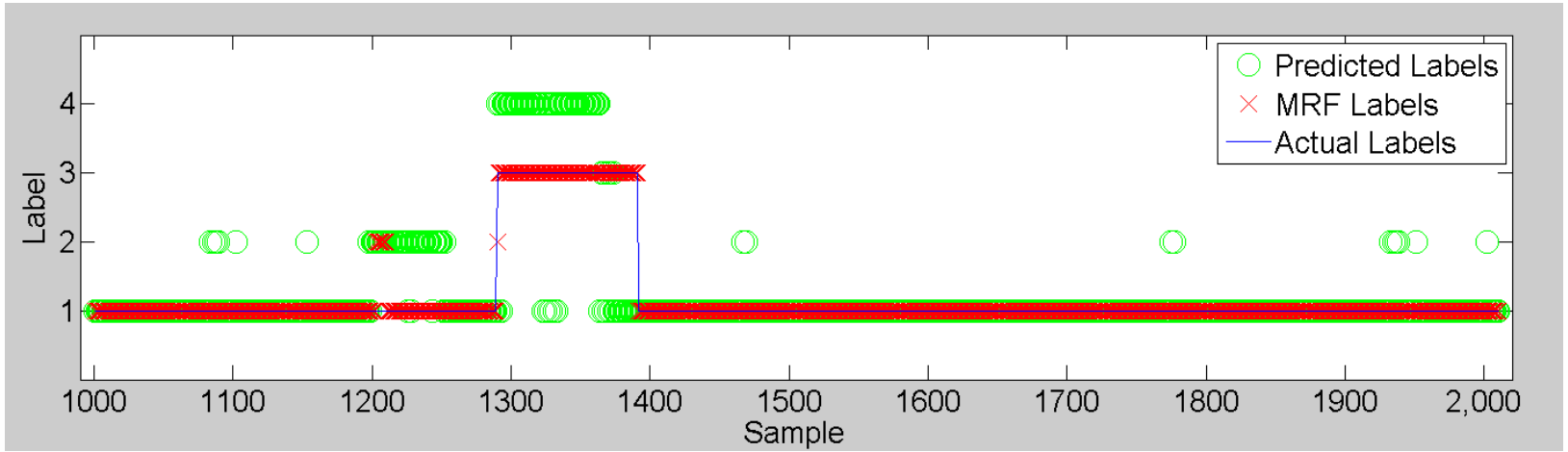


Figure 6.6 (b) MRF experiment label sequence: segment 2

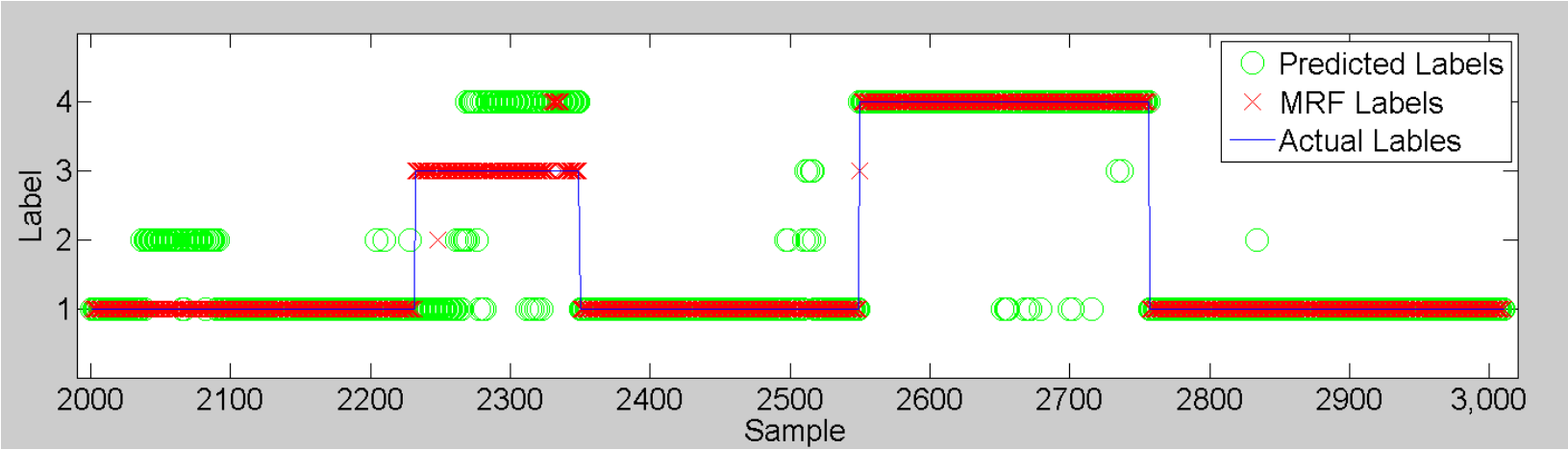


Figure 6.6 (c) MRF experiment label sequence: segment 3

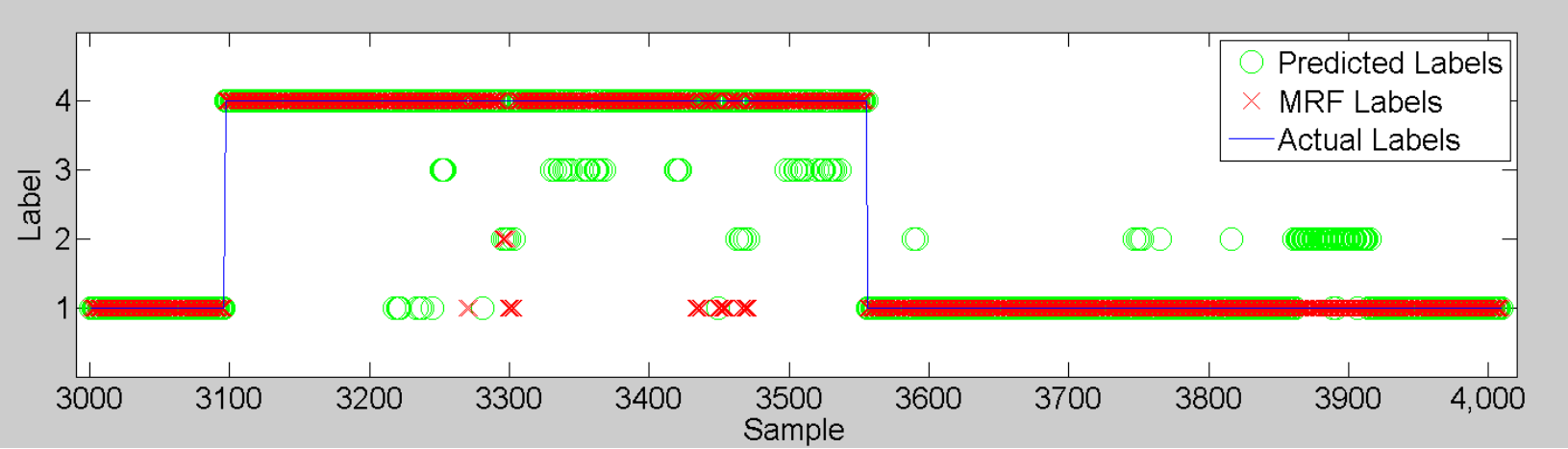


Figure 6.6 (d) MRF experiment label sequence: segment 4

Figures 6.6 show the online experiment results using the MRF algorithm. It indicates how the classification results of the forward-looking LRF are corrected by fusing Multi-sensor classification results. The experiment is split into four parts in Figures 6.6 (a), (b), (c), and (d). In Figures 6.6, the horizontal axis shows the index of the labels while the vertical axis shows the labels that can be 1, 2, 3, and 4 referring to Asphalt, Concrete, Grass, and Concrete road, respectively. The actual road type labels are represented by a continuous blue line. The green circles refer to the predicted labels of the forward-looking LRF. The red crosses are the MRF labels that are the final predicted results of the forthcoming road.

It can be seen that some of the misclassified predicted labels (green circles) are replaced by the MRF labels (red crosses) effectively. Especially for sample index 510-650 in Figure 6.6 (a) where most of the wrong predicted labels are corrected (to be Concrete road), and the instances of the sample index around 1290-1390 in Figure 6.6 (b) and index of 2230-2350 in Figure 6.6 (c) where the Grass road is relabeled.

However, the MRF algorithm cannot always produce very accurate results. Some of the misclassified labels cannot be corrected by this algorithm. For instance, some labels at sample index 1210 in Figure 6.6 (b) and some at index 2340 in Figure 6.6 (c). Due to the misclassification, labels of the neighbouring nodes are immoderate ‘powerful’ and hence the energy function leads to incorrect predicting labels. So this prevents those current nodes to be reset properly.

Moreover, as illustrated by sample index 3305, 3440, and 3475, the new reset labels turn out very different results that are neither the predicted labels of the forward LRF nor the actual ones. Fortunately, this rarely happens. The reason for this is these wrong reset labels are affected by their former labels.

In addition, this method may have a long transition when the vehicle moves from one type of terrain to another. The reset labels around the transition points could be wrong. For instance, when it changes from terrain 1 to terrain 3 around sample index 1290 in Figure 6.6 (b), the MRF label produced is 2 where it should be 1 or 3; and from terrain 1 to terrain 4 at sample index 2550 in Figure 6.6 (c), the MRF label is 3 where it should be 1 or 4. The reason for this is that the reset node (MRF label) in this MRF model is correlated with the former classified labels of the sensors.

Table 6.4 The comparison after using MRF multiple-sensor fusion

Training / Testing	Asphalt @20,30,40,50	Concrete @20,30	Grass @10	Gravel @20,30
Asphalt @0~70	44.9% →99.7 %	9.3% →0.3%	9.2% →0.0%	36.6% →0.0%
Concrete @0~45	13.0% →0.7%	78.2% →99.3%	8.8% →0.0%	0.0% →0.0%
Grass @0~25	26.4% →0.0%	16.0% →0.9%	54.5% →95.5%	3.1% →3.6%
Gravel @0~35	4.3% →3.7%	0.3% →0.6%	1.3% →0.2%	94.1% →95.5%

Average Classification Accuracy: 67.9% →97.5%

Table 6.4 shows the final statistical results and the improvements of the forward-looking LRF based on multi-sensor fusion. Referring to Table 6.1, the classification accuracies of all four road types have been increased. The Asphalt road and Concrete road, which are very hard to distinguish, have been classified very well. The performances on Asphalt road and Grass road have been significantly increased by approximately 50%

points. The performance on Gravel road has been slightly increased even though it could be classified well already before MRF fusion was used. The average classification accuracy is consequently increased from 67.9% to 97.5%. Hence, the forward-looking LRF is now able to accomplish the task of road terrain classification. Its poor prediction accuracy can be significantly improved by employing multi-sensor classification results using the MRF algorithm. Moreover, this classification accuracy 97.5% (multiple-sensor based using MRF algorithm) is higher than each of 69.4% (acceleration based), 88.8% (image based), and 82.3% (downward LRF based), respectively. This indicates that none of the solo sensor produces better classification accuracy than the multiple-sensor method.

6.6 Conclusion

In this chapter, an MRF algorithm has been presented which fuses the multiple sensors' classification results to improve the forward LRF for prediction the forthcoming road terrain types.

A LRF was put on the roof of CRUISE to scan the road surface ahead of the vehicle attempting to predict the road terrain types where the vehicle was about to move on. Unfortunately, it did not work well, especially on the most common type-asphalt road.

A multiple-sensor fusion method was then presented to improve the classification accuracy of the forward LRF. The MRF algorithm was employed to fuse all the sensors' classification results. The final prediction results of the forward LRF were relevant with the accelerometer, the camera, the downward LRF, and its current observation and its former observation. These five relationships were reflected by the five clique potentials while the five classification results were defined as the five nodes of the MRF model (see Figure 6.5). The optimization was computed using the created energy function. The comparison results between the predicting LRF and its corresponding MRF's showed that the MRF

multiple-sensor fusion method was extremely effective and robust to accomplish the task of the road terrain classification.

CHAPTER VII

CONCLUSION AND FUTURE DIRECTION

7.1 Conclusion

The aim of this research is to develop a methodology to identify impending road terrain types by using on-board sensors. An accelerometer, camera, and LRF were employed and investigated respectively. However, all of these sensors have their own advantages and disadvantages. To improve the classification accuracy of the LRF that works for predicting the forthcoming road terrain, an MRF multiple-sensor fusion method was then proposed.

The first sensor tested was the accelerometer which was mounted on the suspension of the vehicle to measure the vibration of the vehicle component. With the speed of the vehicle and one quarter dynamic model of the vehicle, the road profile was calculated. By prior training and testing, spatial frequency features and SVM classifier were selected for the experiment. The acceleration based classification result showed that the method did not work well on an asphalt road which is the most common road in urban areas. Further, the training data set that contained insufficient speed data led to lower classification accuracy.

The second sensor tested was the camera capturing the images of the road surface that the vehicle had driven on. Texture features were extracted from those images for the road terrain classification. The image based classification accuracy was higher than that of acceleration based classification. The obvious weakness of this method was that cameras

only work in well illuminated environments. Speeding vehicles or night time lightning could degrade the quality of the image and hence classification accuracies.

The third sensor tested was the downward LRF scanning the road surface that the vehicle had travelled on. The road surface was reconstructed using laser scanning points along with the speed of the vehicle. The spatial frequency features were extracted again for the LRF based road terrain classification. The experiment results showed that it had less effect when the speed changed. And LRF based method worked better on the asphalt road than acceleration based method whereas worse on the other three types than image based method.

It has been found that each sensor has its own usage limitation and shortcomings. Apart from that, they were not able to predict the forthcoming road terrain types, which was very meaningful for a driving assistant system or an autonomous vehicle.

The forward LRF was then employed. It was mounted on the roof of the vehicle to scan the surface of the forthcoming road. The tested result showed that it did not work well because of the long displacement and range noises. The method of multiple-sensor fusion based on MRF algorithm was proposed to improve the accuracy of road terrain classification.

An MRF model that contained five types of nodes and five cliques which described the relationships between the classification results of the accelerometer, the camera, and the two LRFs was designed. By defining the variable values of nodes and the values of clique potentials, the energy function in this case was then created. The optimization computing method was introduced to solve the energy function. The reset labels of forward LRF showed that fusing the classification results of other three sensors and itself,

the classification accuracy of the predicting LRF could be significantly improved. Therefore, the MRF multiple-sensor fusion method was effective and feasible in the task of road terrain classification.

7.2 Future Direction

The research presented in this thesis has accomplished some tasks but meanwhile also raises more questions. There are several issues of this research arising from this work which can be pursued.

Although the fusion of the four sensors was designed in this thesis, there could be some other sensors involved or replacements. For example, a camera producing greater image quality mounted on the roof of the vehicle could be utilized meanwhile removing the downward one. This camera could be the second sensor that predicts the forth-coming road terrain besides the forward LRF.

Although the experiments in this research are off-line tested, the real-time on-line performances are still incomplete. There is more challenging research needs to be done to transfer software into a different application environment. In order to bring research into a real application, this indispensable work has to be on-line implemented.

REFERENCES

[1] N. Chakraborty, R. Singh Patel, "Intelligent Agents and Autonomous Cars: A Case Study", *International Journal of Engineering Research and Technology (IJERT)*, Vol 2, No 1, pp. 1-7, January 2013.

[2] S. Thrun, "Toward Robotic Cars", *Communications of the ACM*, vol 53, No 4, pp. 98-106, April 2010.

[3] S. Milosavljevic, F. Bergman, B. Rehn, A. B. Carman, "All-Terrain Vehicle Use in Agriculture: Exposure to Whole Body Vibration and Mechanical Shock", Apploed Ergonomics, Vol 41, No 4, pp. 530-535. July 2010.

[4] B. Ghotbi, "Vehicle-Terrain Interaction Models for Analysis and Performance Evaluation of Wheeled Rovers", *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, pp. 3138-3143, 7th-12th, October 2012.

[5] Z. Bing, "Ride Comfort Research of Off-Road Vehicle Based on Soft Terrain", *IEEE 11th international Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD)*, Yiwu, China, pp. 579-584, 17th-19th, November 2010.

[6] G. Koch, T. Kloiber, "Diving State Adaptive Control of an Active Vehicle Suspension System", IEEE Transactions on Control System Technology, Accepted. Published in 2013.

[7] K. Alipour, S. Moosavian, "Postural Stability of Wheeled Mobile Manipulators with Flexible Suspension Considering Tire Friction Model", *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Singapore, pp. 764-769, 14th-17th, July 2009.

[8] S. Biswas, “Design of Steering System Through Optimization of TCD and Steering Effort”, Frontiers in Automobile and Mechanical Engineering (FAME), Chennai, India, pp. 356-362, 25th-27th, November 2010.

[9] K. Iagnemma, S. Dubowsky, “Mobile Robot Rough-Terrain Control (RTC) for Planetary Exploration”, *ASME Biennial Mechanisms and Robotics Conference,* DETC, 2000.

[10] D. Sadhukhan, “Autonomous Ground Vehicle Terrain Classification Using Internal Sensors”, Master Thesis, Florida State University, Tallahassee , US, 16th March 2004.

[11] B. Delong, 4-Wheel Freedom: The Art of Off-Road Driving, Colorado: Paladin Press, 2000.

[12] G. Ishigami, G. Kewlani, K. Iagnemma, “Statistical Mobility Prediction for Planetary Surface Exploration Rovers in Uncertian Terrain”, IEEE International Conference on Robotics and Automation (ICRA), Anchorage, US, pp. 588-593, 3rd-7th, May 2010.

[13] T. P. Setterfield, A. Ellery, “Terrain Response Estimation Using an Instrumented Rocker-Bogie Mobility System”, IEEE Transactions on Robotics, Vol 29, No 1, November 2012, pp. 172-188.

[14] D. Gingras, T. Lamarche, J. Bedwani, E. Dupuis, “Rough Terrain Reconstruction for Rover Motion Planning”, Canadian Conference on Computer and Robot Vision (CRV), Ottawa, Canada, pp. 191-198, 31st May - 2nd June 2010.

[15] H. Xu, Z. Zhang, Y. Wu, “Contact Angle Estimation Based on Kinematics Modeling Analyses for Rover with Caster and Camber”, IEEE International Conference on Robotics and Biomimetics (ROBIO), Tianjin, China, pp. 137-142, 14th-18th, December 2010.

[16] K. Iagnemma, “On-line Terrain Parameter Estimation for Planetary Rovers”, *IEEE International Conference on Robotics and Automation*, Vol. 3, Washington DC, US, pp. 3142-3147, 11th-15th, May 2002.

[17] K. Iagnemma, S. Kang, H. Shibly, and S. Dubowsky, “Online terrain parameter estimation for wheeled mobile robots with application to planetary rovers”, *IEEE Trans. Robo,* Leuven, Belgium, Vol. 20, No 5, pp.921-927, October 2004

[18] R. Jitpakdee, B. Thonburi, T. Maneewarn, “Neural Networks Terrain Classification Using Inertial Measurement Unit for an Autonomous Vehicle”, SICE Annual Conference, Tokyo, Japan, pp. 554-558, 20th-22nd, August 2008.

[19] Y. Gonthier, E. Papadopoulos, “On the Development of a Real-Time Simulator for an Electro-Hydraulic Forestry Machine”, IEEE International Conference on Robotics and Automation, vol. 1, pp. 127-132, 16th-20th, May 1998.

[20] A. González, J. O. Eugene, Y. Li, K. Cashell, “The Use of Vehicle Acceleration Measurements to Estimate Road Roughness”, *Vehicle System Dynamics*, Vol. 46, No 6, June 2008, pp. 483-499.

[21] C. C. Ward, K. Iagnemma, “Speed-independent vibration-based terrain classification for passenger vehicles”, *Vehicle System Dynamics*, vol. 47, No 9, September. 2009, pp. 1095-1113.

[22] C. Weiss, H. Fröhlich, A. Zell, “Vibration-based Terrain Classification Using Support Vector Machines,” *IEEE/RSJ Internatioanl Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, pp. 4429-4434, 9th-15th, October, 2006.

[23] D. Sadhukhan, C. Moore, “Online Terrain Estimation Using Internal Sensors”, *Florida Conference on Recent Advances in Robotics*, Boca Raton, US, 2003.

[24] L. Ojeda, J. Borenstein, G. Witus, R. Karlsen, “terrain characterization and classification with a Mobile Robot”, *Journal of Field Robotics*, Vol. 23, No 2, pp. 103-122, 8th, January 2006.

[25] E. M. DuPont, R. G. Robers, C. A. Moore, “The Identification of Terrains for Mobile Robots Using Eigenspace and Neural Network Methods”, *Florida Conference on Recent Advances in Robotics (FCRAR)*, Miami, US, 25th-26th, May 2006.

[26] C. Weiss, M. Stark, A. Zell, “SVMs for vibration-based terrain classication”, *Autonomoe Mobile System (AMS)*, Fachgespräch, Germany, pp.1-7, 18th-19th, October, 2007.

[27] C. F. Olson, L. H. Matthies, M. Schoppers, M. W. Maimone, “Robust Stereo Ego-Motion for Long Distance Navigation”, *IEEE Conference on Computer Vision and Pattern Recognition*, Hilton Head Island, US, Vol. 2, pp. 453-458, 13th-15th, June 2000.

[28] A. Howard, H, Seraji, “Vision-based terrain characterization and traversability assesment”, *Journal of Robotic Systems*, Vol. 18, No 10, September 2001, pp. 577-587.

[29] D. Thompson, T. Smith, D. Wettergreen, “Data Mining During Rover Traverse: From Images to Geologic Signatures”, *8th Internaltional Symposium on Articial Intelligence Robotics and Automation in Space (ISAIRAS)*, September 2005.

[30] R. Castano, R. Manduchi, J. Fox, “Classication Experiments on Real-World Textures”, *Workshop on Empirical Evaluation in Computer Vision*, Kauai, USA, 10th, Decmber 2001, pp. 1-6.

[31] R. E. Karlsen, G. Witus, “Terrain Understanding for Robot Navigation”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, San Diego, USA, pp. 895-900, 29th October- 2nd November, 2007.

[32] J. Chetan, K. M. Krishna, C. V. Jawahar, “An Adaptive Outdoor Terrain Classification Methodology using Monocular”, *IEEE/RSJ Internaltional Conference on Intelligent Robots and Systems (IROS)*, Taipei, China, pp. 766-771, 18th-22nd, October 2010.

[33] D. Song, C. Yi, I. H. Suh, B. Choi, “Self-Supervised Terrain Classification Based on Moving Objects Using Monocular Camera”, *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Phuket, Thailand, pp. 527-533, 7th-11th, December 2011.

[34] S. Willians, A. M. Howard, “Towards Visual Arctic Terrain Assessment”, *Springer Tracts in Advanced Robotics*, Vol. 62, pp. 91-100, 2010.

[35] K. Nishiwaki, J. Chestnutt, S. Kagami, “Autonomous navigation of a humanoid robot over unknown rough terrain using a laser range sensor”, International Journal of Robotics Research, Vol. 31, No 11, pp. 1251-1262, September 2012.

[36] J. Macedo, R. Manduchi, L. Matthies, “Lada-Based Discrimination of Grass from Obstacles for Autonomous Navigation”, *ISER*, Honolulu, US, 2000.

[37] M. Montemerlo, S. Thrun, “Muti-Resolution Pyramid for Outdoor Robot Terrain Perception”, *19th Natioanl Conference on Artifical Intelligence*, San Jose, US, pp. 464-469, 25th-29th, July, 2004.

[38] J. C. Andersen, M. R. Blas, O. Ravn, N. A. Andersen, M. Blanke, “Traversable Terrain Classification for Outdoor Autonomous Robots using Single 2D Laser”, *Integrated Computer-Aided Engineering*, Vol. 13, pp. 223-232, 2006.

[39] N. Vandapel, D. F. Huber, A. Kapuria, M. Hebert, “Natural Terrain Classification using 3D Lidar Data”, IEEE International Conference on Robotics and Automation, Vol. 5, pp. 5117-5122, 26th, April-1st, May 2004.

[40] J. Lalonde, N. Vandapel, D. Huber, and M. Hebert, “Natural Terrain Classification using Three-Dimensional Ladar Data for Ground Robot Mobility”, Journal of Field Robotics, Vol. 23, No 10, pp. 839-861, November 2006.

[41] C. Urmson, J. Anhalt, M. Clark, T. Galatali, J. P. Gonzalez, J. Gowdy, A. Gutierrez, S. Harbaugh, M. Johnson-Roberson, H. Kato, P. Koon, K. Peterson, B. Smith, S. Spiker, E. Tryzelaar, W. Whittaker, “High Speed Navigation of Unrehearsed Terrain: Red Team Technology for Grand Challenge 2004 (Tech. Rep. CMU-RI-TR-04-37)”, Pettsburgh, PA: Robotics Istitute, Carnegie Mellon University, 2004.

[42] S. Golberg, M. Maimone, L. Matthies, “Stereo Vision and Rover Navigation Software for Planetary Exploration”, *IEEE Aerospace Conference*, Vol. 5, pp. 2025-2036, 9th-16th, March 2002.

[43] C. Urmson, M. Dias, R. Simmons, “Stereo Vision Based Navigation for Sun-Synchronous Exploration”, *Conference on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland, September 2002.

[44] D. Stavens, S. Thrun, “A Self-Supervised Terrain Roughness Estimator for Off-Road Autonomous Driving”, *22nd Conference on Uncertainty in Artificial Intelligence (UAI2006)*, Cambridge, US, 13th-16th, July 2006.

[45] K. Iagnemma, S. Kang, C. Brooks, S. Dubowsky, “Multi-Sensor Terrain Estimation for Planetary Rovers”, *International Symposium on Artificial Intelligence, Robotics, and Automation in Space*, Nara, Japan, 2003.

[46] I. Halacti, C. A. Brooks, and K. Iagnemma, “Terrain Classication and Classier Fusion for Planetary Exploration Rovers”, *IEEE Aerospace Conference*, Big Sky, US, pp. 1-9, 3rd-10th, March 2007.

[47] C. A. Brooks, K. Iagnemma, “Self-Supervised Terrain Classification for Planetary Rovers”, *IEEE Aerospace Conference*, Big Sky, US, pp. 1-9, 3rd-10th, March 2007.

[48] C. Weiss, H. Tamimi, “A. Zell, A Combination of Vision and Vibration-based Terrain Classification”, *IEEE/RSJ Int. Conf. On Intelligent Robots and Systems (IROS)*, Nice, Frence, pp. 2204-2209, September 2008. 19th-23rd, May 2008.

[49] E. G. Gollins, E. J. Coyle, “Vibration-Based Terrain Classification Using Surface Profile Input Frequency Responses”, *IEEE International Conference on Robotics and Automation*, Pasadena, US, pp. 3276-3283, 19th-23rd, May 2008.

[50] L. Lu, C. Ordonez, E. G. Collins, Jr. E. M. DuPont, “Terrain Surface Classification for Autonomous Ground Vehicles Using a 2D Laser Stripe-Based Structured Light Sensor”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, US, pp. 2174-2181, 10th-15th, October 2009.

[51] R. Manduchi, “Obstacle Detection and Terrain Classsification for Autonomous Off-Road Navigation”, *Journal of Autonomous Robots*, Vol. 18, No 1, pp. 81-102, January 2005.

[52] I. Posner, D. Schroeter, P. Newman, “Describing Composite Urban Workspaces”, IEEE International Conference on Robotics and Automation, Roma, Italy, pp. 4962-4968, 10th-14th, April 2007.

[53] J. Wong, Theory of Ground Vehicles, 4th ed., vol 4, New York: John Wiley & Sons, 2008.

[54] D. F. Watson, G. M. Philip, “Triangle Based Interpolation”, *Mathematical Geology*, Vol. 16, No 8, October 1984.

[55] P.D. Welch, “The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms”, *IEEE Trans*. Audio Electroacoust. Vol. AU-15, 1967, pp. 70-73.

[56] I. Daubechies, “Ten Lectures on Wavelets”, *Society for Industrial and Applied Mathematics*, PA, USA, 1992.

[57] D. F. Mix and K. J. Olejniczak, “Elements of Wavelets for Engineers and Scientists, New Jersey”, USA, John Wiley and Sons Inc., 2003.

[58] S. Mallat, “A Theory of Multiresolution Signal Decomposition: The Wavelet Representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No 7, July 1989, pp. 674-693.

[59] S. Mallat, “A Wavelet Tour of Signal Processing: The Sparse Way”, San Diego, USA, Academic Press, Elsevier, 2009.

[60] C. W. Hsu, C.C. Chang, C. J. Lin, “A Practical Guide to Support Vector Classification”, Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, 2003.
http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

[61] I. Jolliffe, “Principal Component Analysis”, 2nd ed., vol 5, New York: Springer-Verlag, 2002.

[62] R. Kohav, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection”, *Appears in the International Joint Conference on Artificial Intelligence (IJCAI)* , Quebec, Canada, 20th-25th, August, 1995, pp. 1137-1143.

[63] R.R. Picard, R. D. Cook, “Cross Validation of Regression Models”, *Journal of the American Statistical Association*, vol. 79, No 387, spetember 1984, pp. 575-583.

[64] Remco R. Bouckaert, Eibe Frank, Mark A. Hall, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten, “WEKA-Experiences with a Java Open-Source Project”, *Journal of Machine Learning Research*, vol. 11, No 2010, October 2010, pp. 2533-2541.

[65] I. J. Good, The Estimation of Probabilities: An Essay on Modern Bayesian Methods. MIT Press, 1965.

[66] D. D. Lewis, “*Naïve (Bayes) at Forty: The Independence Assumption in Information Retrieval*”, 10th European Conference on Machine Learning (ECML), Chemnitz, Germany, 21st-24th, April, 1998, pp. 4-15.

[67] G. P. Zhang, “Neural Networks for Classification: A Survey”, *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applictions and Reviews*, vol. 30, No 4, November, 2000, pp. 451-462.

[68] M. S. Hung, M. Y. Hu, M. S. Shanker, B. E. Patuwo, “Estimating Posterior Probabilities in Classification Problems with Neural Networks”, *International Journal of Computational Intelligence and Organizations*, vol. 1, No 1, January, 1996, pp. 49-60.

[69] S. Miyake, F. Kanaya, “A Neural Network Approach to a Bayesian Statistical Decision Problem”, *IEEE Transactions on Neural Networks*, vol. 2, No 5, September, 1991, pp. 538-540.

[70] V. Vapnik, The Nature of Statistical Learning Theory, 2nd Edition, Springer, New York, November, 1999.

[71] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machine, Cambridge University Press, Cambridge, 2000.

[72] P. Brodatz, “Texture: A Photographic Album for Artists and Designers”, New York, USA, Dover Publications, 1966.

[73] R. M. Haralick, K. Shanmugam, I. Dinstein, “Textural Features for Image Classification”, *IEEE Transactions on Systems, Man and Cybernetics,* Vol. SMC-3, No. 6, November, 1973, pp. 610-621.

[74] L. Soh, C. Tsatsoulis, “Texture Analysis of SAR Sea Ice Imagery Using grey Level Co-Occurrence Matrices”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 2, March, 1999, pp. 780-795.

[75] D. Clausi, “An Analysis of Co-Occurrence Texture Statistics as a function of grey Level uantization”, *Reote Sensing*, Vol. 28, No. 1, 2002, pp. 45-62.

[76] R. Kindermann, J. L. Snell, “Markov Random Fields and Their Applications”, *Americal Mathematical Society*, 1980.

[77] C. M. Bishop, “Pattern Recognition and Machine Learning”, New York: Springer, 2006.

[78] J. Kittler, J. Föglein, “Contextual Classification of Multispectral Pixel Data”, *Image and Vision Computing*, Vol. 2, Issue 1, February, 1984, pp. 13-29.

[79] E. Bauer, R. Kohavi, “An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants”, *Machine Learning*, Vol. 36, 1/2, July/August, 1999, pp. 105-139.