SHIZUOKA UNIVERSITY

DOCTORAL THESIS

Traversability Cost Prediction of Outdoor Terrains for Mobile Robot Using Image Features

Author:
Mohammed A. BEKHTI

Supervisor:
Pr. Yuichi KOBAYASHI

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Kobayashi Laboratory
Department of Information Science and Technology
Graduate School of Science and Technology

July 17, 2020

Declaration of Authorship

I, Mohammed A. BEKHTI, declare that this thesis titled, "Traversability Cost Prediction of Outdoor Terrains for Mobile Robot Using Image Features" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Success is not final, failure is not fatal: It is the courage to continue that counts."

Winston Churchill

SHIZUOKA UNIVERSITY

Abstract

Graduate School of Science and Technology
Department of Information Science and Technology

Doctor of Philosophy

Traversability Cost Prediction of Outdoor Terrains for Mobile Robot Using Image Features

by Mohammed A. BEKHTI

Needless to say that terrain traversability is the most critical challenge mobile robots navigating in outdoor environments face due to the wide variety of terrain configurations. The ability of autonomous systems to accomplish missions in unknown remote territories depends highly on the embedded algorithm to build a reliable representations of the load bearing surface. Supervised visual appearance based methods are the most sought methods due to the abundance of attributes that can be extracted to describe a terrain nature, where in most cases, traversability is estimated for a particular terrain class through data labeling. Traversability is addressed from the vantage point of terrain nature with the assumption all regions ahead of the mobile robot within a certain class share the same properties. Some methods add an additional functionality to detect major terrain artifacts *obstacles* and rules them out of from potential passable regions. However, we argue that it is important to consider not only large-scale uneven artifacts, such as pile of stones, but also small-scale terrain objects that can cause major vibrations. By considering such small-scale unevenness, it becomes possible to reduce accumulated damage to the robot and instability of its cargo. This thesis leverages image properties to improve traversability prediction of non-uniform areas of a terrain a in self-supervised manner. Multiscale analysis is performed on terrain images to detect non-uniformity. Terrains are described by a pair of texture information and vibration data obtained using an inertial unit. These attributes are then used to predict vibration using image features by independent predictors both of which consist of Gaussian Process. Experiment results using real data obtained with our mobile robot in real world environment showed that our approach is reliable to detect non-uniform regions in a terrain to then apply the proper predictor to estimate vibrations. Vibration prediction witnessed a significant improvement compared with classic method of prediction with no non-uniformity detection, with only a an error of 0.23 and 0.26 for uniform and non-uniform terrain, respectively.

Acknowledgements

It is with a profound emotion that I would like to thank all those who, directly or indirectly, have contributed to the success of this work. My first thanks go to Professor Kobayashi, who directed and supervised my work. Thank you for your patience, precious advice, and availability. I hope I have been worthy of the trust you have placed in me and that this work ultimately lives up to your expectations. Anyway, I learned a lot from you, and I am very honored to have had you as a supervisor. Thank you for your encouragement and for being able to communicate to me your enthusiasm and passion for the field of robotics. I could not have imagined better guides to accompany my first steps in the world of scientific research.

My sincere thanks also go to Professor Miura, Professor Iwata and Pr. Inami who did me the honor of accepting to be part of the jury. I also would like to thank Professor Damon M. Chandler for his precious comments and advice.

I warmly thank all the members of Kobayashi Laboratory, whether permanent, PhD students or trainees, whom I have had a lot of fun with during the past years, more specifically Satoshi Ishikawa, Harada Kentaro, Takeshi Matsumoto, Sho Matsunaga, Minakawa Shotaro, Kazuki Matsumura. I would like also to thank my friends outside the laboratory, who have always been there for me at times difficult and who knew throughout these three years to give me the comfort of a family. I think more particularly of Karthikeyan Rajanand, Yash Sharma, Lioe De Xing, Sumeet Shrestha, Anchu Viswan, Prakash Natarajan, Anna Statsenko.

I cannot thank my family enough without whom nothing would have been possible. To Mom and Dad for allowing me to reach this level, for your presence, your support, your encouragement, your listening and your advice. I dedicate this work to you and hope that I have been able to make you proud.

To my brother, your presence by my side has always been essential and comforting. It's good to know that I can always count on you! To my in-laws, thank you for supporting, advising and encouraging me.

I would like to dedicate this work to my spouse. Despite the great distance that separated us for a very longtime, you always have been there for me. Thank you for having encouraged me and for having always believed in me, for having always inspired me and given the desire to surpass myself, for your comfort and your advice in times of doubt and difficult times and for having shared with me my joys and sorrows throughout this thesis. I hope I can be a source of courage and comfort for you as you have been for me.

Finally, I dedicate this work to my bundle of joy, my baby daughter. Knowing I was going to become a father gave me the courage and motivation to work harder to close this chapter and start a new one together. Thank you!

Contents

List of Figures

List of Tables

List of Abbreviations

NASA	National Aeronautics and Space Administration
MER	Mars Exploration Rover
M-TRAN	Modular Transformer
JPL	Jet Propulsion Laboratory
LiDAR	LIght Detection and Ranging
2D	2-dimensional
3D	3-dimensional
FPGA	Field-Programmable Gate Array
IMU	Inertial Measurement Unit
GPS	Global Positioning System
LAGR	Learning for Autonomous Ground Robots
DARPA	Defense Advanced Research Projects Agency
MIT	Massachusetts Institute of Technology
CalTech	California Institute of Technology
SVM	Support Vector Machines
MRF	Markov Random Field
DEM	Digital Elevation Model
CNN	Convolutional Neural Network
ANN	Artificial Neural Network
FFT	Fast Fourier Transform
PSD	Power Spectral Density
SIFT	Scale Invariant Feature Transforms
SURF	Speeded Up Robust Features
UAV	Unmanned Aerial Vehicle
GP	Gaussian Process
LBP	Local Binary Pattern
LTP	Local Ternary Pattern
CoT	Cost of Transport
TNUD	Terrain Non Uniformity Detection
ROI	Region of Interest
MSD	Main Subject Detection
$L^*a^*b^*$	CIELAB Color Space
RGB	Color Space with RGB primaries
BC	Box Counting
FD	Fractal Dimension

fBf	fractional Brownian motion
SFTA	Segmentation based Fractal Texture Analysis
TTBD	Two Threshold Binary Decomposition
RMSE	Root Mean Squared Error
SNR	Signal to Noise Ratio
PCA	Principal Component Analysis
CRF	Conditional Random Field

List of Symbols

C	Configuration Space	
C_{obst}	Obstacle Space	
C_{free}	Free Space	
q_I	Initial Configuration	
q_G	Final Configuration	
D*	Dynamic A*	
A*	Path Search Algorithm	
L	Wheel Base	mm
r	Wheel Radius	mm
v_l	Angular Velocity of Left Wheel	$\mathrm{rad\,s^{-1}}$
v_r	Angular Velocity of Right Wheel	$\mathrm{rad\,s^{-1}}$
$x(t)$	Position in X_w-axis	mm
$z(t)$	Position in Z_w-axis	mm
X_w-axis	x-axis of World Reference	
$\phi(t)$	Heading Angle	rad
t	Time	s
d	Danger Level	
P_i	Power Input	
v	Mobile's Robot Velocity	$\mathrm{m\,s^{-1}}$
CoT_{pow}	CoT based on Power Consumption	
P_{in}	Instantaneous Power Consumption	
m	Mass of the Mobile Robot	g
g	Gravity of Earth	$\mathrm{m\,s^{-2}}$
V	Voltage	V
I_{in}	Instantaneous Current	A
CoT_{mech}	CoT based on Mechanical Load	
E	Energy Expenditure	
L_s	Stride Length	
CoT_{heat}	CoT based on Heat Loss	
W	Mechanical Energy Consumption	
H	Heat Loss	
w	Width of a Block in the Input Image	
$\boldsymbol{B}$	Block in the Input Image	
$C(\cdot)$	Contrast of an Entity	
$STD(\cdot)$	Standard Deviation of an Entity	

$MEAN(\cdot)$	Mean of an Entity	
$CD(\cdot)$	Contrast Distance of an Entity	
CD_{map}	Refined Contrast Distance Map	
w_{map}	Width of Refined Contrast Distance Map	
h_{map}	Height of Refined Contrast Distance Map	
$isTerrainUniform$	TNUD Label	
CD_{gate}	Contrast Distance Threshold	
$\{O_c,x_c,y_c,z_c\}$	Camera Reference Frame	
P	Point in World Reference Frame	
p	Projection of P onto Camera Reference Frame	
$\boldsymbol{K}_c$	Camera Intrisinc Parameter Matrix	
$[\boldsymbol{R}\|t]$	Extrinsic Parameter Matrix	
W_r	Width of Mobile Robot	
$\{O_w,x_w,y_w,z_w\}$	World Reference Frame	
T	Set of Threshold	
n_t	Number of Threshold	
t_r	Threshold value for Image Binarization	
$\boldsymbol{I}^b_{x,y}$	Binary Value of Image Pixel (x,y)	
t_l	Lower Image Threshold	
t_u	Higher Image Threshold	
$\Delta(x,y)$	Border Image of the Binary Image $\boldsymbol{I}^b_{x,y}$	
$N_{x,y}$	8-Connexity of a Pixel (x,y)	
$N(\varepsilon)$	Number of Hyper Cubes	
ε	Length of Hyper Cube	
$\boldsymbol{x}$	SFTA Feature Vector	
a_k	Vertical Acceleration Signal	
a_{k_d}	Amplitude Distance Feature	
α	Camera Tilt Angle	$^\circ$
$\mathcal{D}$	Training Set	
y	Target Value	
$\boldsymbol{x}_*$	New Input Location	
y_*, f_*	New Target, Its Latent Function Value	
$\boldsymbol{w}$	Weights Vector	
$.^T$	Transpose of an Entity	
$\mathcal{N}$	Normal Distribution	
σ_n	Noise Level	
$\boldsymbol{K}$, $k(\boldsymbol{x}_i,\boldsymbol{x}_j)$	Covariance Matrix, Covariance Function	
$\boldsymbol{k}_*$	vector of covariates between a y_* and y	
$\boldsymbol{\theta}=(\sigma,\lambda,\sigma_n)$	Hyperparameters	
d_s	Distance Interval	mm
h	Camera Height	mm
F_s	Sampling Frequency	Hz

l	Length of ROI	mm
μ	Mean of a Signal	
σ	Variance of a Signal	

To my parents, my brother, my wife, and bibisha.

Chapter 1

Introduction

Autonomous navigation of a mobile robot in natural environments have seen a growing interest in recent years, especially under the stimulus of successful planetary exploration missions on Mars under the Mars Exploration Rover (MER) program initiated by the National Aeronautics and Space Administration (NASA). The demand for applications employing autonomous mobile robots in outdoor environments keep increasing. Besides planetary exploration, we can mention: surveying or intervention in environments hostile to humans (radioactive zone, polar) (Wagner et al., 2001), military operations such as mine clearance (Larionova, Marques, and De Almeida, 2006; Ishikawa, 2005), surveillance and protection (Shin et al., 2013). The primary advantage of performing these missions using autonomous mobile robots is to avoid human intervention in hazardous areas, or simply impossible to access. The autonomous operation of these platforms makes it possible to overcome the challenges that may occur in the case of remotely operated mobile robots, such as problems of connection degradation/loss or latency time. Among the various terrestrial locomotion systems available for autonomous terrain navigation, wheeled systems offer the best compromise between energy consumption, ability to transport payloads, easiness of control.

For these platforms to be able to navigate autonomously and safely in outdoor environments, various embedded processes need to be completed in real time. They must first have the ability to perceive the environment to obtain data necessary to compute their location, and to create a model the environment to assess terrain traversability. Based on the latter information, a mobile robot can then calculate the command to send to the motors to reach the final destination.

Autonomous mobile robots navigating in outdoor environments face two major problems. The first one being the lack of a priori information of the terrains ahead. Even when such information exist, it is likely to be imprecise due to the fact that outdoor environments are dynamic *i.e.* subject to changes due to weather conditions, etc. The second one raises from the wide variety of situations that the mobile robot may encounter: more or less rugged terrain, sand dunes, snowy or even icy, muddy or dry earth, slopes, etc. Sample terrains that can be found in outdoor environments are shown in Figure 1.1. Unlike structured outdoor environments where free space is considered traversable, rough outdoor environments it is simply

(a) (b)

Figure 1.1: Typical outdoor terrains. (a) SherpaTT robot traversing undulating terrain abundantly covered with rocks (Cordes, Kirchner, and Babu, 2018); (b) Nomad robot searching for meteorites on an icy terrain in Antarctica (Wagner et al., 2001).

wrong to make the same assumption. The goal for engineers is then to enable mobile robots with an intelligence capable of understanding diverse and priori unknown outdoor environments.

Autonomous mobile robot systems include several sub-modules. The most common components found in fully autonomous mobile robots are shown in Figure 1.2. Depending on the application, mobile robots may be significantly different in terms of internal structure. Even so, some fundamental processes such as terrain traversability analysis are common across all mobile robots. In what follows, we will introduce the different locomotion systems encountered for outdoor navigation, give an overview of sensors used to perceive the environment as well define terrain traversability. Localization and path planning will be introduce in Chapter 2.

1.1 Autonomous Robotic Navigation Systems

1.1.1 Locomotion Systems

One essential element to guarantee a good performance of the mobile robot used in rough outdoor environments is the locomotion structure. Terrestrial robotic locomotion systems can be one of three main categories: wheeled (Siagian, Chang, and Itti, 2013), legged (Wermelinger et al., 2016) (often inspired by humans or animals), and apodal (Guimarães et al., 2016) (bio-inspired modular robots

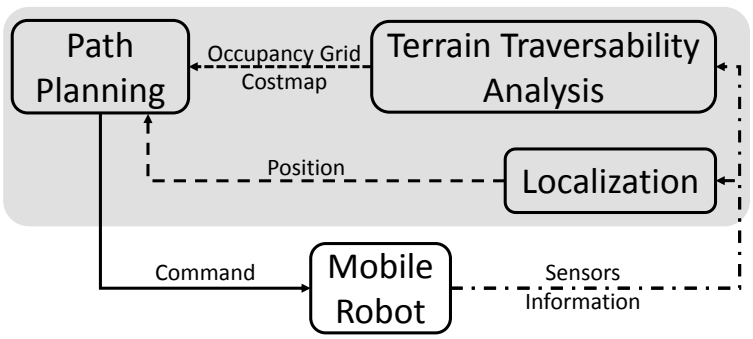


Figure 1.2: Typical components of an autonomous navigation system.

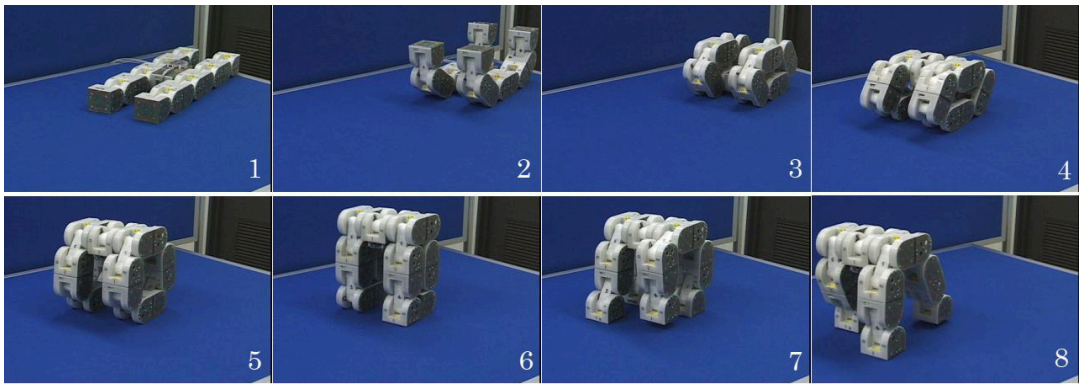


Figure 1.3: Locomotion transformations of M-TRAN (Kamimura et al., 2002)

that use elements of the structure to apply to the ground propulsion forces). Recently, engineers have a growing interest in *hybrid locomotion* as they can combine different types of locomotion structures. M-TRAN (Modular Transformer) presents complete reconfiguration capacities enabling it to traverse in very different modes of locomotion. It is made up of modules each consisting of two half-cylinders connected by a magnetic link. Consequently, the M-TRAN can pass from a form of quadruped robot (proceeding by walking) to that of a caterpillar (moving by crawling) (Kamimura et al., 2002). Transformation of M-Tran into different locomotion configurations is shown in Figure 1.3. In this study, we are interested in wheeled locomotion systems, closer to our applications. Regarding wheeled mobile robots, the variety in the type of locomotion can come from one of the following aspects:

- propulsion system,
- steering system,
- suspension system.

1.1.1.1 Propulsion

The most widespread propulsion for wheeled mobile robots is obtained by rotating the wheels. In case of the Lama robot, propulsion comes from the actuation of internal articulations of the chassis, when it progresses in peristaltic mode as shown in Figure 1.4 (Wettergreen, Thomas, and Bualat, 1997).

At Jet Propulsion Laboratory (JPL), Pirjanian *et al.* developed a unique locomotion systems around the Cliff-bot, where traction is possible by a secondary winch grapple system (Pirjanian et al., 2002). Cliff climbing by the Cliff-bot assisted by two anchor-bots is shown in Figure 1.5.

Figure 1.4: Lama robot on its experiment site (Lacroix et al., 2001).

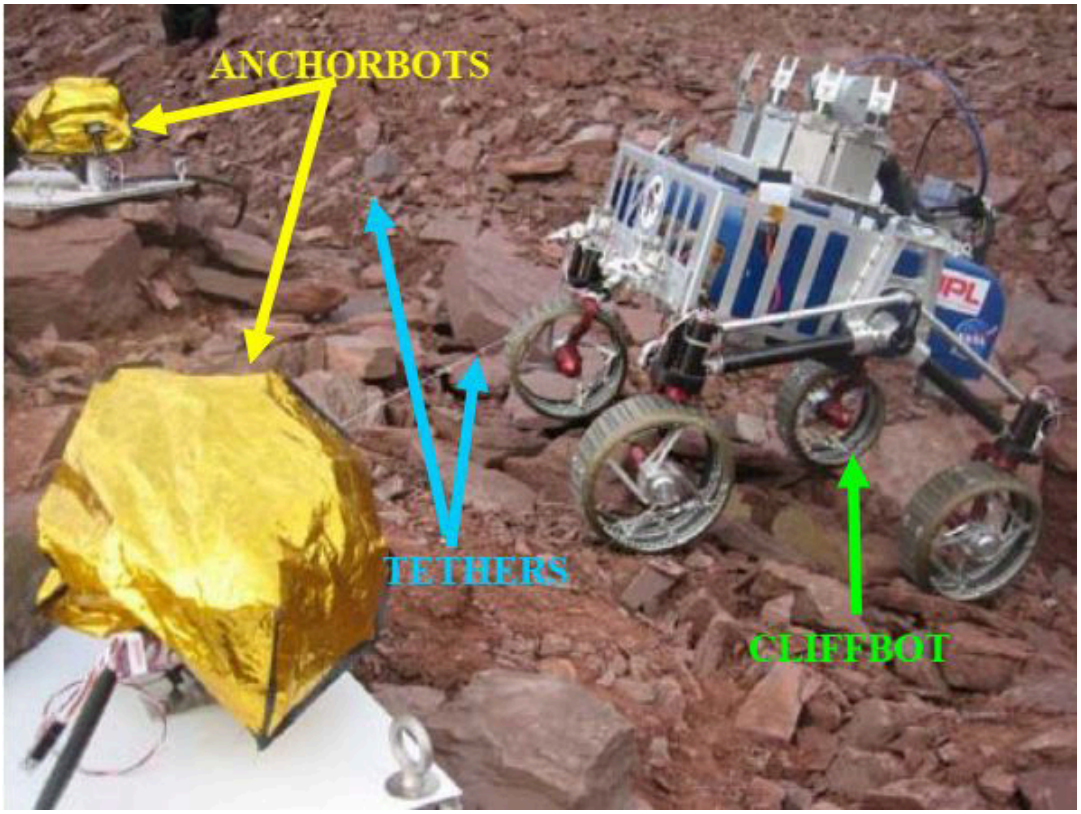


Figure 1.5: Cliffbot performing a cliff climbing aided by two Anchorbots (Huntsberger et al., 2007).

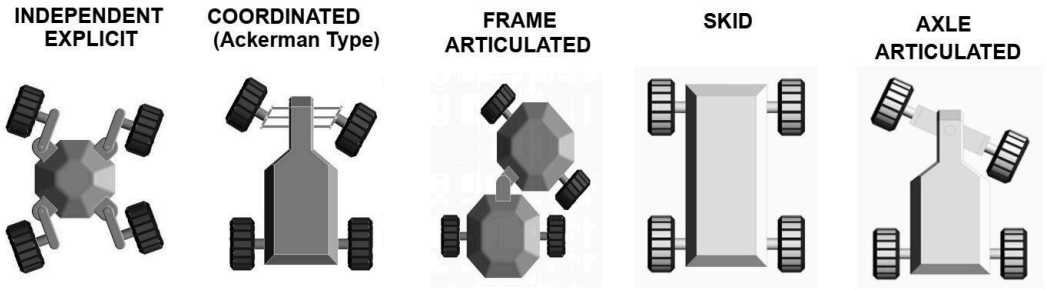


Figure 1.6: Typical steering schemes (Shamah et al., 2001).

1.1.1.2 Steering

Wheeled robotic systems operating in outdoor environments need the ability to change the direction of their motion *i.e.* their trajectory. Several kinematic architectures enabling this are shown in Figure 1.6.

- **Ackerman type**
  This kinematic is commonly used on commercial vehicles. In the event a mobile platform traversing a planar surface, on one of the axels, the coordinated rotation of the wheels is guaranteed by the mechanical coupling. When making a turn, The steering angle of the inner wheel is smaller than the outer one. The center of turning circle is obtained by the intersection of the lines perpendicular to the sagittal plane of the steerable and non-steerable wheels.

- **Articulated type**
  The chassis is made up of several parts that can be oriented relative to each other by connections. Railways vehicles are an example of this configuration with passive connections holding the wagons.

- **Independent explicit type**
  The wheels are independently oriented. This configuration is probably the best mobility option. It enables a mobile robot to execute very difficult maneuvers such as crab motion and rotation around the center of the platform. However, this architecture augments the mechanical complexity. It would requires the implementation of a control system to coordinate the actuators with high precision. Consequently, this system is less robust, heavy, and maintenance greedy.

- **Skid-steering type**
  In order to change trajectory, the idea is to impose different speeds on the wheels located on each side of the chassis. Lateral sliding is an important component needed for the rotation of the mobile robot. The instantaneous position of rotation center depends upon two parameters which are the slip and yaw speeds. Slip makes it difficult to use odometry measurements. Yet,

skid-steered mobile robots are particularly robust, light, and controlling them is relatively simple. For these reasons, we choose this configuration for our study.

1.1.1.3 Suspension

Suspension on an all-terrain locomotion systems increase crossing capabilities as well as stability as it maintains wheels/soil contact and traction. Suspension systems can be either passive or active. In the latter category, Jarrault *et al.* developed a method that take advantage of the kinematic reconfigurability of the mobile robot HyLoS 2 to enhance contact forces and tipover stability while traversing significant terrain blocks (Jarrault, Grand, and Bidaud, 2010). Although, the mobile robot used for experimentation in this thesis has no suspension system, we believe it is robot enough to traverse rough terrains within its physical constraints.

1.1.2 Terrain Perception and Traversability

Perception is a crucial requirement to build intelligent mobile robots capable of autonomously navigating unknown rough environments. The goal is to create a precise environment model to enable obstacle recognition, identifying passable from impassable terrains, and selecting the best route to get to the final destination without endangering its safety and stability. Throughout the years, engineers have explored a wide range of sensors to build terrain models and traversability assessment systems. In the following, we give a glimpse of the sensors used for terrain representation, as well as the mainstream terrain representation and traversability analysis methods.

1.1.2.1 Sensors

The most commonly used sensors for outdoor domains perception are:

- **LiDAR sensors**
  LiDAR sensor provide measurements of obstacles ahead of the mobile robot. However, 2D lasers destine to accomplish a plane scan, can only be used on a flat surface and provided that obstacles are visible at the height of the laser. hence, they are very efficient and widely used in indoor domains with obstacles such as walls, but less common in outdoor applications. 3D lasers provide a more complete representation even in unstructured outdoor environments. Nonetheless, They are still not popular due to the high cost and require a significant computational load due to the large amount of data to be processed.

- **Cameras**
  Cameras are certainly the sensors capable of providing the richest information on the environment, more than ever after engineers started exploiting several

images cues like color and texture. Moreover, mounting two cameras on stereo-vision bench would allow to estimate the depth of each pixel paired in image pairs and thus associate complete 3D coordinates with each point (Gallego et al., 2007). Images acquired through a stereo vision bench helps to compute a 3D point cloud pf the perceived environment, which can be used to build a numerical terrain model. However, to achieve this representation a severe computational cost is induced which often requires a slow progression of the robot. This has prompted engineers to find new ways to accelerate 3D map reconstruction such as associating stereo vision bench and FPGA (Devy et al., 2011).

- **Thermal sensors**
  Thermal sensors, Commonly known as *thermographic cameras*, are devices that use infrared radiations to create an image. While cameras provide significant visual information of the terrain surface, they fail to give an insight of the mechanical properties of layers beneath the surface. Cunningham *et al.* achieved soil strength assessment by exploiting thermal properties of granular materiel terrains (Cunningham et al., 2015). In other works, a thermal sensor was combined with a traditional camera to help extract obstacles (Matthies and Rankin, 2003).

1.1.2.2 Terrain Traversability

According to Papadakis (Papadakis, 2013), traversability is the ability of a ground mobile robot to traverse a terrain region. It is measured by taking into account the terrain model, the mobile robot model, and the kinematic limitations. In terrain traversability analysis, the goal is to generate a traversability map, often called traversability *costmap*, to measure the traversability load for a mobile robot to cross over a terrain area. In most cases, traversability is dependent of the locomotion system used, as different locomotion architectures differ remarkably in terms of size and capabilities. The core of traversability assessment is to use sensors on-board the mobile robot to deduce whether it is safe to traverse. This task would never cause any major issues in structured domains, but, analyzing terrain navigability in outdoor rough environments is very demanding as natural terrains vary on several aspects such as appearance/geometric properties, terramechanical properties, and also a priori unknown. hence building a system capable of building a reliable and faithful model of the load bearing surface is of high importance. To this end, engineers rushed to propose different solutions to tackle this problem.

In model-based frameworks, traversability analysis is performed using mechanical properties or simulation (*e.g.*, (Zhou et al., 2014; Mazhar et al., 2013)). However, the approach relies on a detailed description of terrain properties and is highly computational (Cunningham et al., 2017); therefore, it is not applicable

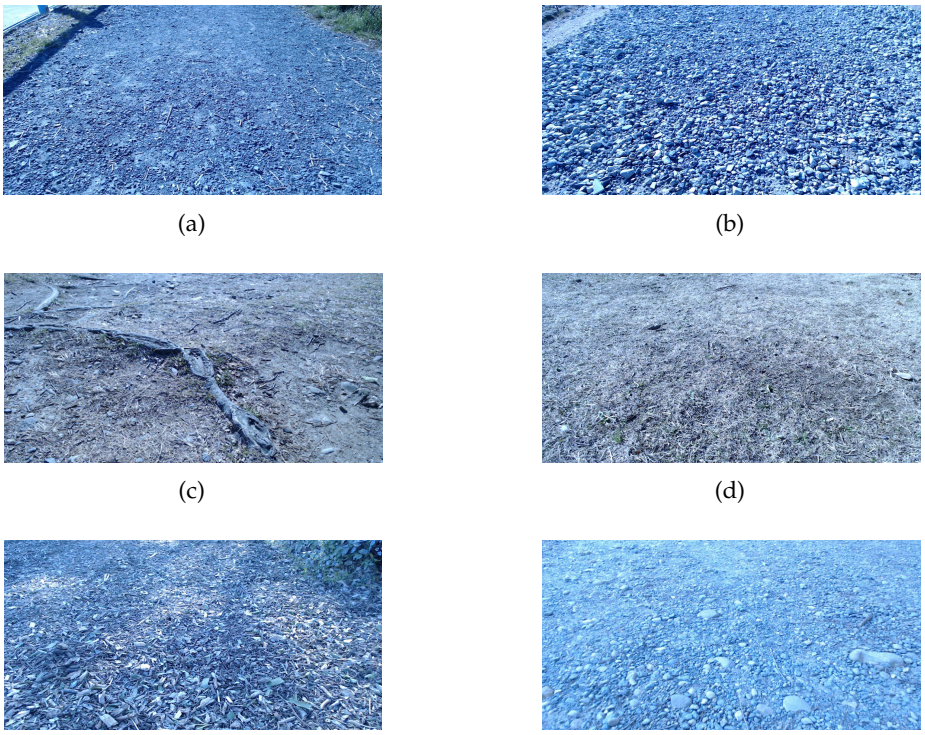
to certain environments. The most popular approach in recent years is to rely on machine learning techniques and terrain data to infer terrain traversability.

While supervised learning methods proved to be effective (*e.g.*, (Hong et al., 2002; Manduchi et al., 2005; Suger, Steder, and Burgard, 2015)), there are several drawbacks to such methods. Manually labeling a large amount of data to train the classifier is tedious and subject to inconsistency. As a result, this approach is not flexible as it is unable to adapt to dynamic environments without more data annotation and training. An alternative to the latter approach, self-supervised learning, based on proprioceptive recordings, such as inertial data, as label information to determine whether terrains is traversable, has been intensively researched (Stavens and Thrun, 2006). The self-supervised approach, also called the near-to-far strategy Ho, Peynot, and Sukkarieh, 2013a; Krebs, Pradalier, and Siegwart, 2010, is promising for making mobile robots adaptive to a variety of environments. A much detailed discussion about the mainstream methods for terrain traversability estimation is given in Chapter 2.

1.2 Problem, Objective and Approach

In this work, we propose a framework to enable mobile robots to autonomously learn terrain traversability using only their sensors in a self-supervised manner. Terrain traversability is predicted as a cost measured with an acceleration sensor mounted on the mobile robot. Focusing on the usage of image features for terrains with small-scale unevenness, a texture-based prediction of traversability was proposed in (Bekhti and Kobayashi, 2016). However, applications of image texture information has not been thoroughly investigated because, in natural environments, terrains are often are non-uniform, therein containing spiky regions such as relatively large stones and roots of trees. This non-uniformity causes traversability cost prediction to be challenging. Detection of such non-uniformity in terrains is proposed based on multiscale local image features. It is shown that we can improve the prediction performance of the texture-based approach. An advantage of the texture-based approach is that the sensor is affordable and can still detect motion features of the traversing robot without high-cost 3D sensing of the terrain geometry. To reduce the difficulty of motion feature prediction due to spiky objects in terrains, the classification of images into uniform/non-uniform is introduced. Predictors generated by a Gaussian process (Rasmussen and Williams, 2006) are independently applied to uniform/non-uniform terrains so that each predictor can be specialized to learn each type of terrain. The proposed framework for improving the prediction performance is evaluated through comparison with existing texture-based motion feature prediction.

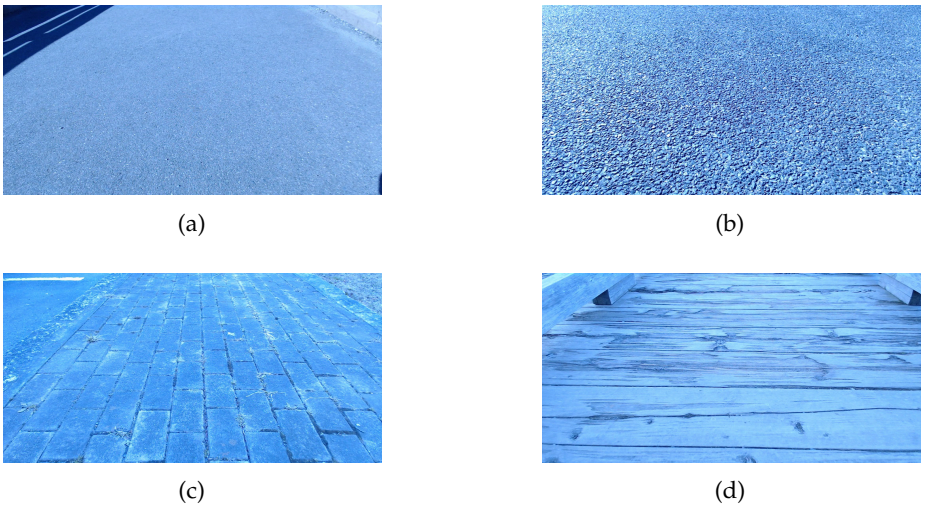
Terrains evaluated in this thesis are not always uniform due to irregular obstacles, such as large stones and roots of trees, and are assumed to be traversable even at a higher cost than homogeneous terrains. Typical obstacle size varies



(a) (b) (c) (d) (e) (f)

Figure 1.7: Natural terrains. (**a**) Gravel; (**b**) Stones; (**c**) Non-uniform; (**d**) Grass; (**e**) Woodchip; (**f**) Non-uniform.

between 40 to 80mm in height, which is a significant mobility challenge for the mobile robot as the wheels radius is 105mm. Terrains are assumed to be rigid. Deformable terrains (Ho, Peynot, and Sukkarieh, 2013a) that endanger the mobile robot's balance, or cause slippage/sinkage Howard et al., 2006, are not covered. In addition, untraversable terrains are also beyond the scope of this thesis. Samples of the terrains investigated in this paper are introduced in Figures 1.7 and 1.8.



(a) (b) (c) (d)

Figure 1.8: Artificial terrains. (a) Slick asphalt; (b) Granulated asphalt; (c) Tiles; (d) Wood.

Chapter 2

Literature Review

Traversability estimation for autonomous mobile robots navigating in rough environments focused mainly on visual- and geometry-based methods. Traversability analysis became the center of extensive research activities following the promotion of outdoor autonomous driving via the Learning for Autonomous Ground Robots (LAGR) program (Jackel et al., 2006) by the Defense Advanced Research Projects Agency (DARPA). Several researchers dealt with *near-to-far* frameworks by training long-range classifiers employing appearance/geometric cues and ground truth traversability quantification either from proprioception or data labeling. Others introduced metric evaluations to traversability. In what follows, we will introduce the two mainstream approaches to traversability: classification and prediction.

2.1 Classification Approach to Traversability Estimation

Proprioceptive traversability measurements are quantification of behaviors that translate from a direct interaction of the mobile robot and the terrain *under-foot*. Such measurements involve wheel torque, wheel sinkage, wheel slip, and terrain roughness. *Near-to-far* traversability prediction framework aiming at training a long-range classifier using image and/or geometry data to estimate these proprioceptive properties is of growing interest. Accordingly, it requires collecting a sustainable amount of ground truth proprioceptive measurements and corresponding exteroceptive cues.

Engineers at Massachusetts Institute of Technology (MIT) and California Institute of Technology (CalTech) developed frameworks to measure traversability in granular environments. At MIT, Halatci *et al.* classified terrains into three dissimilar traversability groups: sandy, rocky, and mixed (Halatci, Brooks, and Iagnemma, 2007). This was achieved by a naive Bayes to merge individual Support Vector Machines (SVMs) employing texture, color, as well as depth from stereo imagery. Brooks and Iagnemma broaden the framework by training on proprioceptive properties to enable self-supervised learning of traversability prediction. Traction coefficient was evaluated by discretizing the range of values in combination with a multi-class classifier. In a beach type terrain conditions,

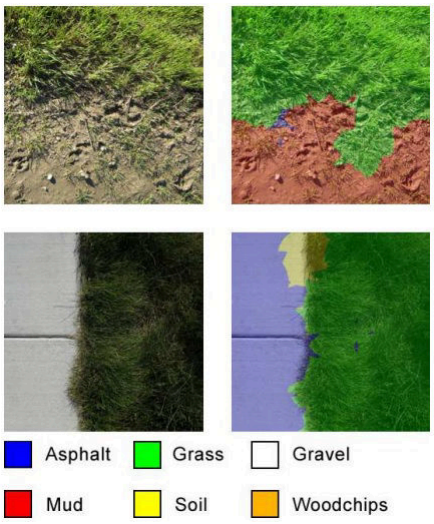


Figure 2.1: Heterogeneous Classification Results(Filitchkin and Byl, 2012)

this work accomplished reasonable results but lacked ground truth necessary for quantitative results (Brooks and Iagnemma, 2012).

At CalTech, Angelova *et al.* defined traversability as a slip prediction problem. Terrain nature is estimated from image texture using the nearest neighbor majority vote of in a dictionary of textons. Given a terrain class, a slip predictor based on receptive field regression algorithm (Schaal and Atkeson, 1998; Schultz et al., 2010) is built relying on pitch and roll of the mobile robot on the terrain. A 20% RMS error performance was achieved through experimentation in sand, gravel, soil, wood chip, and asphalt (Angelova et al., 2007). Above referred method was integrated with a path planning and the resulting architecture was tested at JPL's Mars Yard. Results of slip precision were not reported (Helmick, Angelova, and Matthies, 2009). Both research groups addressed slip-prediction in granular materials which are benign sand environment. However, rough outdoor environments are much more complex and can embedded loose materials and varying surface configurations that can endanger the mobile robot.

Traversability analysis is highly correlated with the nature of terrain the mobile robot is engaging with. Angelova *et al.* (Angelova et al., 2007) train individual traversability estimators for each terrain nature. Certain terrain classifiers perform at pixel level (Filitchkin and Byl, 2012; Manduchi et al., 2005), others on image segments (Angelova et al., 2007; Howard and Seraji, 2001; Seraji and Howard, 2002) and some on 3-D point segments (Munoz, Vandapel, and Hebert, 2009; Vandapel et al., 2004). All these techniques faced challenges to distinguish terrain class at boundaries, as illustrated by Figures 2.1. This emphasizes necessity to have reliable boundary segmentation to avoid having outliers. Smoothness restrictions specified in a Conditional Random Field (CRF) showed an improvement of inference quality (Munoz, Vandapel, and Hebert, 2009).

Terrain classification based on proprioceptive signals aims at enhancing precision and performance by choosing the best composition of sensors, descriptors

and classifiers. Among proprioceptive signals, some engineers rely on vibration while others focus on slippage and stickiness of the terrain. The goal of this work is predicting vibration of far range terrains. To fully comprehend motivation behind the usage of an IMU to measure vibration signals of terrain, a brief review on proprioceptive classification is discussed.

Terrain classification based on vibration was proposed by Iagnemma and Dubowsky using accelerometers as the proprioceptive sensor (Iagnemma and Dubowsky, 2002). Later on, authors used vibrations prompted on a wheel to approximate cohesion and friction angle in a testbed (Iagnemma et al., 2004). The very constrained test-bed involved a single rigid wheel assembled on a static vertical axis. Rotating the wheel at different rates, separate slip ratios are imposed. In practice, wheeled mobile robot have at least four wheels and loads forced upon each of them change according to the load-bearing surface nature. Sadhukhan and Moore used a neural network based classifier to detect terrain nature for a high-speed regular size vehicle (Sadhukhan and Moore, 2004). The size of mobile robot plays a important role as small changes in terrain geometry prompts different responses even at low velocity maneuvers.

In Ojeda *et al.* (Ojeda et al., 2006), comparison of terrain classification and characterization is explained. Pioneer 2-AT, a relatively small mobile robot was used jointly with an Artificial Neural Network (ANN) to classify terrain into distinct types such as dirt, gravel, pavement, sand and grass. Their classifier misclassified dirt with sand and grass due to similar vibration characteristics of those terrain types.

Weiss *et al.* built a terrain classifier based on a Support Vector Machine (SVM) on frequency representations (namely Fast Fourier Transform (FFT) and Power Spectral Density (PSD)) of vibration readings acquired from a cart pulled manually on seven distinct terrain types In (Weiss, Frohlich, and Zell, 2006). To this end, an analysis on a standard data set was necessary to compare the results of the different approaches. A comparison of prior classification works based on frequency description of readings gathered from an ATRV-Jr outdoor robot was presented (Weiss et al., 2007). SVM offers better results than previous terrain classifiers. A common denominator for mentioned methods evaluation being carried out at a constant speed scenario. In contrast, velocity independent terrain identification was tackeled in (DuPont et al., 2008; Ward and Iagnemma, 2009).

It is important to note that proprioceptive terrain prediction and discriminating terrain types using accelerometers and gyroscopes is possible only if the load-bearing surface causes terramechanical behavior in the robot.

Unlike proprioceptive traversability, data annotation refers to feature labeling with ground truth traversability. This can be achieved either by expert designers or heuristically. In (Bajracharya et al., 2009a), large variations from the ground plane fit to a terrain region based on stereo-measured morphology enable obstacle detection. Yet, since stereo-measurements do not offer precise geometric output at far range (*e.g.* for the LAGR robots, the system fails beyond a range of 10 m), a mono-vision

classifier can help match its close-range performance. Hadsell *et al.* predicted obstacles at up to 40 m range from mono-vision system. This was achieved using a deep convolutional neural network (CNN) and stereo training data. Ground truth obstacles were defined heuristically relying on difference from the ground plane fit to a terrain surface in stereo imagery (Hadsell et al., 2009; Hadsell et al., 2008). Vernaza *et al.* used a same near-range obstacle interpretation with sub-modular MRFs to classify at far range (Vernaza, Taskar, and Lee, 2008).

Many techniques acquired ground truth by means of human experts. Happold *et al.* represented the mobility cost of terrains to be the sum of geometric attributes involving height variation, terrain slope, and the ratio of cloud points above the ground plane. Ground truth traversability annotation was carried out by an expert engineer that assigned one of four classes (low, intermediate, high, or lethal) to a terrain cell. Then, the cost was estimated using trained a neural network with color in imagery using the geometric features (Happold, Ollis, and Johnson, 2006). In (Bajracharya et al., 2009b; Bajracharya et al., 2008; Howard et al., 2006), the authors trained a two-class SVM to predict traversability of far terrains in color information using training samples of a near-range classifier. At first, the near-range geometric classifier was trained manually by a designer. Authors also showed the ability to train the near-range classifier through self-supervised proprioceptive sensory information. Berczi *et al.* classified terrains ahead into traversable or untraversable by means of Gaussian Process trained with roughness obtained from stereo imagery measurement. Ground truth was humanly labeled (Berczi, Posner, and Barfoot, 2015).

Both heuristically and humanly defined ground truth are limited to predict traversability in outdoor environments. The issue raises from the inability of humans to adequately and intuitively define what is and what is not traversable by a mobile robot in an unknown domain as proved by the problems faced by the Mars rover.

Exteroceptive terrain classification are either geometry- or appearance-based methods. Single image frames hold color and intensity at each pixel but lack 3-D geometry of the terrain surface. Appearance-based frameworks deal with the problem from an image processing and classification standpoints that output a discrete set of terrain classes instead of a traversability measure. Terrain classification by cameras on-board ground robots belongs to a different category compared to readings acquired from aerial vehicles. However, the underlying practices in on-board and aerial camera image classification share several similarities between one another, literature covering both is introduced below.

Angelova *et al.* started classification with a fast and simple architecture and later on move to finer and more advanced classification algorithms based on color and texture cues (Ward and Iagnemma, 2009). The on-board camera discriminates load-bearing surface type into distinctive classes labeled by the author like sand, soil, grass, woodchips, gravel, asphalt. This method failed to recognize that within a

class of terrain such as grass, very different responses can be induced on the mobile robot. Flat grass or bumpy grass can imply distinct traversability metrics.

Terrains with significant surface granularity may cause images to captured at different angles. which may induce different rotation and scale for images obtained for the same terrain. This raises the importance of the visual descriptor to be invariant to rotation and scale. Fortunately, recent visual descriptors such as Scale Invariant Feature Transforms (SIFT) (Lowe, 2004) and Speeded Up Robust Features (SURF) (Bay, Tuytelaars, and Gool, 2006) enable extraction of robust. Filitchkin and Byl (Filitchkin and Byl, 2012) take full advantage of these features to discriminate a predefined set of natural terrain classes and afterwards select appropriate gait for a quadruped robot.

Aerial classification of outdoor domains came to life with the DARPA PerceptOR program. Aerial LiDAR data enabled prediction of ground mobile robot orientation, additionally detecting vegetation. It is regarded as both geometric- and appearance-based. Sofman *et al.* combined camera imagery cues and range readings from a laser mounted on-board an Unmanned Aerieal Vehicle (UAV), to discriminate terrains into road, grass, tree (Sofman et al., 2006b). This was later utilized to enable far range navigation for a mobile robot (Silver et al., 2006). The approach relies on the geometric and semantic interpretation of heterogeneous data to generate traversability cost chart while every class has a defined a priori traversal cost.

Hudjakov *et al.* classified terrains captured by static aerial imagery into road, house, grass or dirt groups using a convolutional neural network trained by a database of static images acquired from a UAV (Hudjakov and Tamre, 2009). Murphy *et al.* proposed a multi-class Gaussian Process (GP) classifier that compute likelihood of class membership of the image content rather than using labels In (Murphy and Newman, 2010). Texture cues of the terrain images play a major role in discriminating terrain natures. Several texture classification schemes in the literature benefit from sharp images including a homogeneous texture pattern acquired from a static camera angle. However, aerial moving vehicles capture images with blurred artifacts due to motion as well as the vibrations Some image texture descriptors are influenced by the work of Khan (Khan, Masselli, and Zell, 2012). Khan studied the performance of various image descriptors from images taken by a flying vehicle with different resolutions. Alongside two texture descriptors, Local Binary Pattern (LBP) and Local Ternary Pattern (LTP), SURF descriptor was also investigated.

2.2 Prediction Approach to Traversability Estimation

Stavens *et al.* developed a method to estimate terrain roughness that was employed during the DARPA Grand Challenge. Roughness was obtained from IMU feedback on the mobile robot. Local terrain elevation differences in the LIDAR point cloud was correlated to ground truth roughness to support long-range prediction (Stavens

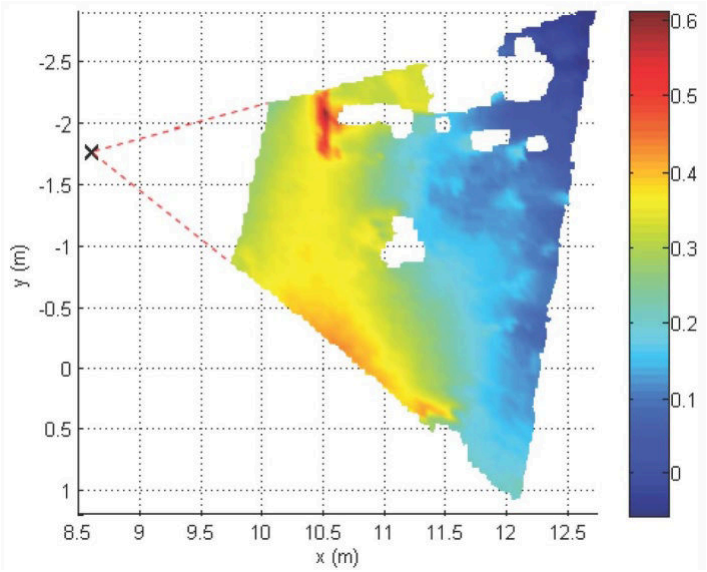


Figure 2.2: DEM of the terrain seen by the mobile robot represented by the cross on the left, colored by elevation. Occluded regions of the terrain appear in white (Ho, Peynot, and Sukkarieh, 2013b).

and Thrun, 2006). Attitude-based traversability prediction depends highly on precise and complete terrain morphology representation. However, most outdoor environments present irregularities of the ground height and existing objects can cause the occlusion of a significant region of the terrain ahead of the mobile robot, as illustrated by Figure 2.2. Ho *et al.* (Ho, Peynot, and Sukkarieh, 2013b; Ho, Peynot, and Sukkarieh, 2013a; Ho, Peynot, and Sukkarieh, 2016) provide a method to extend the elevation map into unknown regions of the terrain to enable worst case mobility to be estimated. The significant hypothesis is the rigidity of the ground. Ho *et al.* expanded the algorithm to predict the deformation of observed terrain based on past experience. The deformable terrains considered were unstable rock and sand. Other frameworks involved texture and color features from camera imagery to learn to predict using proprio-sensors such as suspension angles, bumpers, IMUs, and motor current (Dongshin Kim et al., 2006; Krebs, Pradalier, and Siegwart, 2010).

To estimate mobility of outdoor domains with dense vegetation, Wellington *et al.* pioneered development of algorithms to predict the ground height as well as obstacle detection in heavily vegetated agricultural domains. They used a voxel-column representation in a MRF framework to learn obstacle detection and traversability prediction despite partially the ground being partially occluded by vegetation. Ground truth ground height was measured by the robot driving over the terrain (Wellington, Courville, and Stentz, 2005; Wellington and Stentz, 2004; Wellington and Stentz, 2003).

Learning traversability models in this fashion involves an important amount of training samples that should represent all possible terrains a mobile robot may encounter. However, it is highly that very few samples are acquired in hazardous domains. This increases difficulty to construct a feature-rich terrain model based only on vehicle experience.

Methodologies aiming at predicting traversability from attitude consider the pose the mobile robot while traversing a specific terrain region. Such architectures involve accurate knowledge about the mobile robot as they are often parameterised with respect to certain system information. The estimated measure typically relies on properties of the load-bearing surface and the mobile robot used for the mission. It is assumed that terrain surface is clearly recognizable. In real world, this hypothesis is invalid. A reliable Digital Elevation Model (DEM) that describes elevation variations of the ground is essential for these methods.

Some engineers have used the DEM to compute a set of measurements through a fuzzy logic network, all contributing to the traversability of a terrain region:

- Terrain discontinuity (Goldberg, Maimone, and Matthies, 2002; Howard and Seraji, 2001; Seraji and Howard, 2002). This measure stops the mobile robot from "bottoming out". It is calculated from the maximum difference of adjacent height estimates, and is compared to the mobile robot minimum clearance.

- Terrain roughness (Goldberg, Maimone, and Matthies, 2002; Howard and Seraji, 2001; Seraji and Howard, 2002). This measure inhibits the mobile robot from traversing terrains with important surface irregularities, as it can be endangered with resultant vibrations. It is calculated from the residual error on a plane fit to the region, and is compared to a threshold based on the mobile robot minimum clearance.

- Terrain slope (Goldberg, Maimone, and Matthies, 2002; Howard and Seraji, 2001; Seraji and Howard, 2002). This measure prevents the mobile robot from sliding or tipping. It is calculated from the slope of the plane fit to a terrain grid, and weighted against the maximum tolerable pitch angle of the mobile robot.

- Terrain hardness (Howard and Seraji, 2001; Seraji and Howard, 2002). This metric helps the mobile robot to recognize hard terrains surfaces to limit likelihood of bogging. It is calculated by seeking the terrain type from a set including sand and dirt.

The combined traversability measure is expressed in terms of the mobile robot parameters; however the optimal adjustment of these parameters can be counter-intuitive as the optimal selection for parameters usually differs considerably from the obvious choices (Berczi, Posner, and Barfoot, 2015). Their method targets the detection of rocky regions. No learning was performed and in many cases the rocky regions can be camouflaged. Consequently, their method is not robust.

Other researchers employed a DEM to predict the attitude of the mobile robot while navigating regions (Ho, Peynot, and Sukkarieh, 2013b; Lacroix et al., 2002; Tarokh and McDermott, 2005). Given the mobile robot's attitude, traversability estimate is determined via a simple stability margin metric (Papadopoulos and Rey,

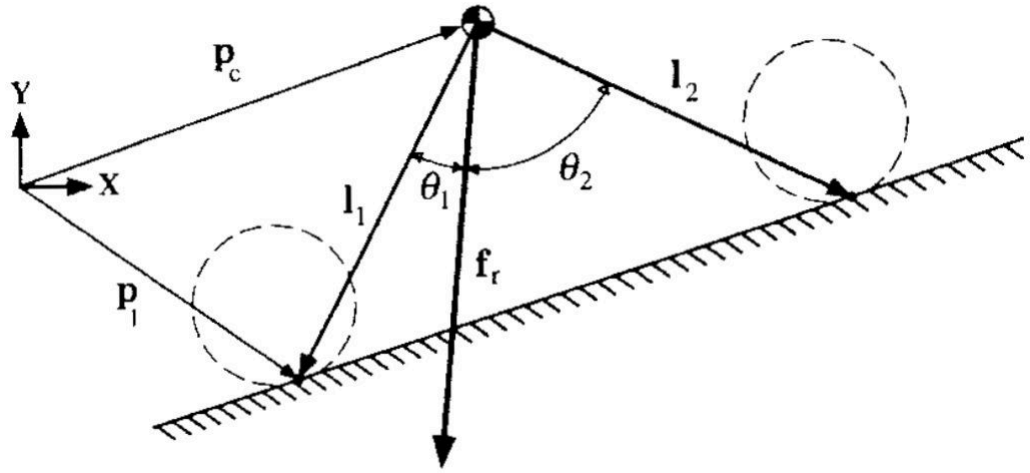


Figure 2.3: Angle metric (Papadopoulos and Rey, 1996).

1996), as shown in Figure 2.3. The stability margin is defined to be the minimum angle between the vectors to the edges of the support points of the robot and the force acting on the robot. As shown 2.3, this is given by θ_1. Critical tipover can be detected if θ_1 approaches the value zero.

When navigating in outdoor environments, mobile robots encounter various levels of traversability. On a flat smooth pavement, mobile robots face no major challenges compared with rough terrains where it is much harder to traverse. Mission success is highly correlated by the mobile robot's ability to efficiently navigate various a priori unknown environment. As this thesis is concerned with the regression of the continuous traversability cost, the following introduces some related works. Two groups of traversability assessment methods are discussed. In the first one, traversability is calculated from proprioceptive readings available to the robot like power consumption and velocity. For the second group, terrain traversability estimation can be constructed directly from captured exteroceptive data such as the ground slope or vegetation density observed in surroundings of the mobile robot. The benefit of the former method is that it is based on the mobile robot's experience traversing a specific terrain, i.e., it can consider changes in terrain traversability due to unfamiliar atypical terrain surfaces. An evenly distributed slope surface is fairly accessible however it can become much harder for a mobile robot to traverse if the load-bearing ground is wet. Exteroceptive methods are flexible to a wider portfolio of mobile robot platforms. On a given terrain, two mobile robots may record distinct proprioceptive readings, but the exteroceptive perception of the environment remains unchanged. Terrain classification into given discrete groups is the simplest strategy to the traversability analysis. While it does not clearly state that the individual classes have different mobility property, it is usually implied. However, a false estimation of the traversability cost may have irreversible outcomes to the mission.

A combination of individual classes and a continuous metric is introduced in (Stelzer, Hirschmüller, and Görner, 2012), where traversability is denoted as

danger level, which is seen as exteroceptive half-continuous-half-discrete measure representing the terrain difficulty. The danger level estimation is calculated based on slope, roughness, and step height of the terrain. Danger level values d vary between $[0, 1]$, where $d = 0$ denote passable terrains and $d = 1$ indicating barely traversable or unknown terrains. Non-traversable terrain morphology is represented by $d = \infty$. A continuous traversability parameter was computed by interpreting density, position, and distribution of ladar point cloud for sensed obstacles (Sofman et al., 2006a). Areas with dense vegetation translate with a high traversability cost value whereas smooth terrains such as roads are represented with a low traversability cost. The criteria is projected onto a logarithmic scale to better approximate prediction errors.

The Cost of Transport (CoT) was first introduced by (Tucker, 1975). It is a continuous proprioceptive measure expressed as

$$CoT_{original} = \frac{P_i}{v}, \tag{2.1}$$

where P_i is the power input and v is the mobile robot's velocity. The metric was borrowed by (Nishii, 2006; Kottege et al., 2015; Mulgaonkar et al., 2016; Xi, Yesilevskiy, and Remy, 2016; Alexander, 2005) in different forms to suit various mechanical architectures. The CoT measure is defined for passable terrains as nonnegative real number, i.e., $CoT \in [0, \infty[$. For nontraversable grounds, the corresponding CoT is either undefined or infinite. This method shares similarity with the danger level measure from (Stelzer, Hirschmüller, and Görner, 2012) as both assigns a continuous value to traversable terrains and infinity to impassable terrains.

Kottege *et al.* (Kottege et al., 2015), proposed a variant definition for CoT based on power consumption. the metric is computed as

$$CoT_{pow} = \frac{P_{in}}{mgv}, \tag{2.2}$$

where P_{in} denote the instantaneous power consumption, m is the weight of the robot, g is the gravitational force, and v is the velocity of the locomotive. CoT based on energy consumption is convenient for electrically driven mobile robots, as it involves instantaneous power consumption P_{in}, which is give by

$$P_{in} = VI_{in}, \tag{2.3}$$

where V is the voltage and I_{in} is the instantaneous current drawn by the motors. Energy consumption based measurement requires knowing power consumed by the mobile robot, and may not be helpful for application where employed locomotion is a biologically inspired platform to which measuring the mechanical load would be relevant.

Instead, Xi *et al.* expressed mechanical based CoT as

$$CoT_{mech} = \frac{E}{L_s}, \tag{2.4}$$

where E denote the energy expenditure and L_s denote stride length (Xi, Yesilevskiy, and Remy, 2016; Alexander, 2005). However, above described metric does not take into account additional energy loss via heat generation when the locomotion is driven.

Nishii (Nishii, 2006), redefined CoT to consider heat loss as

$$CoT_{heat} = \frac{\sum_{i=1}^{n} (W_i + H_i)}{mvt}, \tag{2.5}$$

where W_i denote the mechanical energy consumption, H_i denote heat energy lost, while i denote moving leg (out of n legs), m is the weight of the robot, v its velocity, and t its gait cycle duration.

2.3 Hybrid Approach to Traversability Estimation

Even in presence of ideal models of the terrain and robot, traversability prediction remains a difficult task due to the complexity of the robot-terrain interaction. Many terrains contain complex surface morphologies such as ripples and dunes that impact wheel-terrain interaction. Even on a relatively flat terrain surface, a mobile robot can still slip under the influence of local slopes generated by sinkage. An example of such scenarios is the difficulty for Curiosity to traverse complex polygonal dune configurations (Arvidson et al., 2017). Cunningham *et al.* take advantage of spatial correlations of slip measurements to overcome the inability to detect visual variations in sand terrains. After learning prediction models relying on slip data for visually classified terrain natures, the mobile robot adjusts itself to new terrains using proprioceptive information. On-line local adaptation improves prediction performance for far range terrains (Cunningham et al., 2017). Figure 2.4 illustrates results obtained using local adaptation.

2.4 Localization

Localization is one of the core functions of mobile robotics. In fact, it is important to have an estimate of the position of the mobile robot at all times so that the latter can carry out a trajectory without collision, duly follows the movement instructions received to reach the final destination. Given the significance of the problem in hand, several localization methods have been proposed. In what follows, we briefly introduce some of the localization methods used in outdoor environments.

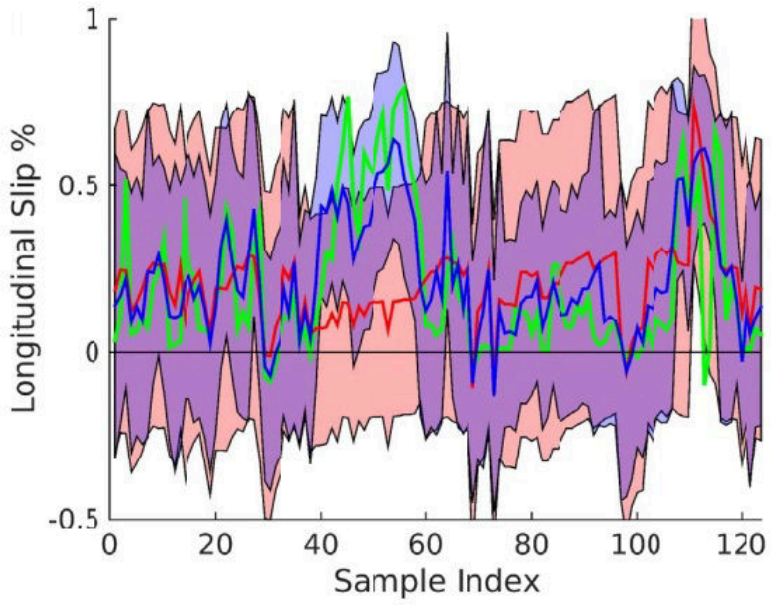


Figure 2.4: Comparison of GP and GP + local information on a subset of Sand samples as classified from visual. Green is ground truth, red is the GP, and blue includes local offset. Shaded areas represent 2.5 standard deviation bounds. Predictions were made within 1 m (Cunningham et al., 2017).

2.4.1 Odometry

Odometry relies on angular displacement measurements of the wheels provided by internal encoders to estimate the relative position of the mobile robot over time. Odometry is preferred among engineers for two reasons: 1. odometry is relatively simple to implement, 2. encoders are commonly found on ski-steering wheeled mobile robots. However, This method suffers from different errors that can be caused by variant factors. Systematic odometry errors, produced at every cycle, are mainly caused by biases in the parameter values of the mobile robot like wheels radius, encoder resolution, sampling frequency, etc (Bostani, Vakili, and Denidni, 2008). Other errors may come from the assumption of rolling without sliding on a flat surface, on which classic odometry formulations rely. In outdoor rough environments, many of the drifts happening induce important occasional errors that are impossible to quantify and correct in practice. Borenstein *et al.* introduced a an algorithm that combines odometry and inertial measurements to improve localization (Borenstein and Feng, 1996). The authors argue that irregularities in the terrain such as bumps impact the mobile robot for short periods of time where encoder and gyro readings varies significantly. Using this observation, the algorithm uses mostly odometry data, and employ gyro signals when encoder and inertial readings differ considerably.

2.4.2 Inertial Localization

This method takes advantage of readings provided by inertial measurement units (IMU), devices composed of 3 gyrometers each measuring the angular speed around

an axis and three accelerators each recording an acceleration along an axis. It is then possible to approximate of the position of the mobile robot boarding an IMU by single and double integrals of data from the gyrometer and accelerometer, respectively. However, this estimate is vulnerable to measurement noise and subject to an important drift if the information is not combined with other data. Most often, engineers implement fusion system where inertial unit data are merged with data from Global Positioning System (GPS) sensors (Sukkarieh, Nebot, and Durrant-Whyte, 1999).

2.4.3 Visual Odometry

Data used in this method are image pairs provided by a stereo vision bench. The process starts from correspondence operation of pixels in the two images then tracking these pixels in the next image frame. This leads to estimate parameters of the real displacement by computing the visual movement (Mallet, Lacroix, and Gallo, 2000). Efficiency of this method depends essentially on the selection of interest pixels used to approximate displacement changes. Lacroix *et al.* proposed an improved version of the classic method by reducing considerably the number of pixels relevant to the correlation operation, allowing a higher processing frequency and a faster estimate of the displacement (Lacroix and Jung, 2004).

2.5 Path Planning

Path planning is historically defined by the the piano mover problem, where the goal is to move the piano from one place to another in the middle of a room containing other objects (Schwartz and Sharir, 1983). Indeed, moving a piano within a room faces a problem: given the shape of the piano, is there a practicable path between the object existing in the room? To resolve this move problem, there must be at least one valid path that the mover must make between a start configuration and an arrival configuration. A movement is said to be valid if it is fully located in a space free of obstacles. Path planning therefore aims at finding a valid route for the moving entity. However, the problem will be very different depending on the object constraints. Path planning is very useful for mobile robotics as well as automotive applications (Diaz de Leon S. and Sossa A., 1998; Kim, Chung, and Park, 2010). It is also utilized to solve issues related to object manipulation (Saut et al., 2007), or molecular disassembly (Cortes, Jaillet, and Simeon, 2007).

2.5.1 Planning with Obstacles

To search a route around obstacles, a mobile robots requires a description of its surroundings. To solve the path planning problem independently of the type of the mobile robot, Lozano-Perez proposed the so called method *Configuration Space* C (Lozano-Perez, 1983). The problem is then defined as follows: give a space C

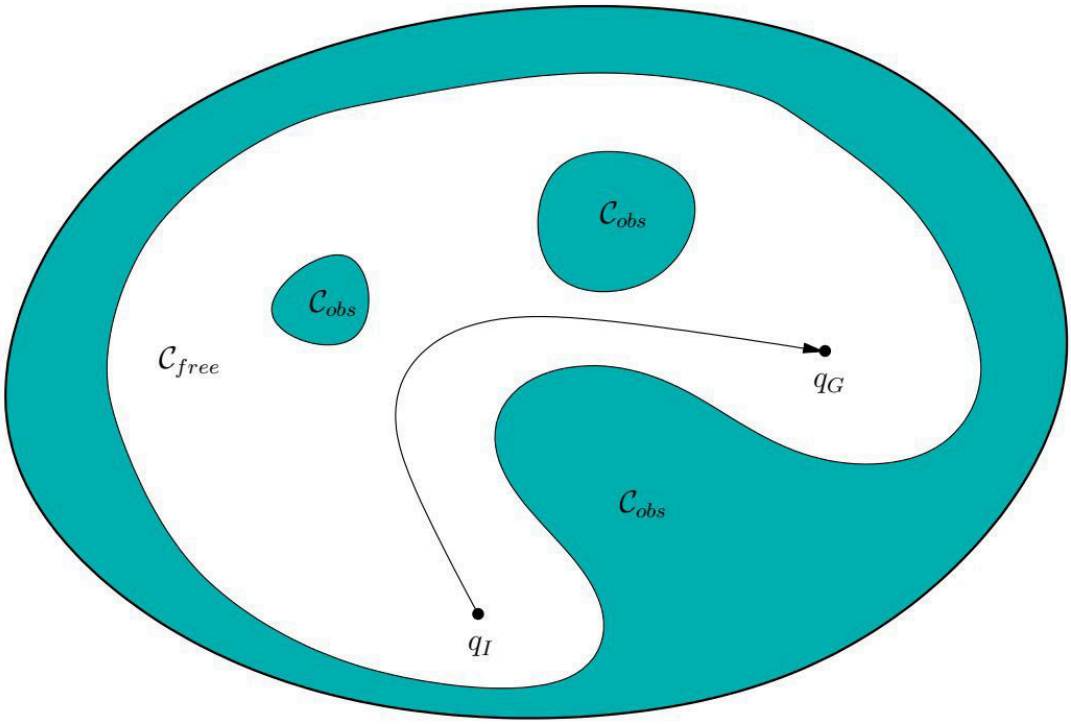


Figure 2.5: Configuration space idea: the goal is to search for a route from q_I to q_G in C_{free} (LaValle, 2006).

composed of obstacle regions C_{obst} and free spaces C_{free}, is there a valid a route between an initial configuration q_I and a final configuration q_G in C_{free}. The entire space is given by $C = C_{free} \cup C_{obst}$. The complexity of the problem raises exponentially with the dimension of C which makes it difficult to deal with. A graphical representation of this formulation is given by Figure 2.5.

In addition to the algebric description of configuration space, most often, occupancy grids are preferred. A two dimentional grid enables efficient storage of a wide area by dividing the space into 2D cells. Beside the widely preferred square cells, engineers introduced different regular polygons with interesting properties (Ryde and Brünig, 2009). Occupancy grid path planning is used very often in mobile robotics and is very successful in indoor domains. Marder-Eppstein *et al.* used an occupancy grid and the D* search algorithm to enable a mobile robot to navigate in an office (Marder-Eppstein et al., 2010; Stentz and Mellon, 1993). The authors called it the office marathon experiment. Occupancy grid generated from configuration space represents the surroundings of a mobile robot in binary form *i.e.* cells are either unoccupied or occupied. Some engineers argue that it is necessary to know the traversability cost of an unoccupied cell and extended the framework to become costmap. Martin *et al.* associated the energy usage commonly known as *Hotel Load* to each cell to in the costmap to describe the influence of the terrain on the mobile robot's motion (Martin and Corke, 2014). Reduction of the tour energy expanditure as well as number of excursions of each cell in the cost map is presented in Figure 2.6.

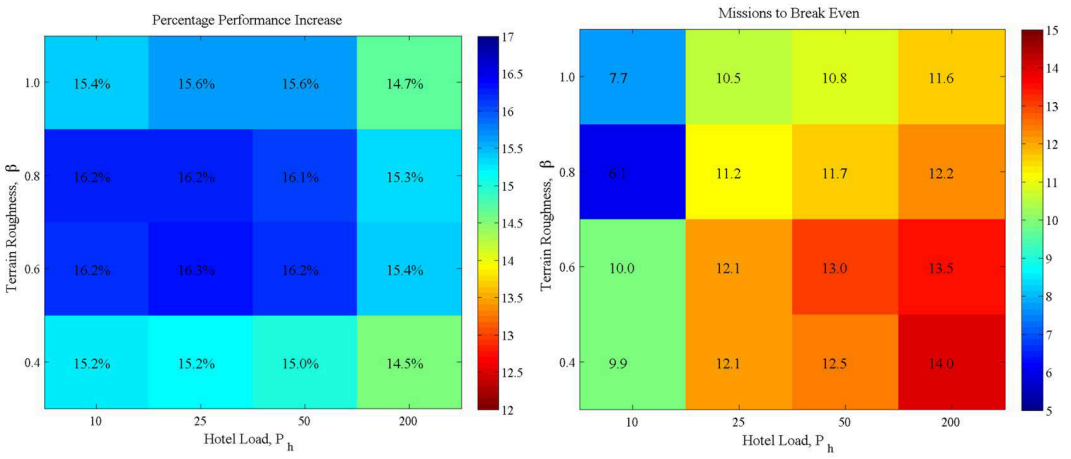


Figure 2.6: The average reduction in tour cost and number of tours until break with varying parameters. (Martin and Corke, 2014).

2.5.2 Graph Search

Many researchers perform path planning by means of graph search algorithms. Graph search aims to find minimal traversability cost from one node to another. Typically, each cell in an occupancy grid is considered as a node. The movement between cells is called an edge and translates into an invariant non-negative cost. In case of obstacles, moving to an occupied cell is materialized by an infinite edge cost.

One of the elementary graph search algorithms found in the literature is the Dijkstra algorithm (Dijkstra, 1959). It performs a uniform cost route search to achieve a complete and optimal path through a graph. Optimal paths search targets the lowest cost, while complete path search aims at finding a path given if it exists. Dijkstra algorithm is not constrained by any assumption on how to search for a route. It propagates in every direction, and if necessary in the opposite direction of the destination which is not always efficient.

To improve efficiency of path search of occupancy grid, the A* and D* graph methods apply heuristics. Hart *et al.* extended the Dijkstra's algorithm by introducing heuristics to govern path search (Hart, Nilsson, and Raphael, 1968; Hart, Nilsson, and Raphael, 1972). Their method predicts the best path through an additive cost function that operates in two stages; the path cost from the start configuration to the current node, and the heuristic estimate for the cost to reach the final configuration. A frequently used heuristic is the distance to the goal, which leads A* to begin the search in the direction of the goal. Several engineers raced to bring improvements to the classic A* algorithm (Korf, 2010), but the most known one is the D* algorithm (Stentz, 1993).

D* is a dynamic implementation of A* and is very efficient in calculating the optimal path. It allows to recalculate the route when new information is available such as obstacle positions. This method start searching in the opposite direction of the goal and propagates the cost to the surrounding nodes. In case of an obstacle, all concerned cells are marked and the algorithm resumes searching the path generated

previously.

2.6 Fractals for Texture Representation

Defining texture is by far a difficult task. This concept is correlated to homogeneity and characterization which ensures the discrimination of the different textures in an image. Indeed, in some cases we want to recognize the same texture in different scenes with, for example, variable lighting conditions. In monochromatic images, a zone with homogeneous levels of grey is unenlightening in terms of texture. This zone is characterized by the mean of grey levels, itself directly dependent on the lighting conditions. This is particularly the case with asphalt-like terrains where texture pattern are almost nonexistent, whereas grass grounds are highly textured. Another important factor involved in texture is the observation scale. A brick wall, for example, at a short observation range will prompt the viewer to consider the inside of the bricks as a particular texture and the separations between the bricks as another element of the scene. Conversely, a more global observation of the wall highlights the brick pattern redundancy to form the wall.

Approaches to texture characterization usually fall into the following categories:

- structural,
- statistical,
- model-based
- transform.

As an example of a structural method, Haralick defined texture by two primitives being the *microtexture* and spatial arrangements known as *macrotexture* (Haralick, 1979). Apart from the analysis aspect, rather adapted to macroscopic textures, this method can be efficient to synthesize a texture. Another advantage of this type of method is it enables a good symbolic representation of the studied texture.

In the statistical approach, the goal is then to represent texture through non-deterministic attributes. The most classical case, these attributes are measured at the pixel level more known as first order measurement, or a pair of pixels as in the co-occurrence matrix more known as measurement of the second order (Haralick, 1979). Second order measures have shown to be effective in the domain of human texture discrimination (Julesz, 1975). The multidimensional co-occurrence matrix was applied in biomedical imaging in (Lerski et al., 1993) and outperformed wavelet packets (Valkealahti and Oja, 1998) for classification of texture.

The model-based approach rely on stochastic models among which fractals. Parameters of the model are estimated and used for texture analysis. In practice, these methods are relatively computationally expensive. Fractal methods have proven to be suitable for the description of natural phenomena as described in (Chaudhuri and Sarkar, 1995; Kaplan and Kuo, 1995).

Transform approaches to texture characterization, texture is represented in another base than the spatial domain of the image. The goal is to find vectors of the base that are the most informative to recognize the texture. This category includes methods based on the Fourier transform, Gabor filters and wavelets (Rosenfeld, 1977; Bovik, Clark, and Geisler, 1990; Lu, Chung, and Chen, 1997).

In this thesis, texture features based on fractals are chosen as terrain descriptor, hence, in the following we give a brief overview of fractals and different methods for calculation. Mathematician Benoit Mandelbrot defined fractal to nominate objects with complex geometry that cannot be represented by a finite dimension. One of fundamental features of fractal objects lays in correlation of metric properties, such as length and area, and the measurement scale. A classical example of this would be the length of coastline (Mandelbrot, 1967). When measuring at a scale d, the total length of a given coastline $L(d)$ is approximated as a set of N segments of length d. hence, smaller details of the coastline that cold not be observed at weak spatial resolution become visible at higher resolutions. As a matter of fact, the measured coastline length $L(d)$ increase when the measurement scale d decreases.

Most often in image processing applications, fractal geometry is used through the fractal dimension. There exist several methods to calculate the fractal dimension, each having a different theoretical foundations. This diversity of calculation algorithms lead often to different dimension for a given object. This is due to the inability to compute the Hausdorff-Besicovitch dimension with the following formula:

$$D_h = \frac{\ln(N)}{\ln\left(\frac{1}{r}\right)}, \tag{2.6}$$

where N denote the number of internal homotheties of the object and r denote the reduction factor. To this end, several methods were proposed to estimate the parameter N. Calculation methods can be grouped into 3 categories: 1. Box-Counting methods, 2. Brownien movement based methods, and 3. area based methods. The first ever developed methods for fractal dimension calculation were based on the Box-Counting. This is certainly the reason why they are still being utilized nowadays. However, they do bare a major drawback. Signals are represented by a mesh of boxes of finite size and this influences the fractal dimension calculation. At this end, different algorithms were developed to eliminate such dependency.

2.6.1 Box-Counting algorithms

BC methods follow the following steps; representing the signal through a mesh of boxes, probability calculation, FD estimation via linear regression. The idea is rather simple and easy to develop, however, they do have certain negative points.

2.6.1.1 Classic Box-Counting algorithm

This measurement method was defined by Russel *et al.*, and is the most employed one (Russell, Hanson, and Ott, 1980). Its general principle is to cover a signal with boxes of size r. The FD is then given by:

$$FD = -\lim_{r \to 0} \frac{\log N(r)}{\log(r)}, \tag{2.7}$$

where $N(r)$ is the number of boxes necessary to completely cover the signal. This equation has few limitations. First, it requires a binary signal. Since most images are coded in grayscale, the binarization results in loss of information. Also, (Normant and Tricot, 1991) have shown that it is theoretically not well defined and that it is valid only for statistically self-similar signals. Moreover, repeating the calculation of the probability $N(r)$ with various sizes of r can precisely produce different values of $N(r)$. Therefore at every iteration the grid should be randomly repositioned, to be independent of the size r (Appleby, 1996). Pruess has also shown that the calculation of the FD can be sensitive to the size of the boxes (Pruess, 1995).

2.6.1.2 Differential Box-Counting algorithms

Chaudhuri and Sarkar proposed an adaptation of the BC method to overcome limitations discussed in section (Chaudhuri and Sarkar, 1995). This new definition is known as differential counting of boxes. Its main advantage is it can be directly applied to gray scale coded images. Thus the binarization step and its drawbacks are eliminated. The signal is partitioned into boxes of different sizes r and the probability $N(r)$ is calculated as being the difference between the maximum and minimum of the gray level in the box. This process is repeated for all boxes and the fractal dimension FD is calculated as in equation 2.7.

2.6.1.3 Extended Counting algorithms

The extended counting method was suggested as a substitute to the classic box counting. The main point is that the box counting is applied to several subsets of the fractal object and the maximum of the dimensions of these is retained as the fractal dimension of the signal (Sandau and Kurz, 1997). This method can therefore be compared to the box counting as the fractal dimension is calculated on binary signals. Although the latter is the most widely used, extended counting method an advantages. Indeed, it calculates a measure of complexity without the need to perform a linear regression. This measurement grows monotonically with complexity and is determined by the most complex region of the signal. It enables the measurement to be less sensitive to rotation or translation of the signal.

2.6.1.4 Fractional Brownian Motion methods

Fractal models based on fractional Brownian motion (fBm) are not stationary models and are often used to describe random processes. Pentland has shown that most of the fractals encountered in physical models are fractional Brownian functions (fBf) (Pentland, 1984). A fractional Brownian function is a general definition of a Brownian motion where the expected value of the difference of intensity of two points is null, but squared difference of intensity of these points is equal to the distance between the two points to the power $2H$ as given below:

$$\begin{cases} E[f(x) - f(x')] = 0 \\ E[f(x) - f(x')]^2 \approx \|x - x'\|^{2H} \end{cases} \tag{2.8}$$

where H denote the Hurst parameter associated with the fractional Brownian motion which indicates the roughness of the resultant motion, where a higher value leads to a smooth motion (Mandelbrot and Ness, 1968). The fractal dimension of an fBf oh dimension n is expressed by:

$$FD = n + 1 - H. \tag{2.9}$$

Since fBf are statistically affine, linear and scale transformations do not affect the fractal dimension measurement. Consequently, a fractal dimension based on fBf is invariant. To measure the performance of the estimator of H, it is necessary to study its bias and variance. An unbiased estimator with minimum variance is efficient if its variance is inferior to a variance from a different estimator. Assuming that the signal studied is an fBf, three algorithms are commonly used to estimate the fractal dimension. They are based on three concepts: the variance (Goodchild, 1980; Soille and Rivest, 1996), the power spectrum (Pentland, 1984) and the maximum likelihood (Dahlhaus, 1989).

2.6.2 Area methods

Area measurement rely on mathematical morphologies (triangles, erosion, dilation, ...) at different scales r and calculate the area $A(r)$ of a surface at a given scale. The fractal dimension is computed through a linear regression of the logarithm of $A(r)$ with respect to logarithm of r. Under this class, the most commonly used algorithms are isarithmic (Shelberg, Lam, and Moellering, 1983), and triangles (Clarke, 1986).

Chapter 3

Traversability Cost Prediction

3.1 Algorithm Architecture

Using data received from sensors, the mobile robot will build a model of the environment to predict terrain vibrations necessary to enable safe autonomous navigation. The overall system architecture is given in Figure 3.1. The architecture consists of the sensors available on the robot, which are a mono-vision system, an acceleration sensor, and odometers. The camera is responsible for gathering images of subsequent terrains. The inertial unit measures acceleration signals generated while traversing. The wheel odometer tracks the position of the mobile platform over time. The proposed method follows offline and online processes to achieve traversability cost prediction.

In the offline process, sensor information required for generating the traversability cost predictor is collected. Multiscale analysis measures a low-level image feature that is contrast distance feature to localize irregularities in terrain images. In terrain non-uniformity detection (TNUD), the resulting feature map from the multiscale analysis, is tested to determine whether the terrain nature is homogeneous or non-homogeneous. Region of Interest (ROI) localization extracts image regions traversed by the mobile robot based on its physical features and the camera model. The texture analysis calculates the texture features using the fractal dimension. Using the acceleration signal, the vibration analysis calculates the motion feature. Based on the output of TNUD, image and motion features are used to approximate either the function for uniform or non-uniform terrains.

In the online process, the same steps accomplished in the above mentioned offline process are executed to obtain the terrain image properties, that is, the terrain image properties, texture feature and terrain category. Based on the terrain category class, the vibrations of the subsequent terrain are inferred using texture information and the learned function.

3.2 Mobile Robot Description

This study was conducted with a widely used mobile platform (Pioneer 3AT) for navigation and traversability projects in outdoor environments (Chavez-Garcia et

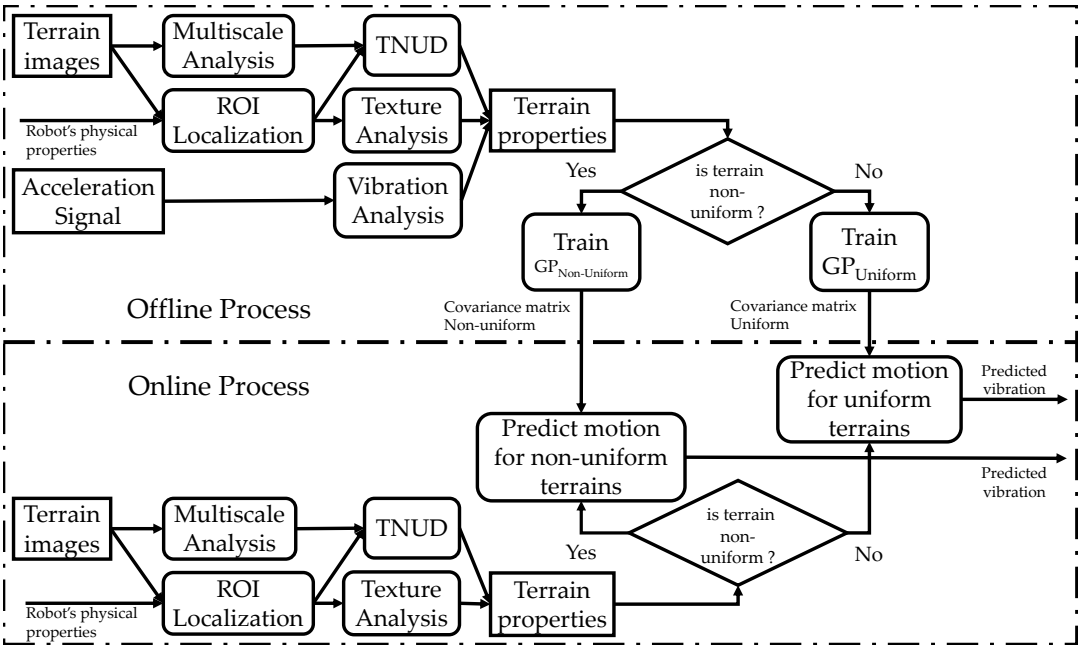


Figure 3.1: Overview of the proposed traversability cost prediction based on TNUD.

al., 2018; Metka, Franzius, and Bauer-Wersing, 2013; Ordonez and Collins, 2008; Collier and Ramirez-Serrano, 2009). The platform is built upon a skid-steering architecture and is capable of traversing bumps up to 100mm, gaps up to 150mm, and slopes up to 19°. The mobile robot carries a vision sensor and an acceleration sensor, as shown in Figure 3.2. The camera takes terrain images, and the IMU registers the acceleration signal generated during a sequence of a run. The goal is to predict traversability metric based on vibrations only from terrain images. The predicted traversability cost will be used to anticipate the motion nature of subsequent terrains and is expected to enable a safe run. The focus is not on which robot offers the best handling of terrain unevenness but on enabling any type of robot, regardless of its construction/configuration, to traverse non-uniform terrains.

In the bird's-eye view given by Figure 3.3, the gray circle represents the mobile robot and the black rectangles are the wheels. The wheel base and the wheels radius are denoted by L and r respectively. The system requires ω_l and ω_r as input, respectively, which are the angular velocities of the left and right wheels. The mobile

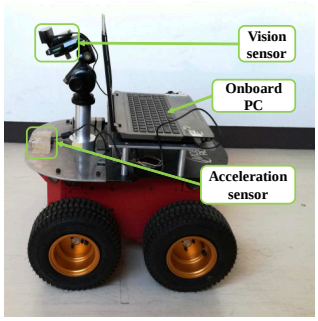


Figure 3.2: Pioneer 3AT mobile robot hardware setup used for experimentation.

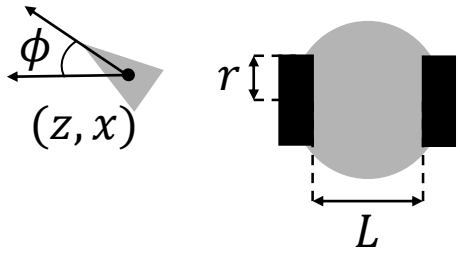


Figure 3.3: Differential drive kinematic model.

robot's kinematic model is described as follows:

$$\begin{cases} z(t) = \frac{r}{2}(\omega_{\mathrm{l}}(t) + \omega_{\mathrm{r}}(t))t\cos(\phi(t)), \\ x(t) = \frac{r}{2}(\omega_{\mathrm{l}}(t) + \omega_{\mathrm{r}}(t))t\sin(\phi(t)), \\ \phi(t) = \frac{r}{L}(\omega_{\mathrm{l}}(t) - \omega_{\mathrm{r}}(t)), \end{cases} \tag{3.1}$$

where x, z, ϕ, are the mobile robot's position and heading angle, respectively. The wheels angular velocities for both sides are similar; therefore, the robot will drive straight according to the world Z_w-axis. Moreover, the effects of steering and acceleration/deceleration are nonexistent. The mobile robot has no suspension system, that is, it is a rigid body. Hence, effects of interaction of the mobile platform on the right and left wheels will appear on the sensor values. Operating in outdoor environments comes with a very important lightning conditions challenge. To simplify the problem setting, experiments were conducted under fair light conditions with neither shadows nor strong backlight.

3.3 Image Feature Extraction

3.3.1 Terrain Non Uniformity Detection

Terrain non-uniformity detection refers to the task of identifying whether the subsequent terrain that lays ahead of the mobile robot is uniformly distributed terrain or contains significantly rugged regions *i.e.* localize which spatial areas in the terrain image correspond to the most visually relevant existing terrain artifact. TNUD is achieved through Main Subject Detection (MSD) which plays a major step in several image processing applications. In image compression, locating the scene defining object would enable the system to allocate more bits to the main subject and its surroundings (Wei, Sang, and Wang, 2009; Yu and Lisin, 2009). Other researchers also showed interest in MSD for image quality assessment (Osberger, Bergmann, and Maeder, 1998; Engelke, Nguyen, and Zepernick, 2008), as well as object recognition (Walther et al., 2002; Rutishauser et al., 2004). While humans are perfectly capable of identifying the main object of interest, this task proves to be difficult for automated systems. Methods for MSD can be categorized into two groups: 1. direct approach, and 2. visual attention prediction. In the latter one,

the objective is to find regions of interest in an image based on visual fixation. Researchers gathered both eye-movement data and scores of visual significance for a considerable volume of images. Experiments showed that a higher number of fixation for the main subject compared with other regions in an image (Wang, Chandler, and Callet, 2010). Instead of predicting visual fixations, some research works attempt to deal with MSD more directly via feature extraction (Ma and Zhang, 2003; Gopalakrishnan, Hu, and Rajan, 2009; Hu et al., 2004). We see TNUD as a direct approach to MSD and argue that an uneven morphological terrain artifact such as a bump, can be different from the background in terms of contrast. Therefore, to perform bumps identification, we compute the lightness contrast distance for each block in the input terrain image. Let $\boldsymbol{I} \in \mathbb{R}^{nRows \times nCols}$ denote the input terrain image where $nRows$ and $nCols$ are the number of rows and columns, respectively. $\boldsymbol{I}$ is divided into overlapping blocks of size $w \times w$ pixels. Bumps tend to have a higher contrast value than their surrounding regions. Therefore, we measure the distance contrast between the local and global contrast values of the blocks and input image $\boldsymbol{I}$, respectively. For this purpose, two additional scale images are produced by re-sizing the original image to half and a quarter by bicubic interpolation. Our image pyramid is a set of three images $\boldsymbol{I} \in \mathbb{R}^{nRows \times nCols}$, $\boldsymbol{I_1} \in \mathbb{R}^{\frac{nRows}{2} \times \frac{nCols}{2}}$, and $\boldsymbol{I_2} \in \mathbb{R}^{\frac{nRows}{4} \times \frac{nCols}{4}}$. As terrain artifacts can be of different sizes and resolutions, spatial operators may need to be scalable. One way of doing this would be to scale the input image instead. Therefore, we can produce an image pyramid using different scale factors to reduce the original image. This way, the operator maintains a fixed size. The smallest image in the pyramid will be analyzed at a larger scale due to the large relative size of the operator. Consequently, the features measured will be much larger and more global. On the other hand, the image at scale factor value of $\times 1$ is analyzed at a smaller scale which will enable to measure tiny features more locally.

Once image pyramid is formed, R, G, B values are converted to L^*, a^*, b^* measured in the Commission Internationale de l' Eclairage (CIE 1976) (L^*, a^*, b^*) color space (CIELAB) as discussed in (Vu and Chandler, 2011). Let R', G', B' denote the nonlinear RGB intensities of an image. L^*, a^*, b^* conversion is obtained by linearizing the R', G', B' to be comparable to the energy of the light. RGB color space is given by:

$$A = \begin{cases} A'/12.92, & A' \leq 0.04045, \\ \left[(A' + 0.055)/1.055\right]^{2.4}, & A' > 0.04045, \end{cases} \tag{3.2}$$

where $A = R, G, or B$. CIE XYZ color space is calculated using obtained R, G, B values as follows:

$$\begin{cases} X = 0.412453 \times R + 0.357580 \times G + 0.180423 \times B, \\ Y = 0.212671 \times R + 0.715160 \times G + 0.072169 \times B, \\ Z = 0.019334 \times R + 0.119193 \times G + 0.950227 \times B. \end{cases} \tag{3.3}$$

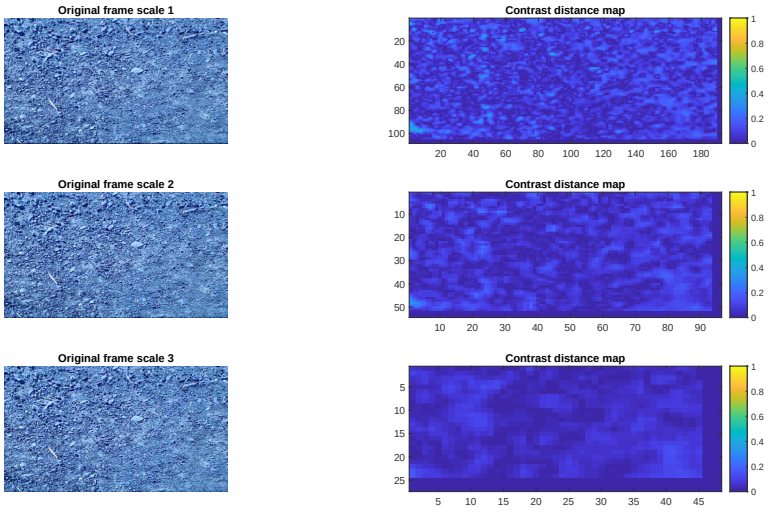


Figure 3.4: Contrast distance feature measurement at three different level for a uniform terrain.

Last, L^*, a^*, b^* values are computed by:

$$\begin{cases} L^* = 116 \times g(Y/Y_r) - 16, \\ a^* = 500 \times [g(X/X_r) - g(Y/Y_r], \\ b^* = 200 \times [g(Y/Y_r - g(Z/Z_r]. \end{cases} \tag{3.4}$$

where $X_r = 0.950456, Y_r = 1, X_r = 1.088754$ are the CIE XYZ tri-stimulus values of the D65 white point; the function g is expressed as follows:

$$g(t) = \begin{cases} t^{1/3}, & t > 0.008856, \\ 7.787 \times t + 16/116, & \text{otherwise.} \end{cases} \tag{3.5}$$

Using the lightness channel, the global lightness contrast of the image is measured by:

$$C(L^*) = STD(L^*)/\max(MEAN(L^*), 1), \tag{3.6}$$

where $STD(L^*)$ and $MEAN(L^*)$ denote the standard deviation and mean values of the lightness channel L^*, respectively. Let $\boldsymbol{B} \in \mathbb{R}^{w \times w}$ denote a block in the lightness channel of the input image $\boldsymbol{I}$, and let $C(\boldsymbol{B})$ denote the lightness contrast measured for the block $\boldsymbol{B}$. $C(\boldsymbol{B})$ is given as :

$$C(\boldsymbol{B}) = STD(\boldsymbol{B})/\max(MEAN(\boldsymbol{B}), 1), \tag{3.7}$$

where $STD(\boldsymbol{B})$ and $MEAN(\boldsymbol{B})$ denote the standard deviation and mean values of the block $\boldsymbol{B}$, respectively. The contrast distance is calculated using the results of Equations (3.6) and (3.7) as follows:

$$CD(\boldsymbol{B}) = \|C(L^*) - C(\boldsymbol{B})\|, \tag{3.8}$$

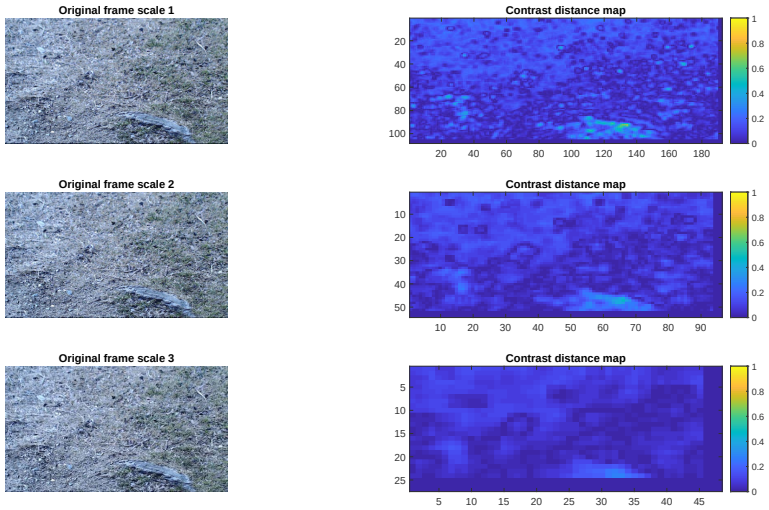


Figure 3.5: Contrast distance feature measurement at three different level for a non-uniform terrain.

where $CD(\boldsymbol{B})$ denotes the lightness contrast distance value of the block $\boldsymbol{B}$. Sample results are given in Figure 3.4 and 3.5.

To detect non-uniformity in terrain image, all scale maps are combined together to compute a refined contrast distance map using a maximum operator. As shown in Figure 3.6, non-uniform image regions tend to have a higher contrast distance value than the surrounding uniform areas.

The refined contrast distance feature map is then weighted against a threshold, and based on results of region of interest extraction only and only if the count of white pixels exceeds a defined gate, a terrain will be declared as non-uniform. Region of interest extraction will be explained later in this section. A summary of steps undertaken to label a terrain image is presented by Algorithm 1.

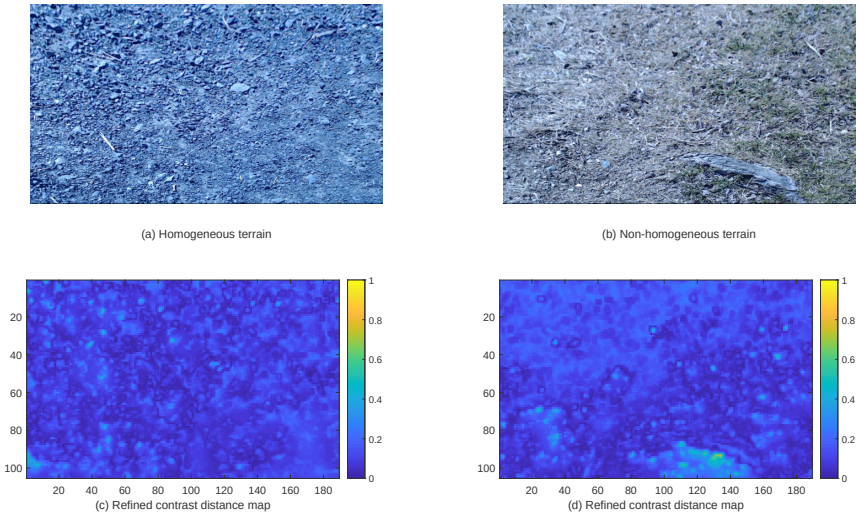


Figure 3.6: Refined contrast distance feature maps.

Algorithm 1: TNUD decision process.

input : Refined contrast distance map CD_{map} of size $w_{map} \times h_{map}$
output: TNUD label $isTerrainUniform$

```
numPixels = 0;
isTerrainUniform = TRUE;
for i = 1 to h_map do
    for j = 1 to w_map do
        if CD_map(i,j) > CD_gate and (CD_map(i,j) ∈ ROI_left or
          CD_map(i,j) ∈ ROI_right) then
            numPixels += 1;
            if numPixels > uniformTerrain_gate then
                isTerrainUniform = FALSE;
                i = h_map;
                j = w_map
            end
        end
    end
end
```

3.3.2 Region of Interest Extraction

From the terrain images, regions traversed by the mobile robot are considered for texture extraction. For this purpose, the camera model in (Hartley and Zisserman, 2003) is employed to define ROIs. As shown in Figure 3.7, the camera is defined by the camera frame $\{O_c, x_c, y_c, z_c\}$ where O_c is the origin. The following equation projects a point from the world reference frame given by $P \in \mathbb{R}^4$ to $p \in \mathbb{R}^3$ (both containing one as the last element) onto the image plane:

$$p = \boldsymbol{K}_c[\boldsymbol{R}|t]P, \tag{3.9}$$

where $[\boldsymbol{R}|t] \in \mathbb{R}^{4\times4}$ is the extrinsic parameter matrix, and $\boldsymbol{K}_c \in \mathbb{R}^{4\times3}$ is the camera parameter matrix obtained through a calibration process.

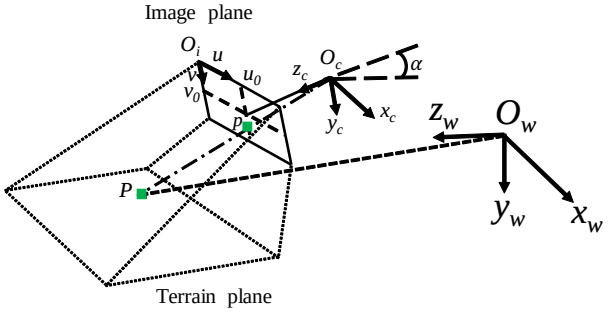


Figure 3.7: Camera model to project a 3D point onto the image plane.

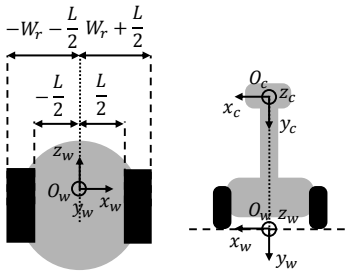


Figure 3.8: Physical properties of the mobile robot and layout representation of world and camera reference frames.

The above-mentioned model will be combined with the physical characteristics of the mobile robot, that is, the wheels base L and width W_r to determine the image regions traversed by the mobile robot. As shown in Figure 3.8, both camera and world reference frames are centered on the platform. The world reference frame is given by $\{O_w,x_w,y_w,z_w\}$ with O_w as the origin. The goal is to find points of inner and outer bounds of the terrain region crossed by the vehicle. All points belong to the same line; hence, only the depth components with respect to the w_z-axis will change. For both the left and right sides, the coordinates of these points are expressed as follows:

$$\begin{cases} P_{ro} = [\frac{L}{2} + W_r, \quad 0, \quad z_w] \\ P_{\text{ri}} = [\frac{L}{2}, \quad 0, \quad z_w] \\ P_{\text{lo}} = [-\frac{L}{2} - W_r, \quad 0, \quad z_w] \\ P_{\text{li}} = [-\frac{L}{2}, \quad 0, \quad z_w] \end{cases}, \tag{3.10}$$

where P_{ro} and P_{ri} denote the points of right outer and inner bounds, respectively. The results of this operation are shown in Figure 3.9. To increase the number of terrain samples, two terrain regions will be extracted from a single image. The warp perspective transformation is performed to remove the trapezoidal shape introduced when taking a picture.

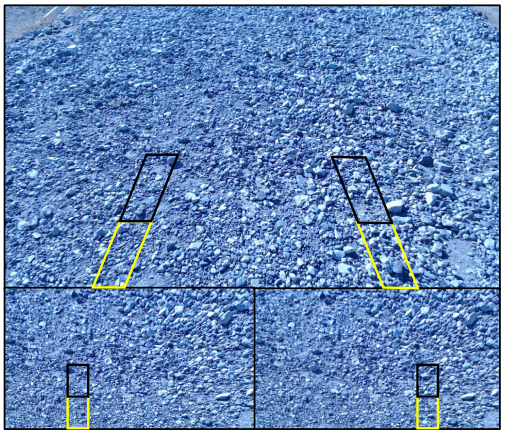


Figure 3.9: Identification of image regions of interest for texture feature extraction.

3.3.3 Texture Extraction

Segmentation-based fractal texture (SFTA) is retained for texture extraction (Costa, Humpire-Mamani, and Traina, 2012). The SFTA method operates in two steps. In the first step, the input grayscale image is split into binary images using the two-threshold binary decomposition (TTBD) (Costa, Humpire-Mamani, and Traina, 2012). In the second step, binary images are used to compute the fractal dimension from its region boundaries. TTBD returns a set of thresholds T calculated by the multilevel algorithm (Liao, Chen, and Chung, 2001). After obtaining the targeted threshold number n_t, pairs of contiguous thresholds $T \cup \{n_l\}$ with n_l being the highest value in the input gray scale image, in combination with pairs of thresholds $\{t_r, n_t\}$ with $t \in T$, are employed to obtain the binary images as follows

$$\boldsymbol{I}^b_{x,y} = \begin{cases} 1, & \text{if } t_l < I(x,y) \leq t_u \\ 0, & \text{otherwise} \end{cases}, \tag{3.11}$$

where $I^b_{x,y}$ denotes the binary value of image pixel (x,y), t_l and t_u denote the lower and the upper threshold values, respectively. Therefore, $2n_t$ binary images will be obtained. In this paper, the SFTA feature vector includes only the fractal dimension of the boundaries. Let $\Delta(x,y)$ denote the border image of the binary image $\boldsymbol{I}^b_{x,y}$, and which is obtained by the following equation:

$$\Delta(x,y) = \begin{cases} 1, & \text{if } \exists (x',y') \in N_{x,y} \\ & \quad s.t.\ \boldsymbol{I}^b_{x',y'} = 0 \wedge \boldsymbol{I}^b_{x,y} = 1 \\ 0, & \text{otherwise} \end{cases}, \tag{3.12}$$

where $N_{x,y}$ is the 8-connexity of a pixel (x,y). $\Delta(x,y)$ takes on a value of if the pixel at location (x,y) in the related binary image $\boldsymbol{I}^b_{x,y}$ has a value of one and has a minimum one neighbor pixel with a value of zero. The border image serves to compute the fractal dimension $D \in \mathbb{R}$ by the box counting method as follows

$$D_0(x,y) = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(\varepsilon^{-1})}, \tag{3.13}$$

where $N(\varepsilon)$ is the number of hyper-cubes of length ε that fill the object. The resulting SFTA feature vector is denoted by $\boldsymbol{x} \in \mathbb{R}^{2n_t}$.

3.4 Motion Feature Extraction

Let $a_k, k = 1, \cdots, K$ be the vertical acceleration signal generated from the wheel/soil interaction when traversing a short range distance, where k is the time step and K is the total time steps. The amplitude distance is used to describe the behavior of the

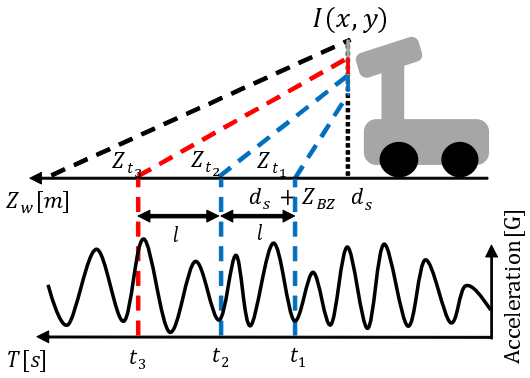


Figure 3.10: Data acquisition.

mobile platform, and is measure by:

$$a_{k_{\mathrm{d}}} = \max_{k=1,\cdots,K} a_k - \min_{k=1,\cdots,K} a_k. \tag{3.14}$$

According to (Bekhti, Kobayashi, and Matsumura, 2014), all motion features (amplitude distance, root mean square, kurtosis, skewness, and crest factor) have a similar level of correlation with the SFTA feature. The amplitude distance feature was chosen as it is easy to implement and to understand.

Two subsequent terrain segments serve for image feature extraction. Accordingly, the corresponding acceleration signal segment for motion feature calculation is paired to the image features by means of the coordinate transformation and odometry. As shown in Figure 3.10, at time $t = 0$, the robot is located at the origin of the world reference frame. The position used to take a new terrain image, denoted here by $Z_{I_i(x,y)}$, is expressed as

$$Z_{I_i(x,y)} = id_s, \quad i = 1, \cdots, N_{\text{image}}, \tag{3.15}$$

where d_s denotes the sampling distance, and N_{image} is the total number of terrain images acquired during a run. Due to the camera tilt angle α, terrains will be covered from a certain position, expressed as

$$Z_{I_i(x,y)} + Z_{\text{BZ}}, \quad i = 1, \cdots, N_{\text{image}}, \tag{3.16}$$

where Z_{BZ} is the blind spot. The motive behind focusing on a short distance range l covered by the images for texture feature extraction is that pixels further from the camera focus point are subject to more noise. Thus, visual information may fail to faithfully represent the environment. The acceleration signal sequence generated when traversing a distance l is used for motion feature extraction, and is limited according to the ROI as follows

$$A_i = [Z_{I_i(x,y)} + Z_{\text{BZ}}, Z_{I_i(x,y)} + Z_{\text{BZ}} + l], \tag{3.17}$$

where A_i denotes the ROI.

3.5 Traversability Cost Regression using Gaussian Process

In supervised regression, let $\mathcal{D} = \{(\boldsymbol{x}_j, y_j), j = 1, \cdots, n\}$ denote a training set with $\boldsymbol{x}_j \in \mathbb{R}^d$ and $y_j \in \mathbb{R}$, of n observed pairs $(\boldsymbol{x}_j, y_j)$. The $\boldsymbol{x}_j$ are indices to the set known as *inputs* and y_j are the output values, called *targets*. Given an input x_*, the goal is to predict the new target y_*. In practice, it is unlikely that $\mathcal{D}$ includes a complete range of $\boldsymbol{x}$. Moreover, an input $\boldsymbol{x}$ might have several y responses. This can cause an induction issue (Mitchell, 1997). That it, a set of assumptions describing relationship between observed and not yet encountered inputs. In machine learning, this is known as the *inductive bias*. The bias used for regression is based on similarity, *i.e.*, similar input x are highly to be correlated with the same output value y than different ones. The goal in regression analysis is to model the relationship between variables x and targets y as follows:

$$f(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{w}, \quad y = f(\boldsymbol{x}) + \varepsilon, \tag{3.18}$$

where f denote the function value, $\boldsymbol{w}$ denote the weights vector for the linear model, and ε denote the additive noise component. The noise terms are considered to obey to an independent identically distributed Gaussian distribution with zero mean and a variance σ_n^2 and is denoted by:

$$\varepsilon \backsim \mathcal{N}(0, \sigma_{\mathrm{n}}^2). \tag{3.19}$$

Definition: *A Gaussian process is a collection of random variables, any of which have a joint Gaussian distribution* (Rasmussen and Williams, 2006). More formally, the Gaussian process regression is a Bayesian algorithm presuming that *a priori* the function values respond as

$$p(\mathbf{f}|\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}), \quad \boldsymbol{\mu} \in \mathbb{R}^n, \boldsymbol{K} \in \mathbb{R}^{n \times n}, \tag{3.20}$$

where $\mathbf{f} = [f_1, f_2, \cdots, f_n]^\top$ contains the latent function values with $f_j = f(\boldsymbol{x}_j)$. For simplicity of notation, we presume $\boldsymbol{\mu} = 0$. The covariance matrix $\boldsymbol{K}$ is an appealing feature of the model, of which inputs are provided by the covariance function defined as $\boldsymbol{K}_{ij} = cov(f_i, f_j) = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. The covariance function highlights the statistical relationship between two points. In practice, the *squared exponential* covariance function is the most popular for similarity measurement. It is expressed as follows:

$$K_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma^2 \exp\left(-\frac{(\boldsymbol{x}_i - \boldsymbol{x}_j)^\top (\boldsymbol{x}_i - \boldsymbol{x}_j)}{2\lambda^2}\right). \tag{3.21}$$

Here σ^2 controls the variance, and λ is the isotropic length scale parameter describing the smoothness of a function. the squared exponential is *stationary isotropic*, that is, the function depends essentially on the relative distance $|x_i - x_j|$

which makes it invariant to translations in the input space. In the literature, σ^2, λ, and σ_{n}^2 are formally known as *hyperparameters* and are denote as $\boldsymbol{\theta} = [\sigma, \lambda, \sigma_n]^T$.

Because the Gaussian process is defined as a set of jointly Gaussian distributed random variables, predicting new target values $\boldsymbol{y}_*$ at for corresponding input locations $\boldsymbol{x}_*$ is equivalent to evaluating $p(\boldsymbol{y}_*|\boldsymbol{x}_*, \mathcal{D})$. we recall that $\mathcal{D}$ is the training set of input/output pairs. In what follows, $\boldsymbol{y}_*$ denote the vector containing l function values y_j^* to be predicted with $\boldsymbol{X}_*$ the matrix of corresponding input vectors $\boldsymbol{x}_j^*, j = 1, \cdots, n$. The joint prior distribution of the training outputs, $\boldsymbol{y}$, and the test outputs $\boldsymbol{y}_*$ according to the prior is

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{y}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \boldsymbol{K} & \boldsymbol{K}_*^T \\ \boldsymbol{K}_* & \boldsymbol{K}_{**} \end{bmatrix}\right), \tag{3.22}$$

where $\boldsymbol{K} \in \mathbb{R}^{n \times l}$, $\boldsymbol{K}_* \in \mathbb{R}^{n \times l}$, and $\boldsymbol{K}_{**} \in \mathbb{R}^{l \times l}$. The latter entities, in case of a single test index point $_*$ are computed as:

$$K = \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & k(\boldsymbol{x}_1, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_n) \\ k(\boldsymbol{x}_2, \boldsymbol{x}_1) & k(\boldsymbol{x}_2, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_2, \boldsymbol{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_n, \boldsymbol{x}_1) & k(\boldsymbol{x}_n, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_n, \boldsymbol{x}_n) \end{bmatrix} \tag{3.23}$$

$$K_* = \begin{bmatrix} k(\boldsymbol{x}_*, \boldsymbol{x}_1) & k(\boldsymbol{x}_*, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_*, \boldsymbol{x}_n) \end{bmatrix} \tag{3.24}$$

$$K_{**} = k(\boldsymbol{x}_*, \boldsymbol{x}_*). \tag{3.25}$$

The predictive distribution of the latent function for Gaussian Process Regression, $\boldsymbol{y}_*$, is given by $\boldsymbol{y}_* \sim \mathcal{N}(\boldsymbol{\mu}_*, \boldsymbol{\sigma}_*^2)$ where $\boldsymbol{\mu}_*$ and $\boldsymbol{\sigma}_*^2$ are given by the followings:

$$\boldsymbol{\mu}_* = \boldsymbol{K}_* \boldsymbol{K}^{-1} \boldsymbol{y}, \tag{3.26}$$

$$\boldsymbol{\sigma}_*^2 = \boldsymbol{K}_{**} - \boldsymbol{K}_* \boldsymbol{K}^{-1} \boldsymbol{K}_*^T. \tag{3.27}$$

Chapter 4

Experiment and Results

Developing a reliable terrain traversability analysis to enable mobile robot systems to navigate autonomously in unknown outdoor environments safely is a very challenging problem. The outline and development process of the proposed method for terrain traversability regression based on TNUD was fully detailed in Chapter 3, whereas this chapter introduces and discusses performance results of vibration prediction.

4.1 Experimental Settings

The experiment was conducted at Sanaru Lake in Hamamatsu, Japan, in a wide portfolio of terrains, where a full database was recorded, including terrain images and acceleration signals. Terrain images were taken at a distance interval $d_s = 500mm$ with a single camera placed at a height $h = 540mm$ and tilted with an angle $\alpha = 31^{\circ}$. The acceleration sensor is configured to register vertical acceleration at a sampling frequency $F_s = 100$Hz and up to $\pm 8g$. This results in an acceleration segment of 125 samples for a region of interest length $l = 250mm$. During data acquisition, the mobile robot was traversing at a constant velocity $v = 200mm/s$. The data set was later on divided randomly into training and test sets. The total number of samples is 2582, from which 90% of samples are allocated for training, and the remaining 10% of samples are used for testing. The refined maps from multiscale analysis were weighted against a lightness contrast distance threshold CD_{gate}. All experiments were performed on an Intel Core i7-3770 3.4 GHz processor with 8GB of RAM.

The goal of the experiment is to verify the ability of TNUD to detect non-uniformity in terrain images and predict traversability cost reliably. We validate the effectiveness of introducing TNUD to the traversability cost prediction by comparing it with the framework introduced in (Bekhti and Kobayashi, 2016), where the Gaussian process was directly applied to cost prediction without non-uniformity detection. For this purpose, we compute the root-mean-squared prediction error

(RMSE), which is defined by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(\boldsymbol{x}_i) - \mu_i)^2}, \tag{4.1}$$

where N is the number of test samples, μ_i denotes the predicted mean vibration of the input image texture $\boldsymbol{x}_i$ and $f(\boldsymbol{x}_i)$ is the corresponding ground truth vibration. But first, we will introduce results of intermediate processes leading to motion prediction.

4.2 Results and Discussion

4.2.1 ROI Extraction

ROI localization as described in section 3.3.1, is the process of recovering the mobile robot's wheels tracks boundaries onto the image plane using equations 3.9 and 3.10. To achieve this, the extrinsic parameter matrix $[\boldsymbol{R}|t]$ and the camera intrinsic parameter matrix $\boldsymbol{K}_c$ are needed. The camera intrinsic parameter matrix was obtained through a calibration process with planar calibration rig images and using the *Camera Calibration Toolbox for Matlab* (Bouguet, 2015). Their respective values are given below:

$$[\boldsymbol{R}|t] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8599 & 0.5105 & -0.5400 \\ 0 & -0.5105 & 0.8599 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.2}$$

$$\boldsymbol{K}_c = \begin{bmatrix} 1370.68 & 0 & 980.71 & 0 \\ 0 & 1369.90 & 539.52 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{4.3}$$

4.2.2 TNUD

Results of terrain nature discrimination using TNUD are given in Table 4.1 and Figure 4.1. Table 4.1 summarizes the number of samples that were used to train our method for uniform and non-uniform terrains, as well as the number of samples to test performance of our method to predict motion information for uniform and non-uniform terrains. As discussed in section 3.3.1, the refined contrast distance map is weighted against a threshold to decide the nature of the terrain ahead of the mobile robot. Consequently, the value of the threshold is very critical to achieve optimal terrain discrimination. We study also the influence of contrast distance threshold on TNUD by varying its values from 0.3 to 0.5 with an interval equal to 0.05.

Table 4.1: Influence of different contrast distance threshold values on number of samples used for training and testing processes.

Training	Contrast Distance Threshold	0.35	0.4	0.45	0.5
	Uniform Terrain Samples	1276	1592	1820	1956
	Non-uniform Terrain Samples	1048	732	504	368
Prediction	Uniform Terrain Samples	154	178	214	220
	Non-uniform Terrain Samples	104	80	44	38

As shown in Figure 4.1, when the contrast distance threshold value decreases, we see an increase in number of terrain samples classified as non-uniform and a decline in number of terrains samples judged to be uniform. An example showing influence of various contrast distance thresholds on TNUD is shown in Figure 4.2. Although the motion feature for terrains in Figure 4.2(a) and Figure 4.2(b) are and , respectively, they are classified as non-uniform for contrast distance threshold values 0.45 and 0.35. Some uniform terrains may have contrast distance map with values higher than the contrast distance threshold, therefore, TNUD discriminates these terrain samples as non-uniform terrains. Experimentation with different contrast distance threshold proved that a threshold value of 0.5 leads to an optimal TNUD discrimination *i.e.* significantly better traversability prediction error. We will show later in this section 4.2.4 how the contrast distance threshold influences also terrain traversability estimation.

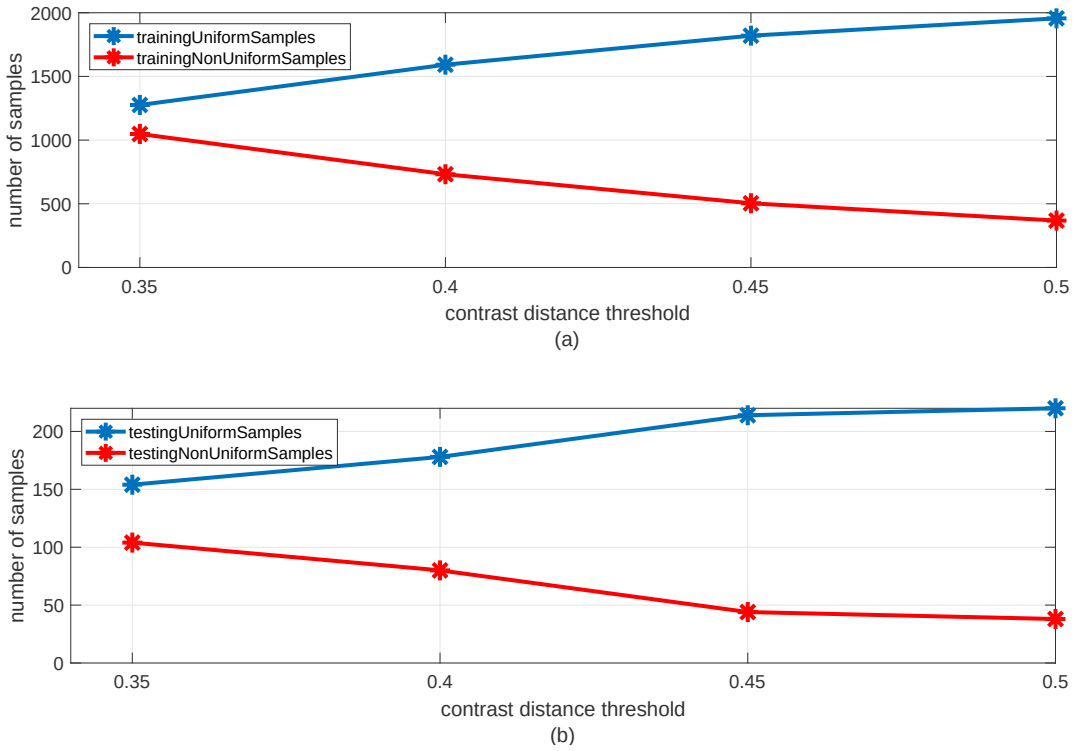


Figure 4.1: Number of samples vs. contrast distance threshold. (a) number of uniform/non-uniform used for training; (b) number of uniform/non-uniform target samples to predict.

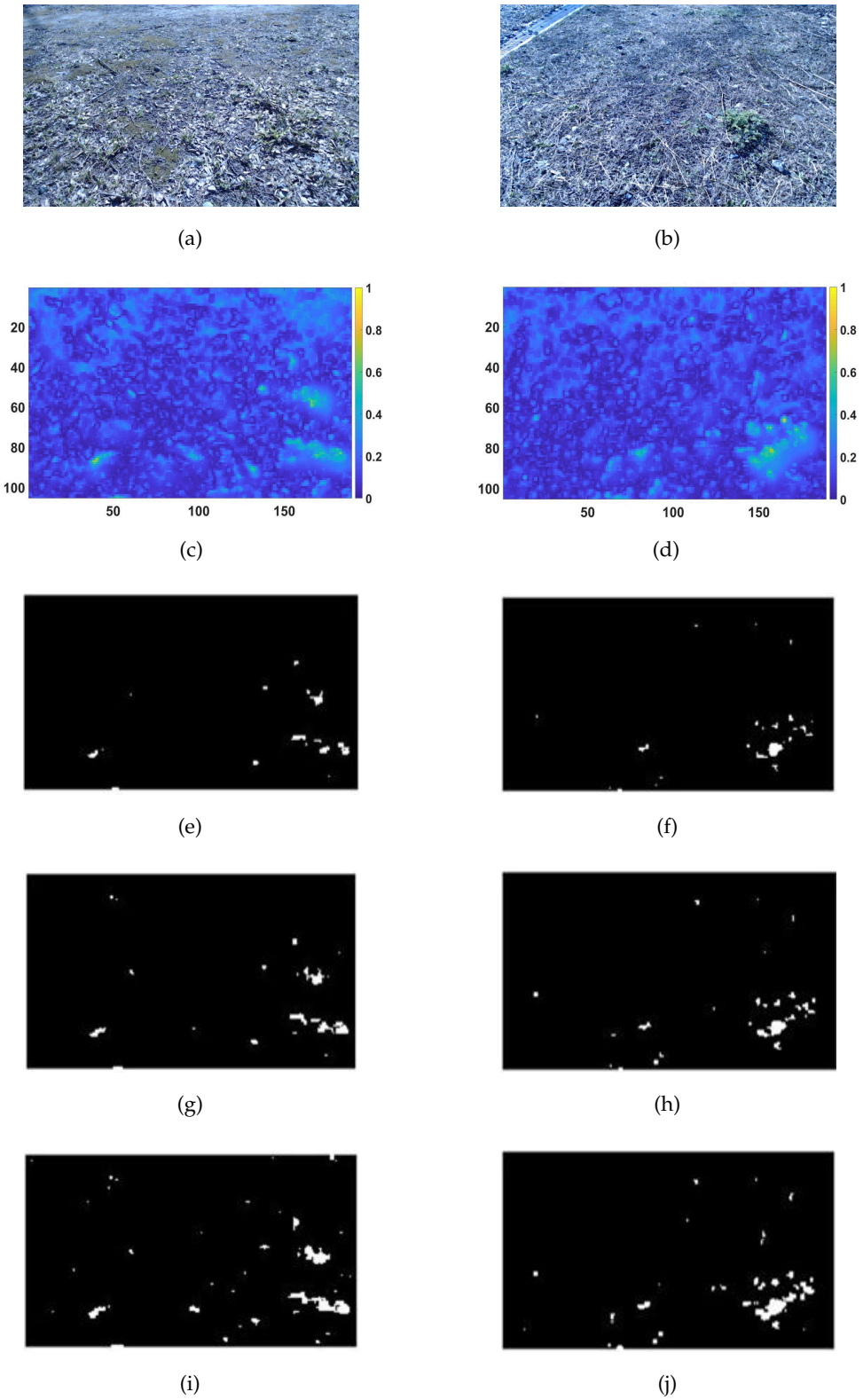


Figure 4.2: TNUD classification failure. (a) and (b) Grass terrain; (c) and (d) contrast distance map; (e) and (f) classification as uniform terrain with threshold value 0.5; (g) and (h) classification as non-uniform terrain with threshold value 0.45; (i) and (j) classification as non-uniform terrain with threshold value 0.35.

Table 4.2: Signal to noise ratio (SNR) of acceleration signal at rest.

Mean	Standard Deviation	SNR
1.0346	0.0101	102.6474
1.0339	0.0102	101.5350
1.0250	0.0096	106.7354
1.0335	0.0099	104.2580

4.2.3 SFTA and Motion Features

The acceleration signal does not undergo any noise filtering operation, and it is clear whether any filtering is performed at the sensor level. To measure the noise contribution to the measurements, signal-to-noise ratio (SNR) of the acceleration signal at rest was computed as follows:

$$SNR = \frac{\mu}{\sigma}, \tag{4.4}$$

where μ and σ are the mean and standard deviation of the acceleration signal, respectively. Results of SNR calculation for four signals are given in Table 4.2. SNR values are high which means that the noise does not impact severely our intended purpose.

the fractal dimension offers a discrimination of terrains texture attributes. Principal Component Analysis (PCA) was applied to SFTA feature vector to reduce the dimension from $\boldsymbol{x} \in \mathbb{R}^{16}$ to $\boldsymbol{z} \in \mathbb{R}^{2}$ to visualize the distribution of the motion information. As shown in Figure 4.3, terrain configurations are scattered in different areas.

4.2.4 Motion Prediction

In this section, we propose to study the effect of contrast distance threshold on motion prediction for uniform and non-uniform terrains, and weigh up performance

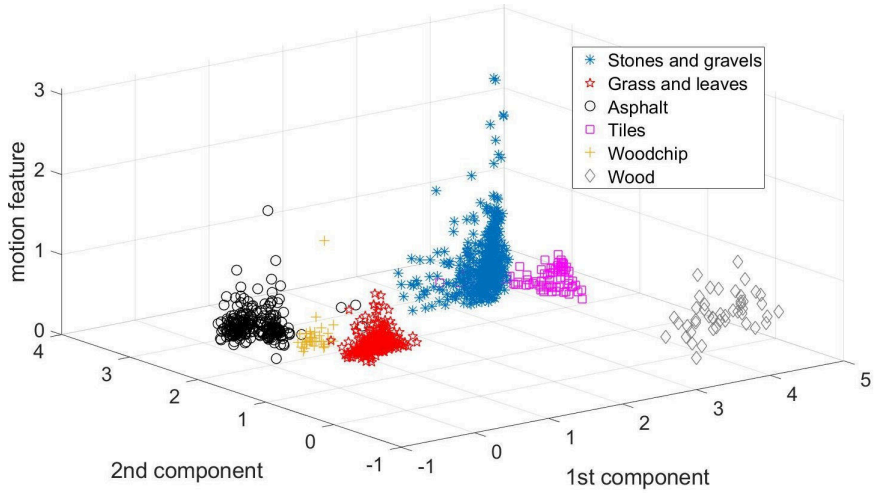


Figure 4.3: Distribution of motion information with regard to SFTA for six terrain types: stones and gravels, grass and leaves, granulated and slick asphalt.

Table 4.3: Influence of contrast distance threshold on prediction error for uniform/non-uniform terrains.

TNUD	Contrast Distance Threshold	0.35	0.4	0.45	0.5
	Uniform Terrains	0.3623	0.3075	0.2730	0.2373
	Non-uniform Terrains	0.4681	0.4411	0.3568	0.268
No TNUD	-	0.3567	0.3567	0.3567	0.3567

output with results obtained in (Bekhti and Kobayashi, 2016). We summarize the regression results in Table 4.3, Table 4.4 and plot Figure 4.4 to ease the interpretation. Table 4.3 introduces prediction error of motion information for uniform and non-uniform terrains with respect to different values of contrast distance threshold as well as prediction error without using TNUD. Additionally, Table 4.4 introduces prediction failure error of motion information for uniform and non-uniform terrains with respect to different values of contrast distance threshold as well as prediction failure error without using TNUD. From Figure 4.4, it is clear that a wrong choice of contrast distance threshold value leads to a weak prediction error performance. The weakest performance is observed at contrast distance threshold value equal to 0.3 where prediction error for uniform and non-uniform terrains is poorer than results obtained without employing TNUD. Same behavior is observed for prediction failure error output which is a lot higher than TNUD is not used. As discussed in section 4.2.2, a very low value of the contrast distance threshold prompts TNUD to classify terrains that are uniform into non-uniform terrains, and thus, impacts negatively the training process. The best performance is observed at a contrast distance threshold value equal to 0.5. We see a significant drop in prediction error for uniform and non-uniform compared with when TNUD is not used. Moreover, a notable reduction in prediction failure error for both uniform and non-unifom terrains compared to the classic framework where terrain discrimination is not applied.

For contrast distance threshold value equal to 0.5, prediction performance for uniform and non-uniform terrain samples are given by Figure 4.6 and 4.5, respectively. The current framework sometimes fails to predict the vibrations in the terrain samples judged by the multiscale analysis to be either uniform or non-uniform. In the case of non-uniform terrains, as shown in Figure 4.6, the Gaussian process outputs negative predictions for vibration since we are not

Table 4.4: Influence of contrast distance threshold on prediction failure error for uniform/non-uniform terrains.

TNUD	Contrast Distance Threshold	0.35	0.4	0.45	0.5
	Uniform Terrains	0.8004	0.6032	0.4686	0.4640
	Non-uniform Terrains	0.9181	1.0584	0.6765	0.4419
TNUD	-	0.6048	0.6048	0.6048	0.6048

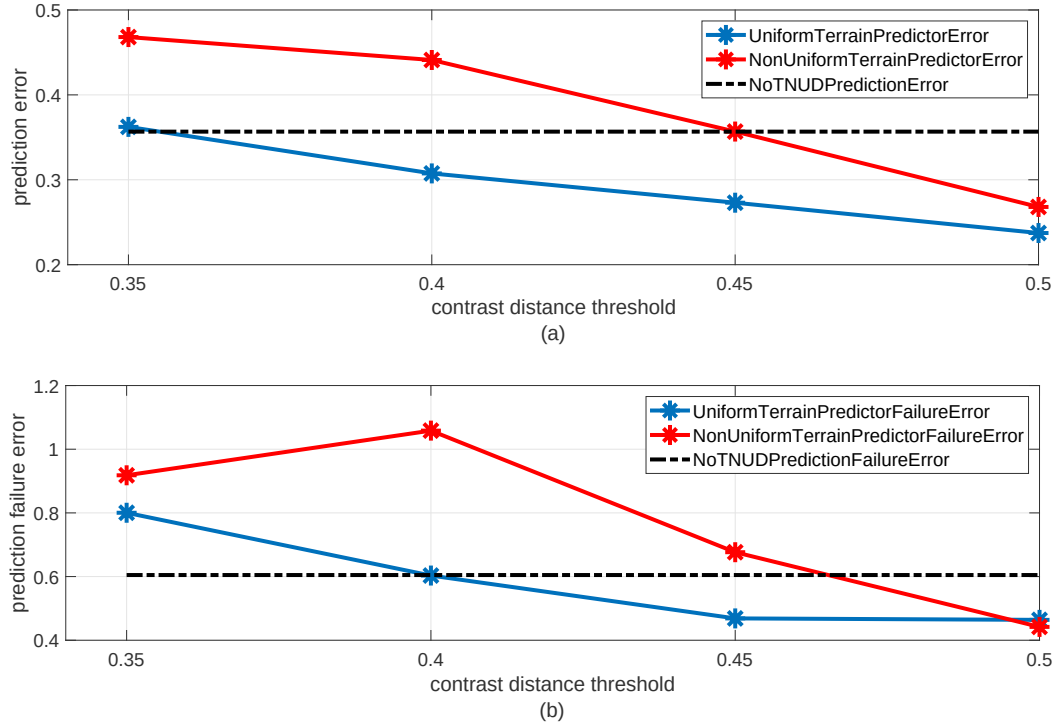


Figure 4.4: Influence of contrast distance threshold on prediction error for uniform/non-uniform terrains.

imposing any restrictions in this regard. Such negative vibration prediction is not consistent with the nature of the feature used in this study, which is always positive, as described by Equation (3.14).

Compared with (Matsumura, Bekhti, and Kobayashi, 2015), where motion prediction with varying speed based on the 3D reconstruction problem setting was investigated, our platform run at a constant speed of 0.2 m/s. We propose to study the effect of speed variations on the prediction of vibrations in future works.

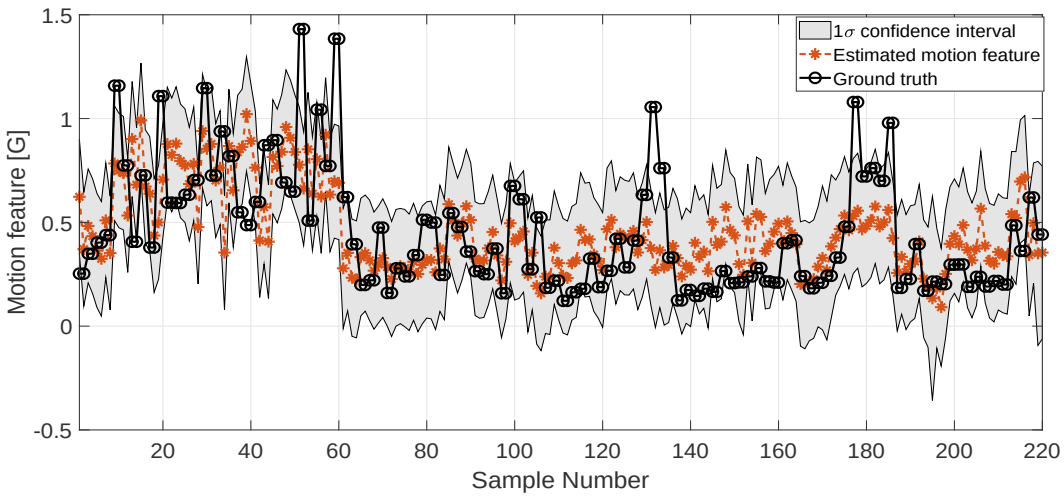


Figure 4.5: Prediction results for uniform terrains for contrast distance threshold of 0.5.

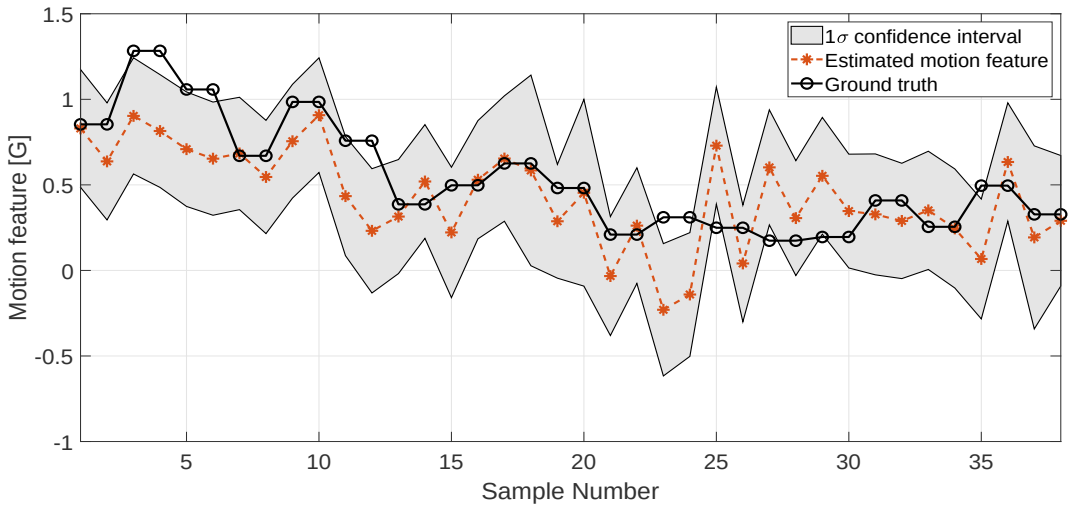


Figure 4.6: Prediction results for non-uniform terrains for contrast distance threshold of 0.5.

4.2.5 Computation Cost

The only computations that are performed on the image are the global contrast measurement and contrast distance maps for scales 1, 2, and 3. The purpose of doing so was to understand and evaluate the behavior of the contrast distance feature toward non-uniformity. Multiscale analysis for scales 1, 2, and 3 suffers a slightly high computation time, as shown in Table 4.5. A scale 1 contrast distance map was generated with an image of size 1080 $\times$ 1920, a scale 2 contrast distance map was generated with an image of size 540 $\times$ 960, and a scale 3 contrast distance map was generated with an image of size 270 $\times$ 480. Since we could confirm that non-uniform regions have a higher contrast distance feature value than uniform regions, we propose in future work to accelerate this process by focusing only on the ROI and not the whole image. The ROI is approximately 10% of the whole image; thus, the computational time will decrease drastically.

Table 4.5: Computation time.

Process	**Computation Time**
Multiscale analysis–contrast distance scale 1	880 ms
Multiscale analysis–contrast distance scale 2	200 ms
Multiscale analysis–contrast distance scale 3	47 ms
Multiscale analysis–fusion (refined contrast distance map)	2.2 ms
Texture extraction	67 ms
Motion feature	47 µs
Prediction for non-uniform terrains	100 ms
Predictor for uniform terrains	22 ms

Chapter 5

Conclusion

This thesis investigates terrain traversability prediction based on non-uniformity detection for uneven outdoor terrain configurations. The work presented has made contribution in improving prediction performance across several terrain natures. This final chapter summaries the research accomplished in this thesis and identifies potential directions for future research.

We proposed to examine how visual properties of terrain can be used to estimate a vibration traversability metric to translate the potential traversal load. Traversability metric is derived from correlation of proprioceptive sensor data obtained from wheel/soil interaction, and exteroceptive data obtained using a mono-vision camera system. An acceelration sensor mounted on our mobile robot acquired vertical acceleration that served to calculate motion feature to characterize the terrain *under foot*. The corresponding image undergoes our proposed TNUD method to detect non-uniformity. Using low level image feature namely contrast distance, we try to detect the irregularities in the terrain texture information. We argue that non-uniform in contrast to uniform surrounding areas translate into a high contrast value. Following this, SFTA descriptor is used to extract texture information from the image based on fractal dimension measurement. Using above mentioned pair of features, we proposed to train two individual terrain traversability predictor based on GP. This metric predictor was implemented on a skid-steering mobile robot to test effectiveness of the proposed scheme. TNUD relies on a cutoff value for the contrast distance threshold to decide on the nature of a terrain *i.e.* uniform or non-uniform. We investigated the influence of this threshold on the quality of discrimination offered by TNUD. It was shown that lower values of this threshold may result in wrong classification of terrain nature, and consequently, affect negatively results for metric prediction. Best results were obtained at a threshold value equal to 0.5, with a significant improvement compared to classic methods that do not used TNUD. Moreover, it was also shown that the texture descriptor based on fractals offers a good representation of terrain natures.

The current framework employs only vibrations to learn terrain nature. While traversing a terrain irregularity, the attitude of the mobile robot change. We think it would be interesting to make use of gyro recordings to strengthen our feature set to represent terrain nature. Moreover, vibration are highly correlated with the velocity

of travel and the architecture of the mobile robot. It would be interesting to integrate the mechanical properties into the prediction model, and investigate influence of travel velocity on the behavior of the mobile robot. We believe this would make the metric invariant to any speed changes.

References

Alexander, R. McNeill (2005). "Models and the scaling of energy costs for locomotion". In: *Journal of Experimental Biology* 208.9, pp. 1645–1652.

Angelova, Anelia et al. (2007). "Learning and prediction of slip from visual information". In: *Journal of Field Robotics* 24.3, pp. 205–231.

Appleby, Stephen (1996). "Multifractal Characterization of the Distribution Pattern of the Human Population". In: *Geographical Analysis* 28.2, pp. 147–160.

Arvidson, Raymond E. et al. (2017). "Mars Science Laboratory Curiosity Rover Megaripple Crossings up to Sol 710 in Gale Crater". In: *Journal of Field Robotics* 34.3, pp. 495–518.

Bajracharya, M. et al. (2008). "Learning long-range terrain classification for autonomous navigation". In: *2008 IEEE International Conference on Robotics and Automation*, pp. 4018–4024.

Bajracharya, Max et al. (2009a). "Autonomous off-road navigation with end-to-end learning for the LAGR program". In: *Journal of Field Robotics* 26.1, pp. 3–25.

— (2009b). "Autonomous off-road navigation with end-to-end learning for the LAGR program". In: *Journal of Field Robotics* 26.1, pp. 3–25.

Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool (2006). "Surf: Speeded up robust features". In: *In ECCV*, pp. 404–417.

Bekhti, M. A., Y. Kobayashi, and K. Matsumura (2014). "Terrain traversability analysis using multi-sensor data correlation by a mobile robot". In: *2014 IEEE/SICE International Symposium on System Integration*, pp. 615–620.

Bekhti, Mohammed Abdessamad and Yuichi Kobayashi (2016). "Prediction of Vibrations as a Measure of Terrain Traversability in Outdoor Structured and Natural Environments". In: *Image and Video Technology*, pp. 282–294.

Berczi, L., I. Posner, and T. D. Barfoot (2015). "Learning to assess terrain from human demonstration using an introspective Gaussian-process classifier". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3178–3185.

Berczi, Laszlo-Peter, Ingmar Posner, and Timothy D. Barfoot (2015). "Learning to assess terrain from human demonstration using an introspective Gaussian-process classifier". In: *IEEE International Conference on Robotics and Automation*.

Borenstein, J. and L. Feng (1996). "Gyrodometry: a new method for combining data from gyros and odometry in mobile robots". In: *Proceedings of IEEE International Conference on Robotics and Automation*. Vol. 1, 423–428 vol.1.

Bostani, A., A. Vakili, and T. A. Denidni (2008). "A novel method to measure and correct the odometry errors in mobile robots". In: *2008 Canadian Conference on Electrical and Computer Engineering*, pp. 897–900.

Bouguet, Jean-Yves (2015). *Camera Calibration Toolbox for Matlab*. URL: http://www.vision.caltech.edu/bouguetj/calib_doc/.

Bovik, A. C., M. Clark, and W. S. Geisler (1990). "Multichannel texture analysis using localized spatial filters". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.1, pp. 55–73.

Brooks, Christopher A. and Karl Iagnemma (2012). "Self-supervised terrain classification for planetary surface exploration rovers". In: *Journal of Field Robotics* 29.3, pp. 445–468.

Chaudhuri, B. B. and N. Sarkar (1995). "Texture segmentation using fractal dimension". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17.1, pp. 72–77.

Chavez-Garcia, R. O. et al. (2018). "Learning Ground Traversability From Simulations". In: *IEEE Robotics and Automation Letters* 3.3, pp. 1695–1702.

Clarke, Keith C. (1986). "Computation of the fractal dimension of topographic surfaces using the triangular prism surface area method". In: *Computers & Geosciences* 12.5, pp. 713 –722.

Collier, J. and A. Ramirez-Serrano (2009). "Environment Classification for Indoor/Outdoor Robotic Mapping". In: *2009 Canadian Conference on Computer and Robot Vision*, pp. 276–283.

Cordes, Florian, Frank Kirchner, and Ajish Babu (2018). "Design and field testing of a rover with an actively articulated suspension system in a Mars analog terrain". In: *Journal of Field Robotics* 35.7, pp. 1149–1181.

Cortes, J., L. Jaillet, and T. Simeon (2007). "Molecular Disassembly With Rrt-Like Algorithms". In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pp. 3301–3306.

Costa, A. F., G. Humpire-Mamani, and A. J. M. Traina (2012). "An Efficient Algorithm for Fractal Analysis of Textures". In: *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 39–46.

Cunningham, C. et al. (2017). "Locally-adaptive slip prediction for planetary rovers using Gaussian processes". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5487–5494.

Cunningham, Chris et al. (2015). "Predicting Terrain Traversability from Thermal Diffusivity". In: *Field and Service Robotics: Results of the 9th International Conference*, pp. 61–74.

Dahlhaus, Rainer (Dec. 1989). "Efficient Parameter Estimation for Self-Similar Processes". In: *Ann. Statist.* 17.4, pp. 1749–1766.

Devy, Michel et al. (2011). "Stereovision Algorithm to be Executed at 100Hz on a FPGA-Based Architecture". In: *Advances in Theory and Applications of Stereo Vision*. Chap. 17.

Diaz de Leon S., J. L. and J. H. Sossa A. (1998). "Automatic path planning for a mobile robot among obstacles of arbitrary shape". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 28.3, pp. 467–472.

Dijkstra, E. W. (1959). "A Note on Two Problems in Connexion with Graphs". In: *Numer. Math.* 1.1, pp. 269–271.

Dongshin Kim et al. (2006). "Traversability classification using unsupervised on-line visual learning for outdoor robot navigation". In: *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* Pp. 518–525.

DuPont, Edmond M. et al. (2008). "Frequency response method for terrain classification in autonomous ground vehicles". In: *Autonomous Robots* 24, pp. 337–347.

Engelke, U., V. X. Nguyen, and H. Zepernick (2008). "Regional attention to structural degradations for perceptual image quality metric design". In: *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 869–872.

Filitchkin, P. and K. Byl (2012). "Feature-based terrain classification for LittleDog". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1387–1392.

Gallego, Antonio Javier et al. (2007). "3D Reconstruction and Mapping from Stereo Pairs with Geometrical Rectification". In: *Advances in Brain, Vision, and Artificial Intelligence*, pp. 318–327.

Goldberg, S. B., M. W. Maimone, and L. Matthies (2002). "Stereo vision and rover navigation software for planetary exploration". In: *Proceedings, IEEE Aerospace Conference*. Vol. 5, pp. 5–5.

Goodchild, Michael (Apr. 1980). "Fractals and the Accuracy of Geographical Measures". In: *Journal of the International Association for Mathematical Geology* 12, pp. 85–98.

Gopalakrishnan, V., Y. Hu, and D. Rajan (2009). "Salient Region Detection by Modeling Distributions of Color and Orientation". In: *IEEE Transactions on Multimedia* 11.5, pp. 892–905.

Guimarães, P. P. S. et al. (2016). "A Bio-inspired Apodal and Modular Robot". In: *2016 XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR)*, pp. 61–66.

Hadsell, Raia et al. (2008). "Online learning for offroad robots: Using spatial label propagation to learn long-range traversability". In: *Robotics: Science and Systems III*. Vol. 3, pp. 17–23.

Hadsell, Raia et al. (Feb. 2009). "Learning Long-Range Vision for Autonomous off-Road Driving". In: *J. Field Robot.* 26.2, pp. 120–144. ISSN: 1556-4959.

Halatci, I., C. A. Brooks, and K. Iagnemma (2007). "Terrain Classification and Classifier Fusion for Planetary Exploration Rovers". In: *2007 IEEE Aerospace Conference*, pp. 1–11.

Happold, Michael, Mark Ollis, and Nikolas Johnson (Aug. 2006). "Enhancing Supervised Terrain Classification with Predictive Unsupervised Learning". In:

Haralick, R. M. (1979). "Statistical and structural approaches to texture". In: *Proceedings of the IEEE* 67.5, pp. 786–804.

Hart, P. E., N. J. Nilsson, and B. Raphael (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths". In: *IEEE Transactions on Systems Science and Cybernetics* 4.2, pp. 100–107.

Hart, Peter E., Nils J. Nilsson, and Bertram Raphael (1972). "Correction to a Formal Basis for the Heuristic Determination of Minimum Cost Paths". In: *SIGART Bull.*, pp. 28–29.

Hartley, Richard and Andrew Zisserman (2003). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Helmick, Daniel, Anelia Angelova, and Larry Matthies (2009). "Terrain Adaptive Navigation for planetary rovers". In: *Journal of Field Robotics* 26.4, pp. 391–410.

Ho, K., T. Peynot, and S. Sukkarieh (2013a). "A near-to-far non-parametric learning approach for estimating traversability in deformable terrain". In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2827–2833.

— (2013b). "Traversability estimation for a planetary rover via experimental kernel learning in a Gaussian process framework". In: *2013 IEEE International Conference on Robotics and Automation*, pp. 3475–3482.

Ho, Ken, Thierry Peynot, and Salah Sukkarieh (2016). "Nonparametric Traversability Estimation in Partially Occluded and Deformable Terrain". In: *Journal of Field Robotics* 33.8, pp. 1131–1158.

Hong, Tsai Hong et al. (2002). "Road detection and tracking for autonomous mobile robots". In: *Unmanned Ground Vehicle Technology IV*. Vol. 4715, pp. 311 –319.

Howard, Andrew et al. (2006). "Towards learned traversability for robot navigation: From underfoot to the far field". In: *Journal of Field Robotics* 23, pp. 1005–1017.

Howard, Ayanna and Homayoun Seraji (2001). "Vision-based terrain characterization and traversability assessment". In: *Journal of Robotic Systems* 18.10, pp. 577–587.

Hu, Yiqun et al. (2004). "Salient Region Detection Using Weighted Feature Maps Based on the Human Visual Attention Model". In: *Proceedings of the 5th Pacific Rim Conference on Advances in Multimedia Information Processing - Volume Part II*, pp. 993–1000.

Hudjakov, R. and M. Tamre (2009). "Aerial imagery terrain classification for long-range autonomous navigation". In: *2009 International Symposium on Optomechatronic Technologies*, pp. 88–91.

Huntsberger, Terry et al. (2007). "TRESSA: Teamed robots for exploration and science on steep areas". In: *Journal of Field Robotics* 24.11-12, pp. 1015–1031.

Iagnemma, K. et al. (2004). "Online terrain parameter estimation for wheeled mobile robots with application to planetary rovers". In: *IEEE Transactions on Robotics* 20.5, pp. 921–927.

Iagnemma, Karl D. and Steven Dubowsky (2002). "Terrain estimation for high-speed rough-terrain autonomous vehicle navigation". In: *Unmanned Ground Vehicle*

Technology IV. Ed. by Grant R. Gerhart, Chuck M. Shoemaker, and Douglas W. Gage. Vol. 4715, pp. 256 –266.

Ishikawa, J. (2005). "Evaluation of Test Results of GPR-based Anti-personnel Landmine Detection Systems Mounted on Robotic Vehicles". In: *Proceedings of the IARP International Workshop on Robotics and Mechanical Assis-tance in Humanitarian Demining, (2005-6)*.

Jackel, L. D. et al. (2006). "The DARPA LAGR program: Goals, challenges, methodology, and phase I results". In: *Journal of Field Robotics* 23.11-12, pp. 945–973.

Jarrault, Pierre, Christophe Grand, and Philippe Bidaud (2010). "Large Obstacle Clearance Using Kinematic Reconfigurability for a Rover with an Active Suspension". In: *Emerging Trends in Mobile Robotics*, pp. 114–121.

Julesz, Bela (1975). "Experiments in the Visual Perception of Texture". In: *Scientific American* 232.4, pp. 34–43.

Kamimura, Akiya et al. (2002). "A Self-Reconfigurable Modular Robot (MTRAN) — Hardware and Motion Planning Software —". In: *Distributed Autonomous Robotic Systems 5*. Ed. by Hajime Asama et al., pp. 17–26.

Kaplan, L. M. and C. . J. Kuo (1995). "Texture roughness analysis and synthesis via extended self-similar (ESS) model". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17.11, pp. 1043–1056.

Khan, Y. N., A. Masselli, and A. Zell (2012). "Visual terrain classification by flying robots". In: *2012 IEEE International Conference on Robotics and Automation*, pp. 498–503.

Kim, D., W. Chung, and S. Park (2010). "Practical motion planning for car-parking control in narrow environment". In: *IET Control Theory Applications* 4.1, pp. 129–139.

Korf, Richard E. (2010). "Artificial Intelligence Search Algorithms". In: *Algorithms and Theory of Computation Handbook: Special Topics and Techniques*, p. 22.

Kottege, N. et al. (2015). "Energetics-informed hexapod gait transitions across terrains". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5140–5147.

Krebs, Ambroise, Cédric Pradalier, and Roland Siegwart (2010). "Adaptive rover behavior based on online empirical evaluation: Rover-terrain interaction and near-to-far learning". In: *Journal of Field Robotics* 27.2, pp. 158–180.

Lacroix, Simon and Il-Kyun Jung (2004). "Simultaneous localization and mapping with stereovision". In: *IFAC Proceedings Volumes* 37.8, pp. 874 –879.

Lacroix, Simon et al. (2001). "Autonomous Rover Navigation on Unknown Terrains Functions and Integration". In: *Experimental Robotics VII*. Ed. by Daniela Rus and Sanjiv Singh, pp. 501–510.

Lacroix, Simon et al. (Oct. 2002). "Autonomous Rover Navigation on Unknown Terrains: Functions and Integration". In: *International Journal of Robotic Research - IJRR* 21.

Larionova, S., L. Marques, and A. T. De Almeida (2006). "Detection of Natural Landmarks for Mapping by a Demining Robot". In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4959–4964.

LaValle, Steven M. (2006). *Planning Algorithms*. Cambridge University Press.

Lerski, Richard A. et al. (1993). "MR image texture analysis–an approach to tissue characterization." In: *Magnetic resonance imaging* 11 6, pp. 873–87.

Liao, Ping-Sung, Tse-Sheng Chen, and Pau-Choo Chung (2001). "A Fast Algorithm for Multilevel Thresholding". In: *J. Inf. Sci. Eng.* 17.5, pp. 713–727.

Lowe, David G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints". In: *Int. J. Comput. Vision* 60.2, pp. 91–110.

Lozano-Perez (1983). "Spatial Planning: A Configuration Space Approach". In: *IEEE Transactions on Computers* C-32.2, pp. 108–120.

Lu, Chun S., Pau C. Chung, and Chih F. Chen (1997). "Unsupervised texture segmentation via wavelet transform". In: *Pattern Recognition* 30.5, pp. 729 –742.

Ma, Yu-Fei and Hong-Jiang Zhang (2003). "Contrast-Based Image Attention Analysis by Using Fuzzy Growing". In: *Proceedings of the Eleventh ACM International Conference on Multimedia*, pp. 374–381.

Mallet, A., S. Lacroix, and L. Gallo (2000). "Position estimation in outdoor environments using pixel tracking and stereovision". In: *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation*. Vol. 4, pp. 3519–3524.

Mandelbrot, Benoit (1967). "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension". In: 156.3775, pp. 636–638.

Mandelbrot, Benoit B. and John W. Van Ness (1968). "Fractional Brownian Motions, Fractional Noises and Applications". In: *SIAM Review* 10.4, pp. 422–437.

Manduchi, R. et al. (2005). "Obstacle Detection and Terrain Classification for Autonomous Off-Road Navigation". In: *Auton. Robots* 18.1, pp. 81–102.

Marder-Eppstein, E. et al. (2010). "The Office Marathon: Robust navigation in an indoor office environment". In: *2010 IEEE International Conference on Robotics and Automation*, pp. 300–307.

Martin, S. and P. Corke (2014). "Long-term exploration tours for energy constrained robots with online proprioceptive traversability estimation". In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5778–5785.

Matsumura, Kazuki, Mohammed Abdessamad Bekhti, and Yuichi Kobayashi (2015). "Prediction of motion over traversable obstacles for autonomous mobile robot based on 3D reconstruction and running information". In: *The Abstracts of the international conference on advanced mechatronics : toward evolutionary fusion of IT and mechatronics : ICAM* 2015.6, p. 205.

Matthies, L. and A. Rankin (2003). "Negative obstacle detection by thermal signature". In: *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*. Vol. 1, 906–913 vol.1.

Mazhar, H. et al. (2013). "CHRONO: a parallel multi-physics library for rigid-body, flexible-body, and fluid dynamics". In: *Mechanical Sciences* 4.1, pp. 49–64.

Metka, Benjamin, Mathias Franzius, and Ute Bauer-Wersing (2013). "Outdoor Self-Localization of a Mobile Robot Using Slow Feature Analysis". In: *Neural Information Processing*, pp. 249–256.

Mitchell, Thomas M. (1997). *Machine Learning*. 1st ed. McGraw-Hill, Inc.

Mulgaonkar, Y. et al. (2016). "The flying monkey: A mesoscale robot that can run, fly, and grasp". In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4672–4679.

Munoz, D., N. Vandapel, and M. Hebert (2009). "Onboard contextual classification of 3-D point clouds with learned high-order Markov Random Fields". In: *2009 IEEE International Conference on Robotics and Automation*, pp. 2009–2016.

Murphy, L. and P. Newman (2010). "Planning most-likely paths from overhead imagery". In: *2010 IEEE International Conference on Robotics and Automation*, pp. 3059–3064.

Nishii, Jun (2006). "An analytical estimation of the energy cost for legged locomotion". In: *Journal of Theoretical Biology* 238.3, pp. 636 –645.

Normant, Fran çois and Claude Tricot (1991). "Method for evaluating the fractal dimension of curves using convex hulls". In: *Phys. Rev. A* 43 (12), pp. 6518–6525.

Ojeda, Lauro et al. (2006). "Terrain characterization and classification with a mobile robot". In: *Journal of Field Robotics* 23.2, pp. 103–122.

Ordonez, C. and E. G. Collins (2008). "Rut Detection for Mobile Robots". In: *2008 40th Southeastern Symposium on System Theory (SSST)*, pp. 334–337.

Osberger, W., N. Bergmann, and A. Maeder (1998). "An automatic image quality assessment technique incorporating higher level perceptual factors". In: *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)*. Vol. 3, pp. 414–418.

Papadakis, Panagiotis (2013). "Terrain traversability analysis methods for unmanned ground vehicles: A survey". In: *Engineering Applications of Artificial Intelligence* 26.4, pp. 1373 –1385.

Papadopoulos, E. G. and D. A. Rey (1996). "A new measure of tipover stability margin for mobile manipulators". In: *Proceedings of IEEE International Conference on Robotics and Automation*. Vol. 4, pp. 3111–3116.

Pentland, A. P. (1984). "Fractal-Based Description of Natural Scenes". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6.6, pp. 661–674.

Pirjanian, P. et al. (2002). "Distributed control for a modular, reconfigurable cliff robot". In: *Proceedings 2002 IEEE International Conference on Robotics and Automation*. Vol. 4, pp. 4083–4088.

Pruess, S. A. (1995). "Some Remarks on the Numerical Estimation of Fractal Dimension". In: *Fractals in the Earth Sciences*, pp. 65–75.

Rasmussen, Carl Edward and Christopher K. I. Williams (2006). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. Adaptive Computation and Machine Learning. The MIT Press.

Rosenfeld, Azriel (1977). "Digital Image Processing and Recognition". In: *Digitale Bildverarbeitung Digital Image Procesing*, pp. 1–11.

Russell, David A., James D. Hanson, and Edward Ott (1980). "Dimension of Strange Attractors". In: *Phys. Rev. Lett.* 45 (14), pp. 1175–1178.

Rutishauser, U. et al. (2004). "Is bottom-up attention useful for object recognition?" In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 2.

Ryde, J. and M. Brünig (2009). "Non-cubic occupied voxel lists for robot maps". In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4771–4776.

Sadhukhan, Debangshu and Carl A. Moore (2004). "Terrain estimation using internal sensors". In: *in Proc. of the IASTED Int. Conf. Robot. Appl*, pp. 447–800.

Sandau, K. and H. Kurz (1997). "Measuring fractal dimension and complexity - an alternative approach with an application". In: *Journal of Microscopy* 186.2, pp. 164–176.

Saut, J. et al. (2007). "Dexterous manipulation planning using probabilistic roadmaps in continuous grasp subspaces". In: *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2907–2912.

Schaal, Stefan and Christopher G. Atkeson (1998). "Constructive Incremental Learning from Only Local Information". In: *Neural Computation* 10.8, pp. 2047–2084.

Schultz, Peter H. et al. (2010). "The LCROSS Cratering Experiment". In: 330.6003, pp. 468–472.

Schwartz, Jacob T. and Micha Sharir (1983). "On the Piano Movers' Problem: III. Coordinating the Motion of Several Independent Bodies: The Special Case of Circular Bodies Moving Amidst Polygonal Barriers". In: *The International Journal of Robotics Research* 2.3, pp. 46–75.

Seraji, H. and A. Howard (2002). "Behavior-based robot navigation on challenging terrain: A fuzzy logic approach". In: *IEEE Transactions on Robotics and Automation* 18.3, pp. 308–321.

Shamah, Benjamin et al. (2001). "Steering and Control of a Passively Articulated Robot". In: *Proceedings of SPIE, Sensor Fusion and Decentralized Control in Robotic Systems IV*. Vol. 4571.

Shelberg, Mark, Nina Lam, and Harold Moellering (1983). "Measuring the Fractal Dimensions of Surfaces". In:

Shin, Youjin et al. (2013). "Autonomous Navigation of a Surveillance Robot in Harsh Outdoor Road Environments". In: *Advances in Mechanical Engineering* 5.

Siagian, C., C. Chang, and L. Itti (2013). "Mobile robot navigation system in outdoor pedestrian environment using vision-based road recognition". In: *2013 IEEE International Conference on Robotics and Automation*, pp. 564–571.

Silver, D. et al. (2006). "Experimental Analysis of Overhead Data Processing To Support Long Range Navigation". In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2443–2450.

Sofman, Boris et al. (2006a). "Improving robot navigation through self-supervised online learning". In: *Journal of Field Robotics* 23.11-12, pp. 1059–1075.

Sofman, Boris et al. (2006b). *Terrain Classification from Aerial Data to Support Ground Vehicle Navigation*. Tech. rep. Carnegie Mellon University.

Soille, Pierre and Jean-F. Rivest (1996). "On the Validity of Fractal Dimension Measurements in Image Analysis". In: *Journal of Visual Communication and Image Representation* 7.3, pp. 217 –229.

Stavens, David and Sebastian Thrun (2006). "A Self-Supervised Terrain Roughness Estimator for off-Road Autonomous Driving". In: *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, pp. 469–476.

Stelzer, Annett, Heiko Hirschmüller, and Martin Görner (2012). "Stereo-vision-based navigation of a six-legged walking robot in unknown rough terrain". In: *The International Journal of Robotics Research* 31.4, pp. 381–402.

Stentz, Anthony and Is Carnegle Mellon (1993). "Optimal and Efficient Path Planning for Unknown and Dynamic Environments". In: *International Journal of Robotics and Automation* 10, pp. 89–100.

Stentz, Athony (1993). "Optimal and Efficient Path Planning for Unknown and Dynamic Environments". In: *International Journal of Robotics and Automation* 10, pp. 89–100.

Suger, B., B. Steder, and W. Burgard (2015). "Traversability analysis for mobile robots in outdoor environments: A semi-supervised learning approach based on 3D-lidar data". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3941–3946.

Sukkarieh, S., E. M. Nebot, and H. F. Durrant-Whyte (1999). "A high integrity IMU/GPS navigation loop for autonomous land vehicle applications". In: *IEEE Transactions on Robotics and Automation* 15.3, pp. 572–578.

Tarokh, M. and G. J. McDermott (2005). "Kinematics modeling and analyses of articulated rovers". In: *IEEE Transactions on Robotics* 21.4, pp. 539–553.

Tucker, V. A. (1975). "The Energetic Cost of Moving About: Walking and running are extremely inefficient forms of locomotion. Much greater efficiency is achieved by birds, fish-and bicyclists". In: *American Scientist* 63.4, pp. 413–419.

Valkealahti, K. and E. Oja (1998). "Reduced multidimensional co-occurrence histograms in texture classification". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.1, pp. 90–94.

Vandapel, N. et al. (2004). "Natural terrain classification using 3-d ladar data". In: *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*. Vol. 5, pp. 5117–5122.

Vernaza, P., B. Taskar, and D. D. Lee (2008). "Online, self-supervised terrain classification via discriminatively trained submodular Markov random fields". In: *2008 IEEE International Conference on Robotics and Automation*, pp. 2750–2757.

Vu, Cuong T. and Damon M. Chandler (2011). "Main subject detection via adaptive feature refinement". In: *Journal of Electronic Imaging* 20.1, pp. 1–21.

Wagner, M. D. et al. (2001). "The Science Autonomy System of the Nomad robot". In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation*. Vol. 2, 1742–1749 vol.2.

Walther, Dirk et al. (2002). "Attentional Selection for Object Recognition — A Gentle Way". In: *Biologically Motivated Computer Vision*, pp. 472–479.

Wang, Junle, Damon M. Chandler, and Patrick Le Callet (2010). "Quantifying the relationship between visual salience and visual importance". In: *Human Vision and Electronic Imaging XV*. Vol. 7527, pp. 160 –168.

Ward, Chris C. and Karl Iagnemma (2009). "Speed-independent vibration-based terrain classification for passenger vehicles". In: *Vehicle System Dynamics* 47.9, pp. 1095–1113.

Wei, Longsheng, Nong Sang, and Yuehuan Wang (2009). "Variable resolution image compression based on a model of visual attention". In: *MIPPR 2009: Automatic Target Recognition and Image Analysis*. Vol. 7495, pp. 188 –194.

Weiss, C., H. Frohlich, and A. Zell (2006). "Vibration-based Terrain Classification Using Support Vector Machines". In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4429–4434.

Weiss, Christian et al. (2007). "Comparison of Different Approaches to Vibration-based Terrain Classification". In: *European Conference on Mobile Robots*.

Wellington, C. and A. Stentz (2004). "Online adaptive rough-terrain navigation vegetation". In: *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*. Vol. 1, 96–101 Vol.1.

Wellington, Carl, Aaron Courville, and Anthony Stentz (June 2005). "Interacting Markov Random Fields for Simultaneous Terrain Modeling and Obstacle Detection". In: pp. 1–8.

Wellington, Carl and Anthony Stentz (Jan. 2003). "Learning Predictions of the Load-Bearing Surface for Autonomous Rough-Terrain Navigation in Vegetation". In: vol. 24, pp. 83–92.

Wermelinger, M. et al. (2016). "Navigation planning for legged robots in challenging terrain". In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1184–1189.

Wettergreen, D., H. Thomas, and M. Bualat (1997). "Initial results from vision-based control of the Ames Marsokhod rover". In: *Proceedings of the 1997 IEEE/RSJ*

International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications. IROS '97. Vol. 3, 1377–1382 vol.3.

Xi, Weitao, Yevgeniy Yesilevskiy, and C. David Remy (2016). "Selecting gaits for economical locomotion of legged robots". In: *The International Journal of Robotics Research* 35.9, pp. 1140–1154.

Yu, Stella X. and Dimitri A. Lisin (2009). "Image Compression Based on Visual Saliency at Individual Scales". In: *Advances in Visual Computing*, pp. 157–166.

Zhou, Feng et al. (2014). "Simulations of Mars Rover Traverses". In: *Journal of Field Robotics* 31.1, pp. 141–160.