The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

University of Cape Town

# Drivable Region Detection for Autonomous Robots Applied to South African Underground Mining

(Omowunmi Elizabeth, **FALOLA** (wunmi.falola@uct.ac.za))

University of Cape Town

This dissertation is submitted in fulfillment of the academic requirements

for the degree of

Master of Science in Computer Science

In the Faculty of Science
University of CapeTown

November 28, 2012

University of Cape Town

# CERTIFICATION

As the candidate's supervisors, we have approved this dissertation for submission.

Supervisor: Dr. Antoine Bagula

Co-supervisor: Prof. Isaac Osunmakinde (University of South Africa)

Signature: ________________

Date: ________________

University of Cape Town

# DECLARATION

I declare that this dissertation is my own work. Where collaboration with other people has taken place, or material generated by other researchers is included, the parties and/or materials are indicated in the acknowledgements or are explicitly stated with references as appropriate.

This work is being submitted for the degree of Master of Science in Computer Science at the University of Cape Town. It has not been submitted to any other university for any other degree or examination.

University of Cape Town

Omowunmi Elizabeth FALOLA

Date

# DEDICATION

To the Almighty God, the master planner, the author and the finisher of faith, whose grace has been sufficient for me.

University of Cape Town

# Abstract

Mining booms have created significant wealth for various countries, including South Africa. The South African mining industry forms a crucial part of the national economic sector. However, the mining industry is constantly faced with the twin needs for safety and improved productivity. In the history of mining, the mortality rate of mine workers has been alarming. Many miners have been reported dead and several hundreds have sustained injuries. Furthermore, the death toll in South African mines is high compared to other major mining countries. Consequently, much effort is currently being directed towards improving safety in South African underground mines.

Autonomous robots exhibit the potential to help address safety issues in mines. It is widely recognised that robots can play a significant role in pre-disaster (pre-emption) and post-disaster (recovery) mine rescue operations. This would inevitably enhance productivity and greatly reduce human exposure to dangerous underground mine environments. Nonetheless, the success of a robot in the mine depends greatly on its visual capability to interpret its immediate environment correctly for navigational purposes. A good visual interpretation of the environment is critical to safe autonomous navigation within the mine. Much effort is currently being directed at South African underground mine safety; however a completely autonomous safety solution in underground mine environments is yet to be realised.

A number of research projects that have been carried out in the field of autonomous navigation and drivable region detection mostly focus on structured and unstructured surface (above-ground) detection. Most of this research presents classification results only in the two-dimensional (2D) plane, which pretends that the world is flat. Furthermore, the few known studies on underground experiments do not fully explore three-dimensional (3D) drivability analysis, which gives

University of Cape Town

an accurate orientation of an image scene analysis.

This dissertation focuses on enhancing autonomous robots' capability to identify drivable regions in underground terrains. A system model that compares the drivability analysis of underground terrains using the entropy model and statistical region merging (SRM) was developed, with a view to presenting an analysis of 2D and 3D results. The approach involves standard image-processing techniques, such as colour and texture feature extraction and region segmentation for underground image classification. A probabilistic method based on the local entropy was employed. The entropy is measured within a fixed window on each frame in order to compute features used in the segmentation process. This research compares the results obtained from the entropy method and SRM approach. Performance evaluation is carried out to provide useful qualitative and quantitative conclusions.

An investigation is conducted on different regions of the mine, such as the shaft, stope and gallery, using publicly available mine frames, with a stream of locally captured mine images. An investigation is also conducted on a stream of underground tunnel image frames, using the XBOX Kinect 3D sensors. The Kinect sensors produce streams of red, green and blue (RGB) and depth images of 640 x 480 resolution at 30 frames per second. Integrating the depth information to drivability gives a strong cue to the analysis, which detects 3D results augmenting drivable and non-drivable regions in 2D.

The results of the 2D and 3D experiment with different terrains, mines and tunnels, together with the qualitative and quantitative evaluation, reveal that a good drivable region can be detected in underground terrains. Since the promising results are benchmarked with publicly available image frames and methods, this work would contribute to further autonomous path planning and navigation in underground terrains in order to facilitate autonomous safety tasks.

University of Cape Town

# Acknowledgements

My profound gratitude goes to the Almighty God, the author and finisher of faith, the greatest time keeper, the master planner, whose love, mercy and grace have been sufficient for me throughout the research period. God, I give you all the glory.

This research would not have been successful without the support of my supervisors, Dr Antoine Bagula of the University of Cape Town (UCT) and Prof. Isaac Osunmakinde of the University of South Africa (UNISA). Thank you for the constructive criticism during each phase of this research work. Also, to the authorities at the University of CapeTown, the African Institute for Mathematical Sciences (AIMS) and the Google Anita Borg Memorial Scholarship, I say thank you for the financial opportunities given to me during my study. Much appreciation to the UCT Robotics and Agents Research Laboratory for the useful information supplied on the XBOX Kinect sensors device.

I am forever indebted to my parents, Mr and Mrs. Falola, and siblings, for their constant support and encouragement throughout my academic pursuit. I cannot imagine myself without you. With a sincere heart, I appreciate my better half, the larger part of me, Dr. Adeniyi Isafiade, and my precious angel, Isafiade Grace Oluwatamilore; you are wonderful, thank you for your unending love and encouragement, may the Lord continue to nurture us on every side. On a resounding note, I love you all.

University of Cape Town

# Contents



University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

# List of Figures



University of Cape Town

University of Cape Town

University of Cape Town

# GLOSSARY

**Drivable Region:** A safe region, within an image frame, for autonomous navigation

**Mine Frame:** A section of an underground mine video stream, usually an image snapshot.

**Mine Shaft:** vertical excavations sunk adjacent to an ore body.

**Mine Stope:** An excavation in the form of steps made by the mining of ore from steeply inclined or vertical veins.

**Mine Gallery:** A long enclosed passage, such as a hallway or corridor.

**Ore Body:** Material extracted from underground mining.

**Robotic Platform:** A means of deploying robots for certain tasks.

**Robots' Sensor:** A device used by robots to sense/view/perceive an environment.

**Robots' Mission:** A task assigned to an autonomous robot.

**Safety Potentials:** Capabilities displayed by robots for safety tasks.

**Video Stream:** Video image data captured in an environment, an unbroken flow (succession).

**Vision Model:** A model for robots' visual sensation.

**Vision Task:** A function to be performed for something to be seen.

**Visual Sensation:** Robots' perceptual experience of seeing.

**3D plane:** 3 dimensional cordinates, likened to real world $(x, y, z)$ cordinate system.

University of Cape Town

# Chapter 1

# INTRODUCTION

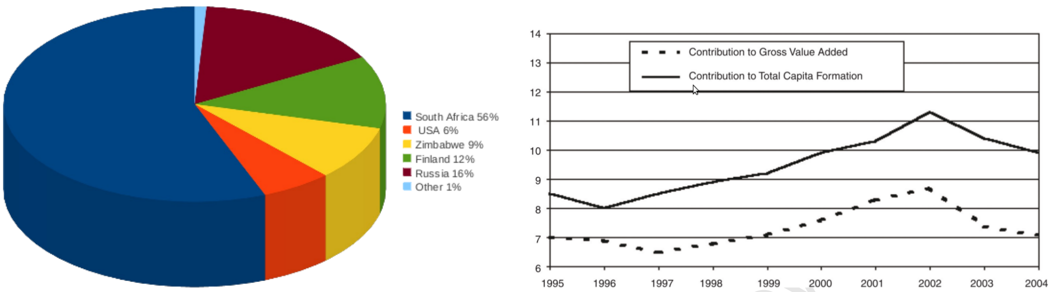
University of Cape Town

This chapter gives a general introduction to this study and the motivation for this work. The objectives and goals of this research are also stated. The motivation for this research is the twofold needs for safety and productivity faced by the mining industry, which have not been explored in depth by the research community. Furthermore, an overview of the research organisation is documented in order to orient the intended audience about current research.

## 1.1 Introduction

The mining industry in South Africa forms a crucial part of the national economic sector. The industry is a world-class one and a cornerstone of the South African economy [21]. According to Figure 1.1 , South Africa supplies over 80% of the world's platinum group metals (PGMs) need. The industry makes a significant contribution to the overall economic activity [69] which includes, but is not limited to, the production of useful minerals, foreign exchange earnings and the development of meaningful job opportunities for several thousands of South African citizens. In addition to the aforementioned benefits, the South African mining industry employed 493,000 workers as of 2007 and the industry represents 18% of South Africa's $588 billion USD gross domestic product (GDP). Thus mining remains crucial for the country's economy [34] [71]. Nonetheless, the industry is constantly faced with the needs for safety and efficient productivity

[37]. In the quest to address safety issues in the mines, autonomous robots could play an important role. Robots can be used for checkpoint and safety inspection tasks in mines. However, an effective vision model is critical for safe autonomous navigation in an underground terrain. This work serves to assist robots' mission in an underground mine by enhancing autonomous robots' capability to identify drivable regions within the mines.




(a) Identified World Resources: PGMs (South Africa Minerals Bureau)

(b) Mining Contribution to Gross Value Added and Fixed Capital Formation

Figure 1.1: South African Mining's Contribution to the Economy (Department of Minerals and Energy)

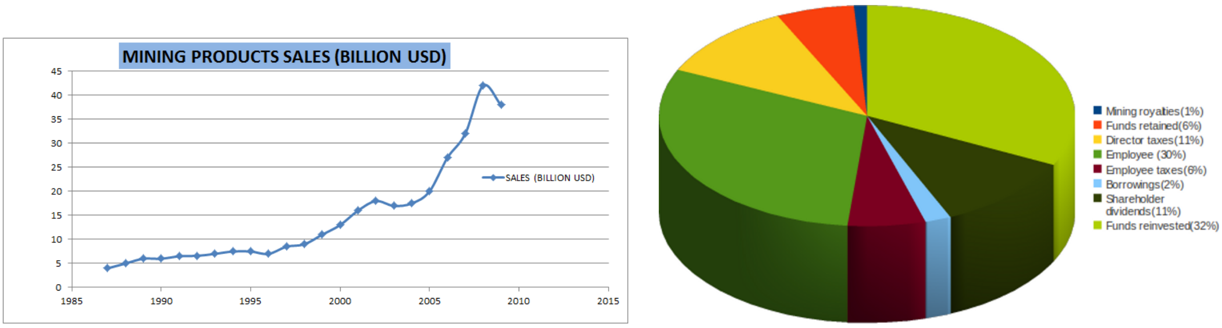
University of Cape To

### 1.1.1 The Mining Context, its Motivation and Challenges

Mining is the activity of extracting minerals from the earth's surface and is ranked as the basic or one of the primary industries of early civilization [18]. The vast amount of resources generated from the mines and the benefits derived from the mining industry cannot be over-emphasised. Figure 1.2 shows the sales trend of mine produce and the distribution of value created by the mining industry [92] and Table 1.1 shows some mine resources required as well as their importance to the human race. There is no doubt that the resources utilised by modern civilization are continually supplied by the mining industry. The resources generated from mines, which have formed part of everyday human life, include the following, to mention only a few:

- Oil and gas
- Gold

- Platinum and coal.




(a) Sales Trend of Mine Produce (B1 Information 2009, Department of Minerals and Energy)

(b) Value Distribution of Mining

Figure 1.2: Sales Trend and Value Distribution of Mining

In 2009, according to the Chamber of Mines of South Africa [83], the industry contributed:

- 8.8% directly, and another 10% indirectly, to the country's GDP.
- Over 50% of merchandise exports, if secondary beneficiated mineral exports are counted
- About 1 million jobs (500 000 directly)
- About 18% of gross investment (10% directly)
- Approximately 30% of capital inflows into the economy via the financial account of the balance of payments
- 93% of the country's electricity-generating capacity
- About 30% of the country's liquid fuel supply
- Between 10% and 20% of direct corporate tax receipts (together worth R10.5 billion).

While these benefits are derived from the mining industry, many risks are associated with the business of the extraction of an ore body. Mining accidents occur in the process of mining minerals or metals. A high risk of accidents in an underground terrain poses a major challenge

University of Cape Town

Table 1.1: Some Mine Mineral Requirements and their Uses

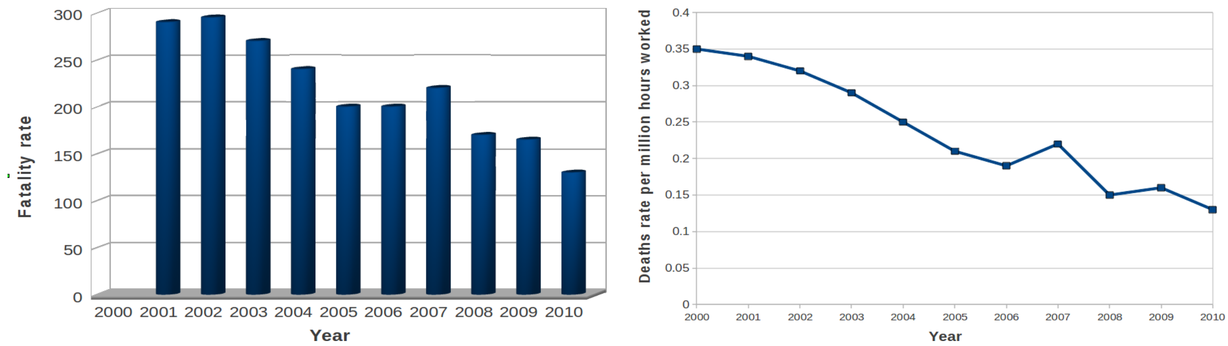
| **Mineral Needs of Humans'** | **Uses** |
|---|---|
| Energy | Heat and power |
| Ornaments and decoration | Cosmetics, jewelry and arts |
| Tools and utensils | Food and shelter |
| Electronics | Computers and communication |
| Currency | Monetary exchange |

to the productivity and profitability of the mining industry. Thus, the twin needs for safety and efficiency in the mining industry have called for the serious attention of researchers and practitioners in recent time.

In general, there are two categories of mining; surface mining and underground (hard rock/sub-surface) mining. Generally speaking, surface mining has less risk associated with it than underground mining, which is considered more hazardous, as it poses a higher risk of accidents and is prone to more disasters owing to its mode of operation [84]. In particular, underground mining deals with excavation, which involves openings for human entry below the earth's surface. It thus leads to the occurrence of fatal accidents and deaths, unlike sub-surface mining. The operations of the miners include blasting, operation of power tools, mechanical and electrical maintenance and similar tasks. These operations are usually carried out in a challenging and an inherently hazardous environment with a high level of noise and dust, extreme temperatures and unstable ground conditions. Accidents in mines can lead to injuries, subsequent ill-health and loss of life. Figure 1.3 shows the fatality rate in the South African mining industry.

In the history of mining, the mortality rate of miners has been alarming; numerous accidents and disasters have been recorded [53]. Many miners have been reported dead and several hundreds have sustained injuries. Over 256 miners were reported dead and more than 64,000 were injured in mining accidents in the decade between 1988 and 1998. Many more lost limbs or eyes, or were otherwise disabled. Miners are sometimes trapped in the underground mine environment [7]. According to a report in 2011 [69], fatalities were still unacceptably high, in line with the trend experienced in the recent past. The occupational health and safety performance of the South

University of Cape Town




(a) Yearly Fatality (b) Deaths Per Million Hours Worked

Figure 1.3: Annual Mine Fatality Rates (SA Chamber of Mines)

African mining sector is lacking in safety when compared to other major mining countries such as the USA, Canada and Australia [53]. Mine stakeholders have directed much effort to mine safety in the past decade as the consensus remains that one death is one too many. The annual death toll from accidents on South African mines has called for the serious attention of researchers and practitioners. Consequently, there is ongoing research on improving safety in South African mines.

The aforementioned statistics are a compelling argument for prioritising safety in mining environments. Safety in the mining industry is non-negotiable. Mining health and safety in South Africa are governed by two important Acts: The Mine Health and Safety Act of 1996 (MHSA) [24] and Occupational Health and Safety Act (OHASA) 85 of 1993 [52]. The "Mine Health and Safety Act 1996" stipulates that employers must ensure safety and maintain a healthy and safe mining environment for workers. While OHASA further stipulates that mine owners must ensure, as far as is reasonably practicable, that nothing about the erection of the mine environment poses a threat to miners' health safety. These Acts commit mine owners and managers to prioritising safety in order to boost productivity and profitability.

It is estimated that the safety performance of the South African mining industry must improve by at least 20% per year to reach, the average performance of Australia, US and Canada [37] by 2013. Underground mines' operation is associated with severe safety problems and the environments of mines tend to degrade fairly rapidly as mining progresses. Thus, the environment requires

University of Cape Town

continuous investigation/monitoring for safe, effective and timely decisions of the management. Miners are susceptible to hazardous situations such as fall of ground, respiratory diseases associated with a high level of dust, ergonomic hazards, falls in shafts, explosives, trucks and trams and unstable ground conditions. Table 1.2 shows the statistics of some fatality agents in mine accidents and deaths [80]. A study carried out on separating people and risk [93] reveals that a fall in shaft/excavation, fall of ground and trucks and tramming are part of the top fatality agents. High fatality rates subject the mine industry to poor productivity. If a mine is unable to become more productive, it will go out of business, causing economic loss to the mining industry and the nation as a whole. Thus, creating a safe working environment for miners is not negotiable and would make mines remain as productive as possible in order to remain economically viable.

Table 1.2: Some Accidents and Deaths on the Transvaal Gold Mines

| Fatality Agent | 1907-08 | | | | 1956 | | | |
|---|---|---|---|---|---|---|---|---|
| | Accidents | | Deaths | | Accidents | | Deaths | |
| | No | % | No | % | No | % | No | % |
| Fall of Ground | 303 | 22.6 | 192 | 25.6 | 4,808 | 21.02 | 271 | 41.31 |
| Falls of Material | 186 | 11.2 | 58 | 7.7 | 3,860 | 16.87 | 44 | 6.71 |
| Falls in Shafts, Excavations, etc. | 113 | 6.4 | 67 | 9.0 | 202 | 0.88 | 42 | 6.40 |
| Trucks and Trams | 151 | 8.3 | 26 | 3.5 | 5,291 | 23.13 | 102 | 15.55 |
| Movements of Cages, Power, Fire, Water | 202 | 12.2 | 123 | 16.4 | 249 | 1.09 | 76 | 11.59 |
| Explosives | 271 | 25.0 | 237 | 31.6 | 206 | 0.9 | 46 | 7.01 |
| Other Machinery | 108 | 6.0 | 24 | 3.2 | 350 | 1.53 | 12 | 1.83 |
| Other | 141 | 8.3 | 22 | 2.9 | 7,910 | 34.58 | 63 | 9.60 |
| | **1,475** | **100.0** | **749** | **100.0** | **22,876** | **100.0** | **656** | **100.0** |

University of Cape Town

### 1.1.2 Possible Remedy

One way of addressing the twin needs for safety and efficient productivity in an underground terrain is to provide miners with knowledge of their working environment via a robotics platform, as demonstrated in Figure 1.4. Knowledge of the safe zone within the mine environment heightens confidence about the occupational environment. With regards to safety in the South African mining industry, research at the Council for Scientific and Industrial Research (CSIR), South Africa [30] continues. It has been noted that subsurface (underground) mining follows a sequential operation/process - drill, blast and clean. After the blasting operation, a safety inspection of the underground rock mass stability is necessary just before the miners remove the ore (clean). At the moment, the safety inspection (dangerous) task is carried out by humans, which is very risky. Robots can be used for the inspection task instead [86]. It is imperative for mine operators to define the risk zone continuously to avoid near misses and accidents in mines and this can be achieved through the use of robots.

Environmental exposure, such as inhalation of poisonous gases, and geographical information, such as cracks in the walls [77], are examples of two important categories of information that can be obtained through the use of robots in mines. The objective is to enhance the security of operations for mine workers and consequently improve mine productivity and profitability. This information would serve as cautionary or danger information for situation awareness and decision-making in the mines, as highlighted in Figure 1.4. The outcome of this work would contribute immensely to safety goals in an underground mine by accelerating path planning and effective traffic awareness for autonomous robots to enable them to carry out safety tasks effectively.

University of Cape Town

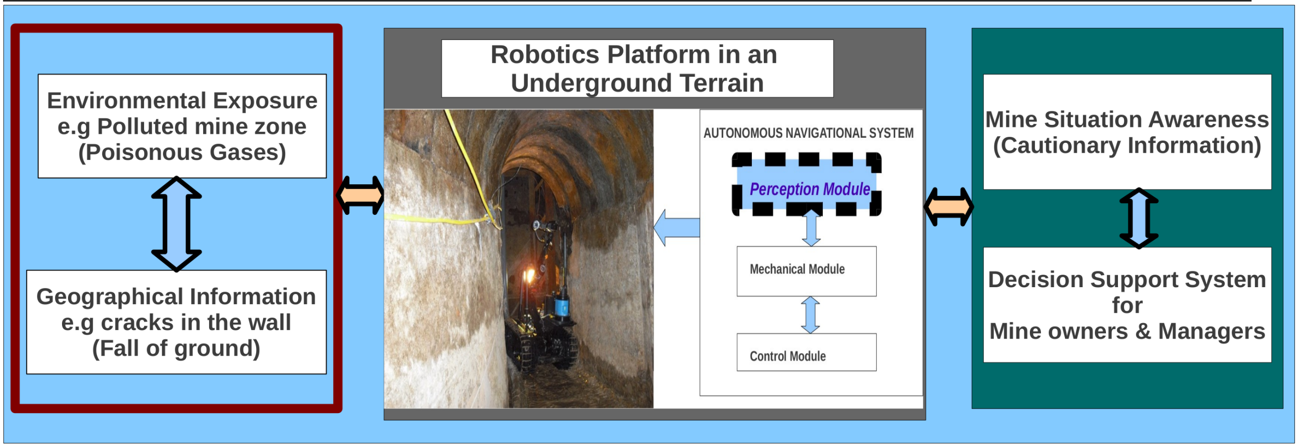



Figure 1.4: Overview of Robotics Potential in Underground Mine Safety

Safety is one of the key factors driving the trend to automation, which has attracted significant attention in recent years [7]. The significant potentials offered by autonomous robots [67] have made them highly relevant in today's world. To provide effective safety measures in mines, robots would play a crucial role. For example, robots can be sent to areas considered dangerous for human operation. However for a robot to traverse a course effectively, its visual capability plays a key role. A significant part of artificial intelligence deals with planning or deliberation for a system that can perform mechanical actions such as moving a robot through some environment. This type of processing (visual sensation) typically needs input data provided by a computer vision system, acting as a vision sensor and providing high-level information about the environment and the robot [10]. How the robot perceives and interprets its immediate environment is very important.

University of Cape Town

### 1.1.3 Robots' Perception

Robots' accurate perceptions (perception module in Figure 1.4) and interpretations of the immediate environment are a critical part of safe autonomous navigation [10]. While, several research efforts have been directed at surface (ground-based) terrains and autonomous outdoor navigation in the 2D plane [20, 25, 39], underground terrains have received little attention and presentation of image classification results in the 3D plane, as shown in Figure 1.5, is relatively scarce.

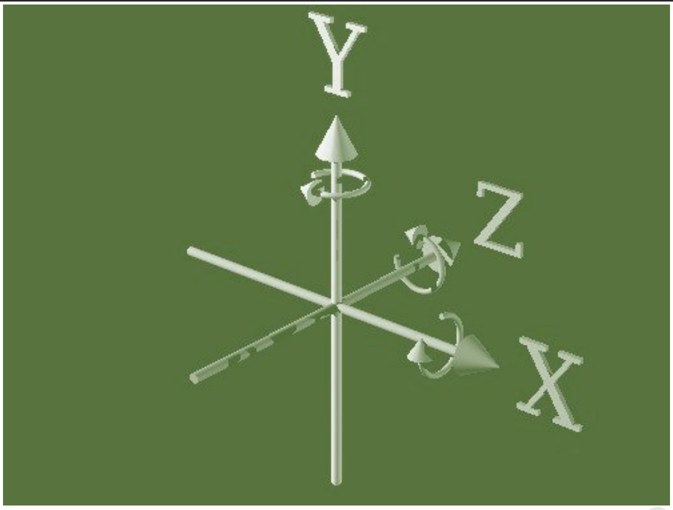



Figure 1.5: An Illustration of a 3 dimensional (3D) Plane

Drivable region detection for robots in mines is a sound basis for navigation or path planning and has received little attention in research, probably because of its roughness. Figure 1.4 depicts the idea of an autonomous robot's navigation and safety potentials in a mine. This research focuses on the perception module (which comprises sensors with high-quality visual capabilities), a critical component in autonomous navigation; the mechanical and control modules fall beyond the scope of this study. The perception module aims to capture observations of the environment (standard and high-resolution imagery), based on the robot's current position $(x, y, z)$, and to specify which region is safe for the robot's navigation. A major focus in this research is the enhancement of robots' capability of identifying drivable regions in underground terrains.

University of Cape Town

## 1.2 Research Rationale and Question

### 1.2.1 Research Rationale

Much research has been directed at autonomous navigation and drivable region detection in different domains of interest. However, autonomous drivable region detection in underground terrains has generally not been explored in depth. Robotics technology shows potential to help address the problems of safety and efficient productivity in an underground environment. However, robots need to execute certain vision tasks if they are to be successful in their operation. Thus, this research is directed at drivable region detection for autonomous robots in an underground terrain. The major contribution of this dissertation focuses on enhancing autonomous robots' capability to identify drivable regions in underground terrains.

### 1.2.2 Research Question

Road-following and autonomous path navigation require the ability to discriminate between the road and surrounding areas and is a well-studied visual task [61]. However, its application to underground mining has only been explored superficially by both the research and practitioner communities, thus research work on this topic continues. Consequently, this study focuses on providing an effective navigational aid for autonomous robots by explicitly devising means through which drivable regions can be identified in an underground environment. In particular, standard image processing techniques, such as feature extraction and image segmentation [73, 91, 90] using entropy-based textural information of image clusters [8], are used in this study. Performance evaluation is also carried out for useful qualitative and quantitative conclusions. Since the key objective of this research is focused on enhancing robots' capability to identify drivable regions in underground environments, the following are the research questions:

University of Cape Town

**Research question one**

*How can a vision model enhance autonomous robots' capability to identify drivable regions in underground environments? That is, how can it produce a visual underground image with clear distinction between drivable and non-drivable regions.*

A number of fairly recent studies have shown promising results in the field of drivable region detection [17, 47, 61]. Most of these studies are domain-specific and depend on the application at hand. The motivation for this work is to enhance the security of operations in underground terrains, that is, exploring possible ways of discriminating between drivable and non-drivable regions in underground terrains for autonomous navigational mission. The use of texture characterisation (entropy model) and image region segmentation (statistical region merging [SRM] algorithm) would be of immense help in achieving this goal. Knowledge of the image textural distribution would accelerate pixel classification on a stream of mine frames and the ability to reconstruct the structural components of an image scene seems feasible.

#### Research question two

*What level of confidence can be expressed on the model in terms of its performance - does the utilisation of the statistical region merging SRM and the entropy models in developing a vision model lead to improved performance?*

This is important in computer vision applications such as drivable region detection for robots, where potentially critical decisions are made by the robots based on the rendered view of the underground image frames. The SRM and entropy algorithms are considered in this research because their statistical/mathematical features show promise. The visual comparison of various 2D experiments is presented for useful qualitative conclusions. Quantitative evaluation of the results obtained also gives useful quantitative conclusions on the models developed.

University of Cape Town

#### Research question three

*How can 3D results be presented to augment 2D drivability results in the proposed model?*

It is common practice in research, especially that on computer vision applications, to present classification results in the 2D $x, y$ plane only. These results pretend that the world is flat and do not give accurate orientation of an image scene. Thus, a 3D approach to image classification results is a better alternative, as it explains the 3D orientation of an image scene. The augmentation of 2D results with 3D drivability maps for autonomous robots is promising, since it reveals the ground truth labelling of the image classification. Integrating the depth maps from the XBOX Kinect 3D sensors into the 2D image classification results in a useful 3D image scene.

### 1.2.3 Contributions and Outline

Most research work is directed at an analysis of the outdoor environment or structured roads [79] where road markings, lanes and kerbs can be used as discriminating factors. In addition, most studies have only considered a 2D drivability analysis, which is not sufficient to define the true structure of an image scene, unlike a 3D drivability analysis.

In this research, publicly available underground mine images are investigated with a stream of locally captured mine images [11][58]. An investigation is also carried out on underground tunnel image frames, using an XBOX Kinect 3D sensors, which produce streams of RGB and depth images. Using the entropy and SRM models on the multi-sensory data from the two Kinect (laser and infrared) sensors, promising results are obtained where drivable regions and non-drivable regions are clearly distinguished. Evaluation is also carried out for useful qualitative and quantitative conclusions, as well as future adoption. The major contributions of this research are as follows:

1. The development of a system model which applies the entropy and SRM methods to drivability analysis of underground terrains using publicly available mine images and a stream of images captured locally in an underground mine.

2. The augmentation of 2D results with 3D drivability maps for autonomous robots, which would need to climb steps in the mines, and benchmarking with publicly available images and methods.

University of Cape Town

To the best of the researcher's knowledge, despite being widely used to detect drivable regions for surface navigation, the entropy and SRM models have never been applied to an underground environment. In addition, not enough research work has presented 3D views of drivability analyses for underground environments [16, 87]; the few known study on 3D presentations only focused on road images.

## 1.3 Dissertation Outline

This dissertation is organised as follows; Chapter one presents the general introduction and motivation for this study. Chapter two presents a brief review of relevant literature in the field of autonomous processes, drivable region detection and image processing. Chapter three presents the methodology and framework in detail. Chapter four presents the 2D experimental results. Chapter five follows with the 3D results and a review of the outcome measures for useful qualitative and quantitative conclusions. Finally, chapter six concludes the research and future work is also presented.

## 1.4 Declaration of Recent Publications

The following recent articles have been published from the research work.

### Refereed Conference Publications

- Falola, O., Osunmakinde, I. and Bagula, A. A Comparative Drivability Analysis for Autonomous Robots in Underground Mines Using the Entropy and SRM Models. *In International Conference Proceedings of the South African Institute for Computer Scientists and Information Technologists (SAICSIT)-ACM 2012.*, ISBN: 978-1-4503-1308-7, pp 31-40, Pretoria, South Africa.

- Falola, O., Osunmakinde, I. and Bagula A. Supporting Drivable Region Detection by Minimising Salient Pixels Generated Through Robot Sensors. In *Proceedings of the Twenty-First Annual Conference of the Pattern Recognition Association of South Africa (PRASA),* MIAPR, ISBN: 978-0-7992-2470-2, Nov 2010; pp. 87-92., Cape Town, South Africa.

University of Cape Town

## 1.5 Chapter Summary

The discussion in this chapter offers a general overview of the study, namely the needs for safety and productivity in the mining industry. While robots have a great deal of potential to deal with safety issues in mines, an effective vision model is critical to their success. However, addressing robots' vision in underground terrains has received little attention in research. Hence, this research focuses on enhancing robots' vision in underground terrains in order to assist their safety mission. An extension in this work is to detect 3D results augmenting drivable and non-drivable regions in 2D. The general outline of the dissertation is also explicitly stated. The comprehensive discussion of the research outline is presented in the succeeding chapters.

University of Cape Town

# Chapter 2

# RESEARCH BACKGROUND AND LITERATURE REVIEW

University of Cape Town

The problem of improving the vision of robots for autonomous navigation has gained significant attention over the years. Several approaches have been adopted to address this problem, most of which are domain-specific, paying less attention to underground terrains, such as mines, which present unique and terrain-specific human hazards. This chapter reviews related work in this domain of interest, to gain an understanding of how previous findings are used in drivable region detection while emphasising the lack of attention to these processes in underground terrains, as well as presenting results in 3D. Different approaches, from the literature, to safety in underground mines are discussed in this chapter. The mathematical background to the entropy and statistical region merging (SRM) models is also documented in order to demonstrate the underlying principles of these models as employed in current research.

## 2.1 Background and Significance of the Study

Mining booms have created significant wealth for various countries, including South Africa [69]. As noted earlier in Figure 1.1, even though the South African mining industry is a key player in the global industry, the industry is continually faced with the twin needs for safety and productivity.

The inherent risks/hazards in the mining industry lead to injury, loss of lives, subsequent ill-health, body deformation and extensive human destruction [52]. Consequently, much effort is currently being directed at improving safety in mines .

### 2.1.1 Relevance of Robotics in Underground Terrains

The potential exhibited by autonomous robots as demonstrated in the literature makes them an appropriate candidate to help address the dual needs for safety and efficient productivity constantly faced by the mining industry. Robotics technology can be useful in underground terrains in diverse ways, including timely preventive interventions and targeted environmental operations, as demonstrated in Figure 1.4.

As mentioned in the previous chapter (Figure 1.4), knowledge of the mine working environment, via a robotics platform, would go a long way in assisting mine owners and managers to make timely preventive decisions concerning safety issues in the mines. A practical application is the work of Stacy *et al.* [85] who developed a system that uses microgravity for the detection of subsurface voids in the earth's crust. The idea is to use a robotics platform to enhance geologists' awareness of possible locations of subsurface voids resulting from discontinuities in the density levels of various materials below the earth's surface.

Robots could be used for various pre-disaster (auto inspection)/targeted environmental operations, such as safety inspection tasks [77], for example, investigating toxic regions in the mines and alerting the mine stakeholders to the level of danger the miners might be subjected to. Inhalation of toxic gases could harm the health of the miners and even result in respiratory or lung diseases. It has been noted that horizontal surface deformation may occur in an underground environment owing to mining activity [40]. However, vertical surface deformation may also occur. Thus robots could be used for the inspection of cracks in the wall, which might have resulted from the seismic activities/factors caused by the previous mining activity carried out [30].

In a recent study, Teleka *et al.* [86] investigated the "Automation of the 'Making Safe' Process in South African Hard-Rock Underground Mines". The work was motivated by the fact that

University of Cape Town

there is a need for a safety inspection of the stability of the underground rock mass, after the blasting operation, just before the ore is removed and cleaned. The dangerous safety inspection task currently executed by humans can otherwise be carried out via a robotics platform. This ultimately provides useful precautionary statements to the miners before they start their operation and protects them against the dangerous underground mine environment. Robots could also be used for post-disaster rescue operations, for example, when miners are trapped in the mines [84]. The goal is to reduce miners' exposure to the dangerous underground mine environment and consequently improve productivity. However, robots need to robustly execute its vision tasks in order to carry out their operations effectively. Hence, an effective vision model is critical to the success of robots in mines.

### 2.1.2 Robots' Vision - The Perception Module

The human visual system carries out recognition on image objects with ease. Activities such as sensing the environment, behavioral sequences and self-propulsion, as shown in Figure 2.1, are carried out effortlessly by humans. However, designing computer vision algorithms to imitate this is a challenging task [59]. Road-following and autonomous path navigation requires the ability to discriminate between the road and surrounding areas and is a well-studied visual task [60, 65]. However, its application to underground mining has been addressed only poorly by both the research and practitioner communities [10], thus research on the topic continues.

A robust navigational model consists of a variety of modules such as the perception (vision), navigation (mechanical) and control (behavioural) modules. The significance of these modules can be seen in emerging real world applications, i.e. The Defence Advanced Research Project Agency challenges [12, 10], ranging from drivable region detection for robots to obstacle detection and avoidance in different terrains or environments. These modules are integrated to form a larger system for completely autonomous path planning and navigation. The basic architecture of an autonomous system is shown in Figure 2.1 which is similar to the trend of the navigational schematic described by Miranda *et al.* [61] in Figure 2.2. Thus, it becomes clear that the perception module (which comprises sensors with high-quality visual capabilities), is a critical

University of Cape Town

component in autonomous navigation. While the mechanical and control modules does not follow the scope of this study.

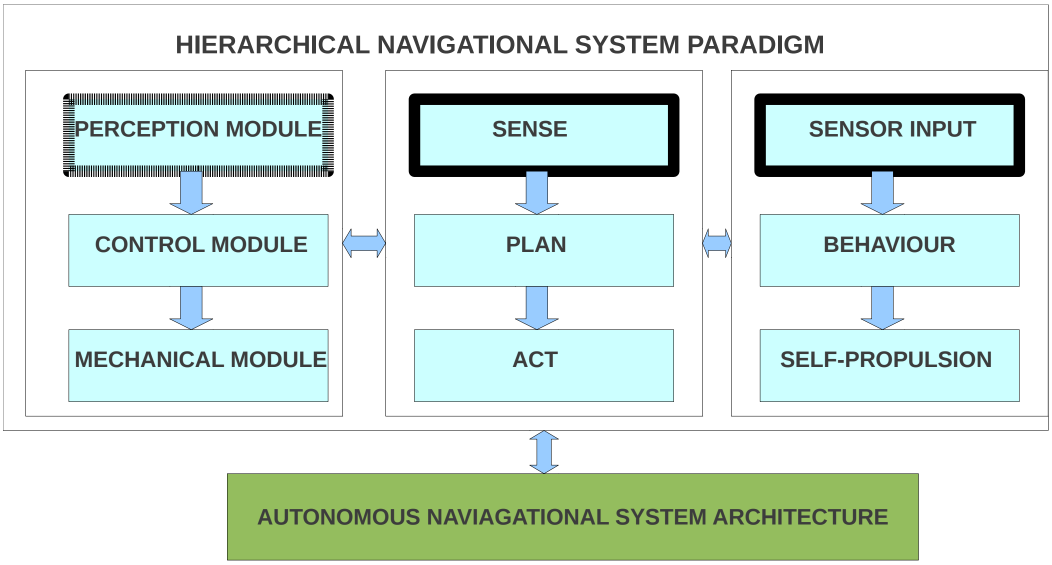



Figure 2.1: Schematic of the Architectural Components of an Autonomous Navigational System

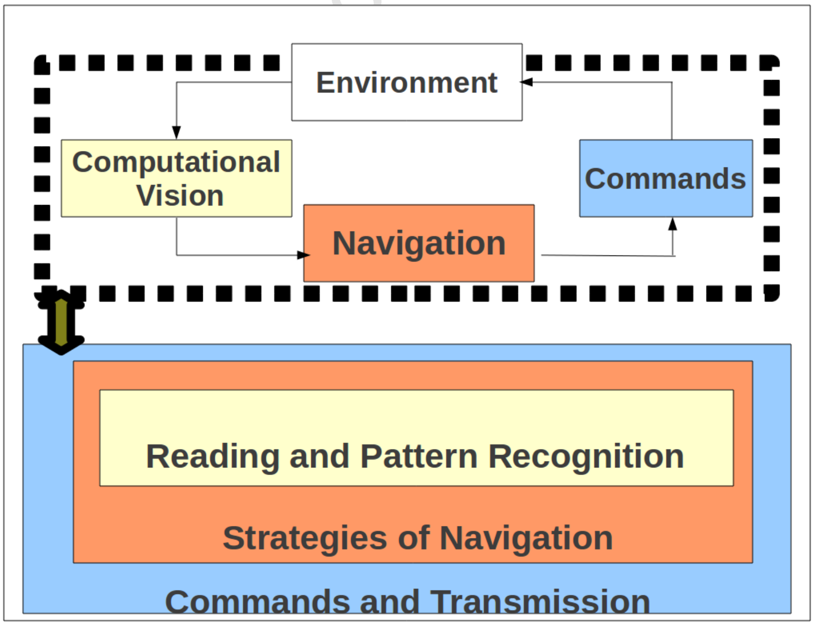



Figure 2.2: Components of a Navigation System for Autonomous Robots

## 2.2 Review of Related Literature

The advancement in robotics and automation has led to a significant increase in autonomous processes. This has in turn prompted the research community to develop new methods for autonomous navigation and even improve on existing ones [59] [60] [65]. In the recent past, there has been some development in the field of image processing and segmentation, region classification and object detection for autonomous processes [17][22][79]. The state-of-the-art in image processing and computer vision has shown that robotics technology would be of immense benefit in tackling underground mine safety. However, its application to underground environments has received little attention.

Greve *et al.* [31] report a new approach to the classification of image regions. They use a wavelet standard deviation descriptor for texture patterns. The idea is to find a label for each region in any given image that can later be used in image management applications or retrieval systems, e.g. to tag an inserted photo image automatically. The ultimate aim is to reduce the problem of image region classification to a problem of texture pattern classification. Their work is directed at surface (ground-based) and outdoor image classification.

A lane detection algorithm for autonomous driving is presented by Young *et al.* [90]. Their algorithm is based on three features, termed three-feature based automatic lane detection, as demonstrated in Figure 2.3. Then they used two horizontal windows for automatic lane detection by identifying left and right boundaries in the image based on their defined region of interest. The last step involved lane inference based on the $L$ possible lane candidates observed in the previous processing. Figure 2.3 presents the overview of their algorithm. They also used the inverse perspective transform operation in converting the image coordinate plane with the perspective effect into a plane-view image geometry. The three basic features considered in their work employed Sobel feature extraction, which is used for image enhancement after image preprocessing. Several types of filters for features extraction exist, such as the Prewitts, Roberts, Sobel and Canny [76][81]. However, the choice of filter depends on the application at hand.

Joaquin *et al.* [47] propose an approach to a visual-based sensory system for an autonomous navigation through orange groves. They used colour camera with auto iris and VGA resolution

University of Cape Town

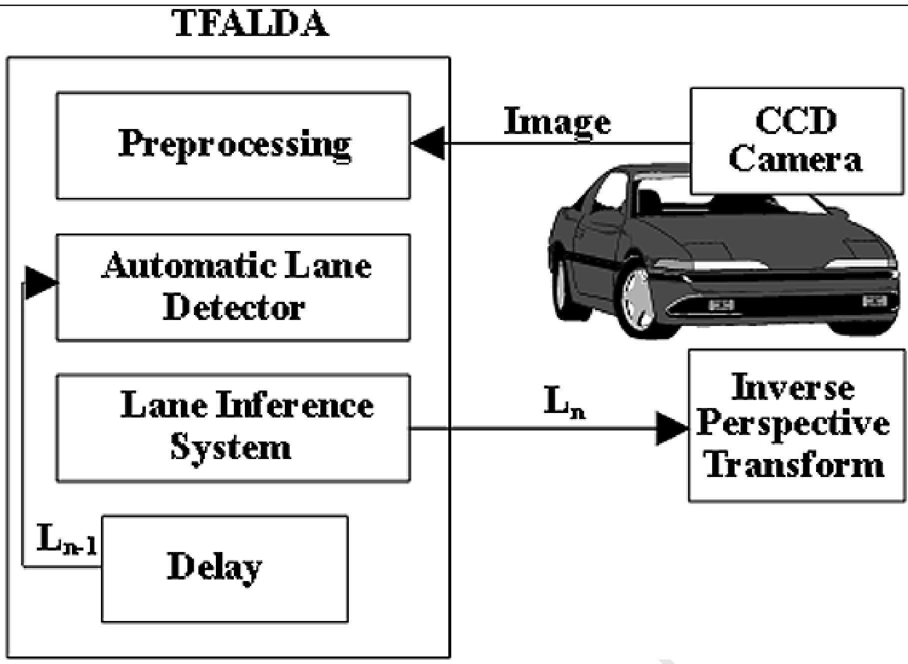



Figure 2.3: Block Diagram of the Three-Feature Based Algorithm

for the image capture and a neural network (multilayer feedforward network) to classify the ensembles together with hough transform. The aim of their work is to establish the desired path for autonomous robot within an orange grove. For example, in agricultural robotics where autonomous robots are used for weed detection or spraying fungicides. The findings from their research show promising results.

Derek *et al.* [22] address the issue of recovering surface layout from an image. Their work presents a partial solution to the spatial understanding of the image scene (environment), which aims to transform a collection of an image into a visually meaningful partition of regions and objects. Using statistical learning based on multiple segmentation framework, they constructed a structural 3D scene orientation of each image region. They went further to conduct experiments on indoor scenes which correspond to underground tunnel images in this research.

Shengyan *et al.* [95] put much effort into road detection using a support vector machine (SVM) based on online learning and evaluation. The focus in their work is on the problem of feature extraction and classification for front-view road detection. According to Jian *et al.* [46], the SVM is defined as a technique motivated by the statistical learning theory, which has shown its

University of Cape Town

ability to generalise well in high-dimensional space. SVM attempts to separate two classes by constructing an N-dimensional decision hyper-plane that optimally maximises the data margin using the training sample. In the problem of road detection, the SVM classifier is used to classify each image's pixel into road and non-road classes based on the computed features.

A close application to underground terrains is the work of Bakambu *et al.* [6]. They presented an autonomous platform for navigation and surveying within networks of tunnels, like those typically found in underground mines and caves. The system works in two alternative modes: surveying mode or navigation mode. In the surveying mode, the researchers gathered range data for map building using a remotely located supervisor who instructs the platform to move through successive sections of the network. While in navigation mode, the supervisor specifies high-level missions using the previously acquired survey maps. A motion planner then translates each mission into a set of consecutive navigation actions separated by natural landmarks. Mission execution consists of autonomously detecting landmarks, self-localising, and performing the planned navigation actions. The path-following controller uses a kinematical model of the motion vehicle, which depends on position $(x, y)$ and orientation $\theta$, while wall-following was ensured by following the middle axis of the drift.

Valarmathi and Aruna [91] proposed a novel approach to extract image features such as contour extraction and edge detection for image segmentation with self-organising properties for a network of adaptive elements. They used a type of neural network called Kohonen's self-organising maps. The extra spatial information about a pixel region obtained by using the unsupervised training algorithm verifies that the neighbouring pixels should have a similar segmentation assignment unless they are on the boundary of two distinct regions.

Over the years, researchers have made tremendous progress in the field of computer vision, including autonomous navigation and drivable region detection [17, 20, 25, 39, 47]. However, underground terrains have not been explored in depth and 3D analysis of image frames is yet to be fully explored and optimised [16], as most research work only presents 2D results. It is evident that a 2D approach to robots' vision does not adequately address the challenge of geometrical image structure in the world model, since it does not fully explore the 3D scene relationship. Most early studies on drivable region detection focused on structured road images and very little

University of Cape Town

attention was paid to unstructured roads and not much significant attention to underground terrains (mines/tunnels).

A vision based approach to detecting drivable space in traffic scene understanding is presented by Chih-Ming *et al.* [41]. Their work focuses on road detection for intelligent vehicles. An important step in their detection approach was to generate a feature similarity image (mean, standard-deviation and entropy) from the road images by using a statistical feature analysis process. Using a Breadth-First Search (BFS) model in the feature similarity image, road regions are detected. Inverse perspective transform is then finally used to detect drivable regions within the tested road images. Figure 2.4 shows the results from their research.

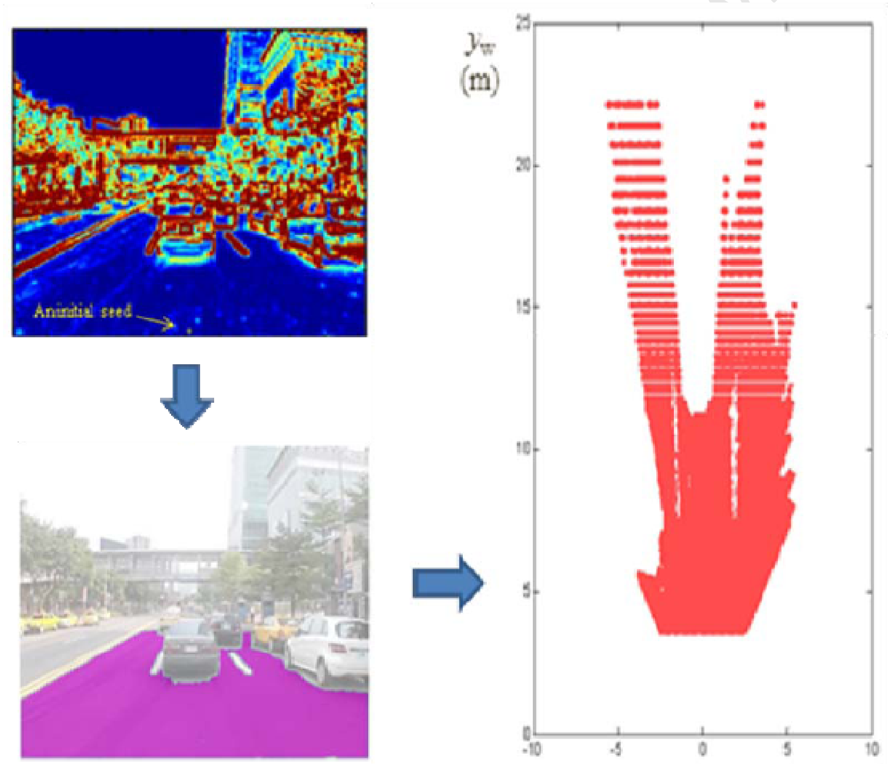
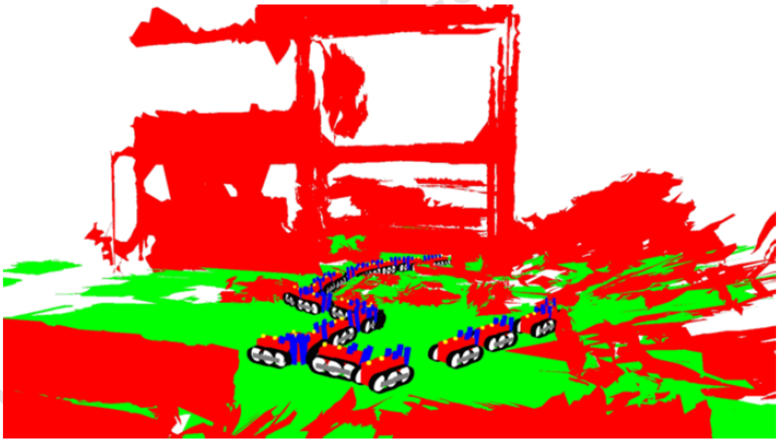



Figure 2.4: Studies Result on Road Detection in World Coordinates System

Neuhaus *et al.* [63] investigated terrain drivability analysis in the 3D laser range data for autonomous robot navigation in unstructured environments. They used grid-based principal component analysis (PCA) and hierarchical PCA algorithms for classifying regions as either drivable or not and conducted a further examination using a novel algorithm, which determines the local terrain roughness. While their study is directed at unstructured environments, it only applies to drivability analysis in surface unstructured terrains and not underground terrains. Similarly, Narunas *et al.* [62] presented a work in progress on path-planning for unmanned ground vehicles. 3D-plane simultaneous localisation and mapping (SLAM) is employed in their path-planning processing. They presented an impressive result (see Figure 2.5), comprising 3D maps integrated over time on long-range navigation.

(a) Unstructured image of collapsed car parking lot



(b) 3D visualisation with green region as drivable and red region as non-drivable

Figure 2.5: Detection Results on Path Planning for Autonomous Robots

From the literature, it is observed that most of the early studies on autonomous drivability analysis focus on outdoor navigation and ground-based detection. Underground terrains have received less attention, probably because of their roughness, compared to structured and unstructured ground-based terrains. While autonomous navigation in an underground mine environment has been studied for more than 12 years [75], a robust algorithm that is applicable in different terrains is yet to emerge. Thus, it remains an ongoing key challenge.

## 2.3 Automation and Safety in Underground Environments

In the service of promoting and enhancing autonomous processes in real time, several applications have emerged. Drivable region detection and path planning for autonomous robots have been widely addressed by researchers using different approaches [62]. There are several motivations for studying and developing techniques or applications for drivable region detection, which depend on the nature of the application at hand. However, the application of these techniques to drivable region detection in underground terrains has been poorly addressed. While robots have a great deal of potential to tackle safety issues in mines, underground environments pose some constraints on robotics technology.

Automation in an underground environment is a current and challenging task. This is due to the fact that underground environments are inherently unstructured and subject to constraints such as limited wireless information and unstable ground conditions. Bandyopadhyay *et al.* [7] investigated a "Wireless Information and Safety System for Underground Mines" and noted that the application of wireless communication systems are highly restricted in underground mines by the high attenuation of radio frequency waves in mines. These limitations were also corroborated in a recent study by Green *et al.* [30], who investigated sensing technologies that could advance the development of autonomous vehicles in an underground mining environment by exploring a combination of 3D cameras (SR 4000 and XBOX Kinect) and a thermal imaging sensor (FLIR A300) in order to create 3D thermal models of narrow mining stopes. The objective of their work was to identify the technical risk areas that are barriers to the implementation of underground robots. Their aim was to provide information that can be used in determining the risk of rock fall in an underground mine, which is a major cause of fatalities in underground narrow-reef mining.

The state-of-the-art technology in mine automation includes a discussion on the various research projects and activities that have been carried out on the subject of automation of excavating

University of Cape Town

machinery [36]. They stressed that the main tasks involved in autonomous loading are excavation, navigation, obstacle detection and obstacle avoidance. Nonetheless, most practical operations by robots in the mine have only been teleoperated [32] by a human and drivable region detection for robots in the mine has scarcely been addressed.

There has been some developments in robotics technology over the years, with variations in application and phases of automation. The first-generation applications involve modification of conventional devices such as load-haul dump (LHD) vehicles. The second phase of automation involves the creation of dedicated teleoperated robots, while the third generation paradigm shift deals with autonomous robots in an exclusively robotic environment. Robotic technology has been extended to automated navigation of an LHD unit of underground mines. A roof-mounted reflected line and a camera were used to guide the vehicle along the mine passages. Another initiative is the work of Clark *et al.* [19] that presented the development of an underwater robot system capable of mapping out and navigating underwater tunnel systems. They used SLAM techniques on a small underwater robot to explore and map ancient cisterns located on islands, which act as water storage systems for private homes and churches.

Safety in underground mines has been addressed from different perspectives. A variation is presented by Aminossadati *et al.* in [1], where computational fluid dynamics modeling is used to simulate the airflow behaviour in underground crosscut regions, where brattice sails are used to direct the airflow into these regions. Their aim is to assist in understanding the behaviour of ventilation air and in determining the optimum size of brattice curtain (sails), which provide a highly effective contaminant removal from the unventilated mine areas. They suggested that their work would help the mine ventilation designers meet mine safety requirements. With improved mine safety and productivity in mind [29], an innovative alternative to manual procedures is described for the application of carbon fibre and resin injection in concrete surfaces in tunnels. Vision and laser telemeter sensing are integrated into the tool to assure precise inspection and maintenance operations. Lavigne *et al.* [51] investigated mapping GPS-deprived underground mining environments with the eventual goal of using these maps for navigation. The major goal of the work was to overcome the requirement for human input and to make the map-building process autonomous. This is also addressed in [38] by Hlophe. On the other hand, Teleka *et*

University of Cape Town

*al.* [86] seek to address mine safety issues via a robotic platform. Their aim is to automate the dangerous safety inspection task required after the blasting operation in mines. This prevents humans from being exposed to the dangerous underground environment, as discussed earlier.

3D image analysis has not received much attention compared to 2D image analysis. The 2D $x, y$ image plane assumes that the world is flat and does not give accurate information about the orientation of an image scene, unlike 3D. One of the few studies on 3D imaging is the work of Andreasson *et al.* [2]. They focus on methods to derive a high-resolution depth image from a low-resolution 3D range sensor and a colour image. They use colour similarity as an indication of depth similarity, based on the observation that depth discontinuities in the scene often correspond to colour or brightness changes in the camera image. This work hinges on the work of Thrun *et al.* [88], which deals with acquiring accurate and very dense 3D models in excavation sites and mapping of underground mines using laser range finders. Figure 2.6 presents results obtained in their work.

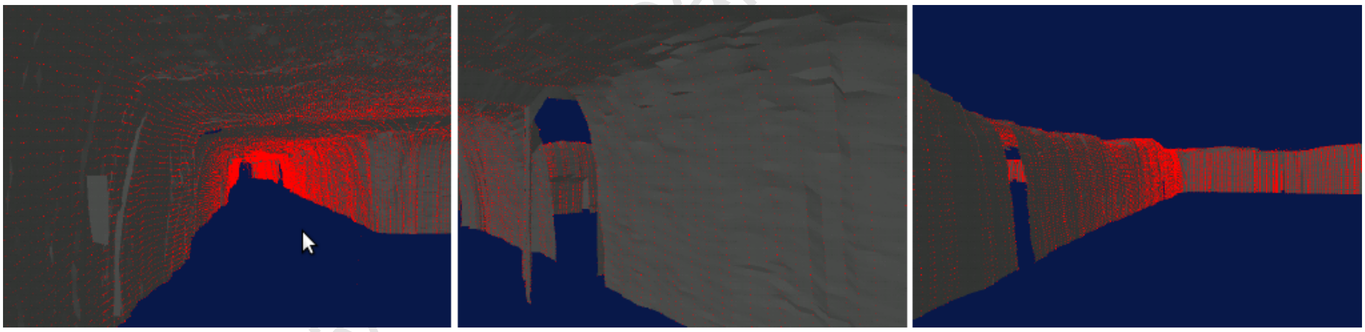


Figure 2.6: Sequence of 3-D Visualisations of a Volumetric Mine Map

This research focuses on enhancing robots' visual capability in underground terrains by exploring drivability analysis of the underground terrain. The entropy and SRM models are used for extracting features (colour and texture) because they show promise in the feature/statistical property, thereafter image region segmentation and classification are carried out. The goal is to accelerate autonomous mine safety inspection tasks and consequently improve mine productivity.

University of Cape Town

## 2.4 Image Data Acquisition

Over the years, whenever terrain drivability or negotiability information was needed, image or laser-based strategies were widely employed. Several papers have been published on the deployment of a laser range finder sensor [2, 63, 79] and other sophisticated cameras/devices, ranging from mono to stereo cameras [87]. Figure 2.7 presents hardware sensor devices deployed by Haselich *et al.* [35] for environmental perception. Most of these sensors are relatively expensive and have precision limitations of some sort. Pingbo *et al.* [68] noted that while much research has addressed laser scanner noise models, comparatively little is known about other artifacts, such as the mixed pixel effect (usually caused by the presence of noise), colour-dependent range biases, and specular reflection effects. They investigated the mixed pixel effect and the related challenge of detecting depth discontinuities in 3D data by presenting a comparative analysis of five discontinuity detection algorithms on real data sets. The mixed pixel effect, which results in erroneous range measurements, is caused by the presence of noise (unwanted signal) usually through the integration of light from multiple sources at different ranges by a single pixel [28]. Their findings indicate that no algorithm performs exceptionally well. However, low-cost range sensors are an attractive alternative to expensive laser scanners in application areas such as indoor mapping, surveillance, robotics and forensics [50].

University of Cape Town


Figure 2.7: Deployed sensor for perception: A 3D laser Velodyne HDL-64E S2 (left) and two different commercially available cameras (Logitech HD Pro Webcam C910 on the upper right and Philips SPC1300NC on the lower right)

## 2.5 Image Analysis

After a device has been used to capture an observation (image scene) of an environment, the image must be analysed in order to draw useful conclusions from it. An image is an array of pixels arranged in columns and rows. A pixel (picture element) is the smallest discrete component, usually stored as a collection of bits (binary digits) [72]. At each pixel's location, a colour value is stored. In a (8-bit) greyscale image, which is normally referred to as black and white, each picture element has an assigned intensity from the set $\{0, 1, \cdots, N-1\}$ where $N = 256$. If it is a coloured image, say an RGB image, sets of 24 bit (tri-stimulus) values, corresponding to the different channels of the colour intensity component are stored.

Recognising real world objects in a complex and dynamic scene is the purpose of general image understanding. Generally speaking, in image analysis, whenever information about an image structure/content is needed, image segmentation is the first option that readily comes to mind [54]. Image segmentation is the partitioning of an image into regions of interest or their constituents parts, based on some property, in order to gain better understanding of the image structure. Whenever image segmentation is needed, certain approaches or methods are necessarily required.

University of Cape Town

Two such approaches are discussed below:

- Edge detection [73]
- Region growing or boundary estimation [64].

Edge detection involves the identification of areas of sharp or abrupt discontinuities in image features [81] and many methods for edge detection exist. Typical examples are the Sobel, Canny, Prewitts and Roberts detectors [76], but these are not the only ones available. Each of these detectors has unique operators, usually a smaller pixel neighbourhood window, associated with it. On the other hand, region growing involves the detection and/or merging of regions sharing similar statistical feature properties in an image [13]. However, the nature of the application at hand and the complexity of the system considered would determine what approach to segmentation to adopt. These methods are generally considered for features extraction for image retrieval.

Several types of segmentation techniques have been developed. According to Kartikeyan *et al.* [49], segmentation methods can be classified into two classes: local behaviour-based (LB) and global behaviour-based (GB). Typical examples of LB methods are region extraction and edge detection methods. The LB methods analyse the variation of spectral features in a small neighbourhood, while GB methods group the pixels based on the data in the computed feature space [54]. Histogram threshold and clustering are two examples of GB. A notable segmentation technique is the Otsu method [66]. The Otsu segmentation technique has been widely adopted in the field of image processing and computer vision [61]. The Otsu approach is to carry out image thresholding on histogram structure-based images using the intensity function. The threshold that minimises the weighted within class variance and maximises the between class variance is regarded as the optimal threshold.

### 2.5.1 Entropy Model

Image analysis can be conducted with a variety of models, such as the entropy model, which uses a different technique from the edge detection based method. Edge detection techniques aim

University of Cape Town

at extracting edge features, while the entropy of an image is usually used for the extraction of texture features. The concept of entropy dates back to information theory, where it is usually referred to as Shannon's entropy [78]. Entropy in general is defined as the number of binary symbols needed to code a given input given the probability of that input appearing on a stream. More information on the subject of entropy can be found in the work of Jaynes [45], and the reviews carried out by Sebastiano [8].

The entropy of an image is a statistical measure of randomness that can be used to characterise the texture of the input image. Texture is a measure of local spatial variation in the intensity of an image. Entropy is an important concept in image segmentation [74]. One way to apply the entropy concept to image segmentation is to calculate the graylevel transition probability distributions of the co-occurrence matrices for an image and a thresholded bilevel image, respectively, then find a threshold which minimises the discrepancy between these two transition probability distributions, i.e. their relative entropy. The threshold rendering the smallest relative entropy will be selected to segment the image. The entropy of a square neighbourhood around each image's pixel is an important feature in the study of Chih-Ming *et al.* [41]. A major step in their road detection approach is to extract spatial/textural features and generate a feature similarity image of the road images.

The entropy of the mine images is computed using Equation (2.1) such that every pixel in the entropy filtered image (EFI) is measured within a fixed window (9 $\times$ 9 window in this case), which accounts for a reasonable percentage of the textural distribution of each pixel region [8]. This window size is considered approriate after testing with the 3 $\times$ 3 and 16 $\times$ 16 which seem not to give a satisfactory result. High entropy indicates a high variance in the pixel values, while low entropy is associated with fairly uniform pixel values.

For each pixel's neighbourhood window, the entropy is evaluated using Equation (2.1).

$$k_i = \sum_{v} -q_v \times \log_2(q_v) \tag{2.1}$$

where $q_i$ represents the probability that a random pixel $p$ chosen from the window centred at $p_v$ will have intensity $i$. The computation is done using the non-zero values of the histogram samples

University of Cape Town

probability, say $q_v$, for every point, $h$, in the sample histogram, as shown in Equation (2.2).

$$\textit{samples-probability } (q_v) = \frac{h}{\textit{length of histogram}} \tag{2.2}$$

The entropy filter measures the relative change of entropy in a defined or sequential order [13]. For each pixel $p(i,j)$ in the EFI, corresponding pixels $p_1, p_2, ..., p_N$ exist for each mine image. The local entropy $k_i$ measured within a fixed window, for each pixel $p_i$ in each image, is computed and the weighted average $p$ is computed as shown in Equation (2.3).

$$p = \frac{\sum_{i=1}^{N} p_i k_i}{\sum_{i=1}^{N} k_i} \tag{2.3}$$

John *et al.*[48] presented the maximum entropy method as the generally acceptable procedure for assigning prior probabilities, since it gives acceptable results in simple cases. It also applies to probability distributions, and generally to any additive or positive distribution such as an image, giving a direct justification for the use of entropy in imaging. Their work focuses on Bayesian interpretation of maximum entropy image reconstruction.

In this research, the entropy approach is used to estimate the probabilistic textural information of image frames for further processing. Since entropy is a measure of randomness, it provides a way to compare different regions (drivable and non-drivable regions) of the mine frames for further processing.

University of Cape Town

### 2.5.2 SRM Algorithm

On the other hand, the SRM algorithm is a method of image analysis that uses the region-growing approach to segment an image [54]. A region in an observed image $I$ is a group of connected pixels with some homogeneity in feature property. Image segmentation refers to the process of partitioning a digital image into multiple regions (sets of pixels). Segmentation is a collection of methods allowing one to interpret parts of the image as objects by transforming the pixels into visually meaningful partition of regions and object [73]. The object is all that is of interest in the image and the rest of the image is considered as the background. For an image $I$ and homogeneity predicate $H_p$, the segmentation of an observed image $I$ is a partition $K$ of $I$ into a set of $G$ regions, $R_1, R_2, \dots R_G$, such that the following conditions hold [9]:

a. $H_p(R_g) = true\, \forall g$

b. $H_p(R_g \cup R_h) = false\, \forall\, adjacent(R_g, R_h)$

c. $\bigcup_{g=1}^{G} R_g = I\ with\ g \neq h\ and\ R_g \cap R_h = \emptyset$

SRM models segmentation as an inference problem by performing a statistical test based on a merging predicate and has been widely used in medical imaging and remote sensing imagery [13, 15, 64, 54]. Richard *et al.* [64] present an elaborate theoretical analysis of the SRM algorithm in order to analyse the underlying principles. SRM is applied to skin imaging technology in [15] so as to detect borders in a dermoscopy image in an attempt to analyse a skin cancer (melanoma).

In region merging, regions are iteratively grown by combining smaller pixels. SRM uses a union-find data structure or merge-find set, which is defined as follows:

- Find: Determines if two elements (pixels) are in the same subset.
- Union: Merges two subsets (sub-region) into a single subset (region) based on some criteria.

A major limitation of SRM is overmerging, where an observed region may contain more than one true region. It has been shown that the overmerging error is more or less insignificant, as the algorithm manages accuracy in segmentation close to optimum [64]. The idea is to reconstruct the statistical (true-similar) regions of an observed image instance.

The algorithm relies on the interaction between a merging predicate and the estimated cluster, $Q$, specified. The merging predicate, $P(R, R')$, on two candidate regions, $R, R'$, is depicted in Equation (2.4) with an extension in Equations (2.5) and (2.6).

$$P(R, R') = \begin{cases} \text{true} \quad \textit{if} \quad \forall c \in (\text{R,G,B}), |\bar{R}'_c - \bar{R}_c| \leq T \\ \text{false otherwise} \end{cases} \tag{2.4}$$

University of Cape Town

$$T = \left| \sqrt{k^2(R) + k^2(R')} \right| \tag{2.5}$$

$$k(R) = g\sqrt{\frac{1}{2Q\,|R|}\ln(6|I|^2 R_{|R|})} \tag{2.6}$$

$R_c$ is the observed average colour channel $c$ in region $R$ and $R_{|R|}$ represents the set of regions with $R$ pixels.

Let $I$ be an observed image with pixels $|I|$ that each contains three $(R, G, B)$ values belonging to the set $\{0, 1, \cdots, g-1\}$ where $g = 256$. The observed image $I'$ is generated by sampling each statistical pixel for the three $RGB$ channels. A statistical pixel corresponds to the intensity value of the pixel. Every colour level of each pixel of $I'$ takes on value in the set of $Q$ independent random variables with values of $[0, g/Q]$. $Q$ is a parameter that describes the statistical complexity of $I'$, the difficulty of the problem and the generality of the model [54]. The optimal statistical regions in $I'$ satisfy the property of homogeneity and separability.

- Homogeneity property: In any statistical region and given any colour channel, the statistical pixels have the same expectation value.
- Separability property: The expectation of any adjacent statistical region differs in at least one colour channel

Equation (2.7) defines the sort function [64], where $p'_a, p_a$ represent pixel values of a pair of adjacent pixels of the colour channel.

$$f(p, p') = \max_{a \in R,G,B} |p_a - p'_a| \tag{2.7}$$

Other approaches to image analysis can be seen in the work of Angelova *et al.* Different means exist to classify the terrain around an autonomous robot platform. Angelova *et al.* [4] put much effort into fast terrain classification by using a variable-length representation approach to build a learning algorithm that is able to detect different natural terrains. They used a hierarchy of classifiers to classify different natural terrains, such as the sand, soil and mixed terrain. Their ultimate goal was to classify different terrains for autonomous navigation. From another perspective, Angelova *et al.* [5] endeavoured to assist an autonomous vehicle to drive safely on slopes by learning autonomously about different terrains and their slip characteristics using a probabilistic framework. By using automatic mechanical supervision collected from sensors on board an autonomous vehicle, they learned terrain traversability properties from visual input. The ultimate goal was to ensure safe autonomous navigation.

## 2.6 Chapter Summary

This chapter reviews the related literature in the field of autonomous drivable region detection and image processing, evaluates existing knowledge and specifically identifies the gaps that this research is intended to fill. Various approaches/perspectives to addressing safety in underground terrains are also discussed. While autonomous robots exhibit a great deal of potential to tackle safety issues in mines, an effective vision model is critical to their success. However, from the literature, it is observed that most early studies on autonomous drivability analysis focuses on outdoor navigation and surface (ground-based) detection. Underground terrains have received little attention, probably owing to their roughness and environmental/technological constraints, compared to structured and unstructured surface terrains. The different techniques that have been used in drivable region detection, as well as their application areas, have also been discussed. These techniques, such as PCA, SVM, the variable-length representation approach and SLAM, have a number of applications in outdoor and surface terrains and mostly present results in 2D. The 2D results which, are common practice, assume/pretend that the world is flat, which is false, and do not fully describe the 3D relationship of an image scene.

The gap which this research is intended to fill is indicated, together with the theoretical background of the entropy and statistical region merging models as applied to current research. This research focuses on enhancing autonomous robots' capability to identify drivable regions in underground terrains with a view to presenting the results in 3D.

University of Cape Town

# Chapter 3

# RESEARCH RATIONALE AND DESIGN

University of Cape Town

This chapter re-establishes the research rationale and provides the detailed methodology of this research. The entropy and SRM approach to drivability analysis is also documented in order to demonstrate its procedure and applicability to current research.

## 3.1 Research Rationale

This study is motivated by the needs for safety and efficient productivity constantly faced by the mining industry, since mine produce has inevitably become part of the daily human supply. The aforementioned needs can be addressed through the use of robotics technology, since robots exhibit very useful potentials to assist in this domain of interest. The idea is not to replace mine workers by robots, but to improve safety in their working environment. One way of implementing this is by using robotics technology to provide useful precautionary statements to the miners before they begin their operation. Hence, this work serves to assist autonomous robots to identify drivable regions in the underground environment, since the success of the robot is highly dependent on its visual capability to interpret its immediate environment correctly.

## 3.2 Research Design: Proposed System Model for Drivability Analysis

### 3.2.1 Proposed System Model

This research conducts a drivability analysis of underground terrains for autonomous robots by devising means through which drivable regions can be identified in underground terrains. This consequently improves the vision of a robot, allowing it to navigate in only drivable regions in a mine frame, thereby minimising accidents while executing its tasks. Figure 3.1 depicts the proposed system model, which is described in detail in the next sections.

The model presents the stages, ranging from data capture to classification results, required to tackle the segmentation problem addressed in this research. In the model, after the data capture is done, the proposed algorithms (entropy and SRM) are applied to the images, in order to detect the drivable region. Furthermore, performance evaluation in terms of the accuracy of detection is carried out. The qualitative and quantitative performance is checked for accuracy as presented in Figure 3.1.

Owing to the fact that gaining access to underground mines is somewhat difficult, an XBOX Kinect device has not been used to capture depth information of mine images. Hence, there are no 3D visualisations for the mine images presented in the model. However, the segmentation problem is still tackled on mine frames/images since it is one of the major focus in this research.

University of Cape Town

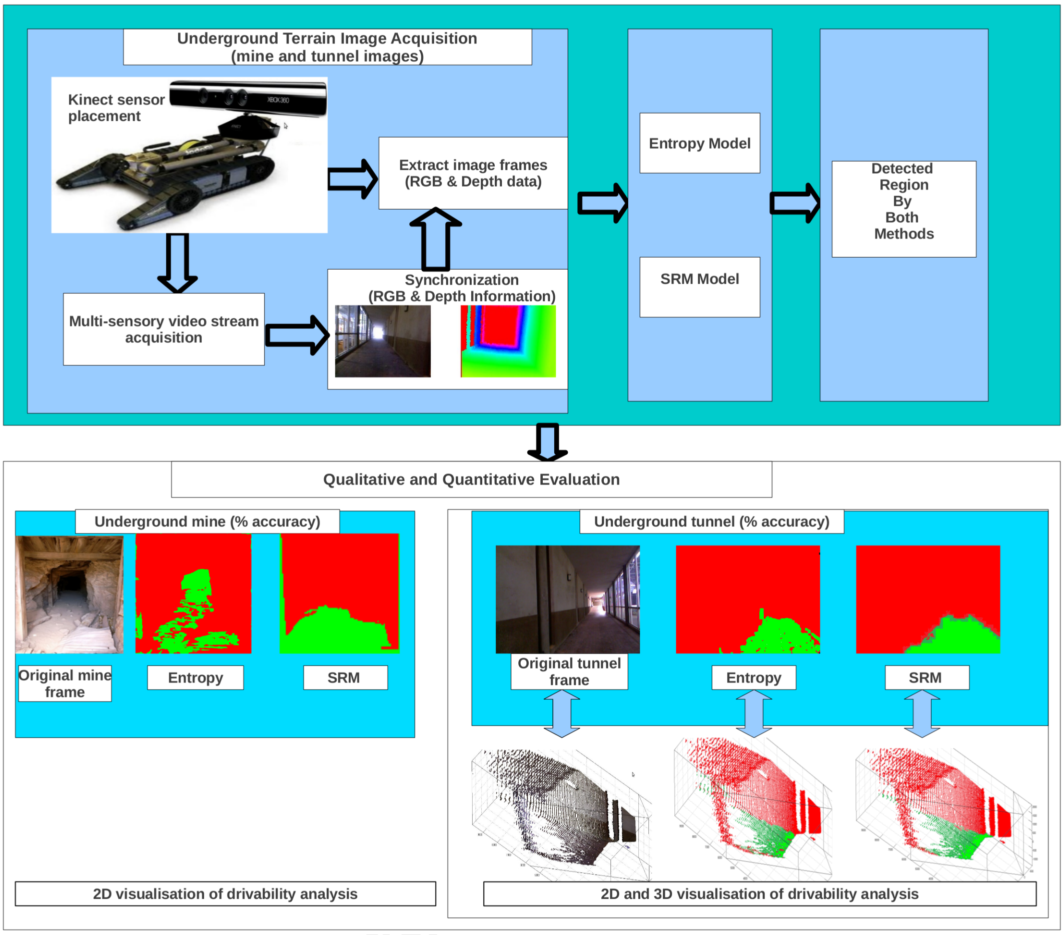



Figure 3.1: Schematic of the System Model for Drivable Region Detection (Perception Module)

University

### 3.2.2 Underground Terrain Image Acquisition

In this work, the tested mine frames combine a selection of publicly available frames with a stream of mine frames captured in a local mine using common photo cameras. Different regions of mines, such as the shaft, stope and gallery, are investigated. An XBOX Kinect 3D sensor, a low cost [30], high quality multi-sensory image device, is used to capture underground tunnel image frames. It is important to emphasize that the author has not used an XBOX Kinect sensor in an underground mine environment. However, the device is used in an underground tunnel like environment in order to explore 3D visualisation of the classification results as presented in Figure 3.1. Details of the 3D Kinect sensor device and the capture process are presented in Chapter 5.

### 3.2.3 Entropy Approach to Drivability Analysis

Figure 3.2 describes the interlinked streams of the entropy approach used in this research.

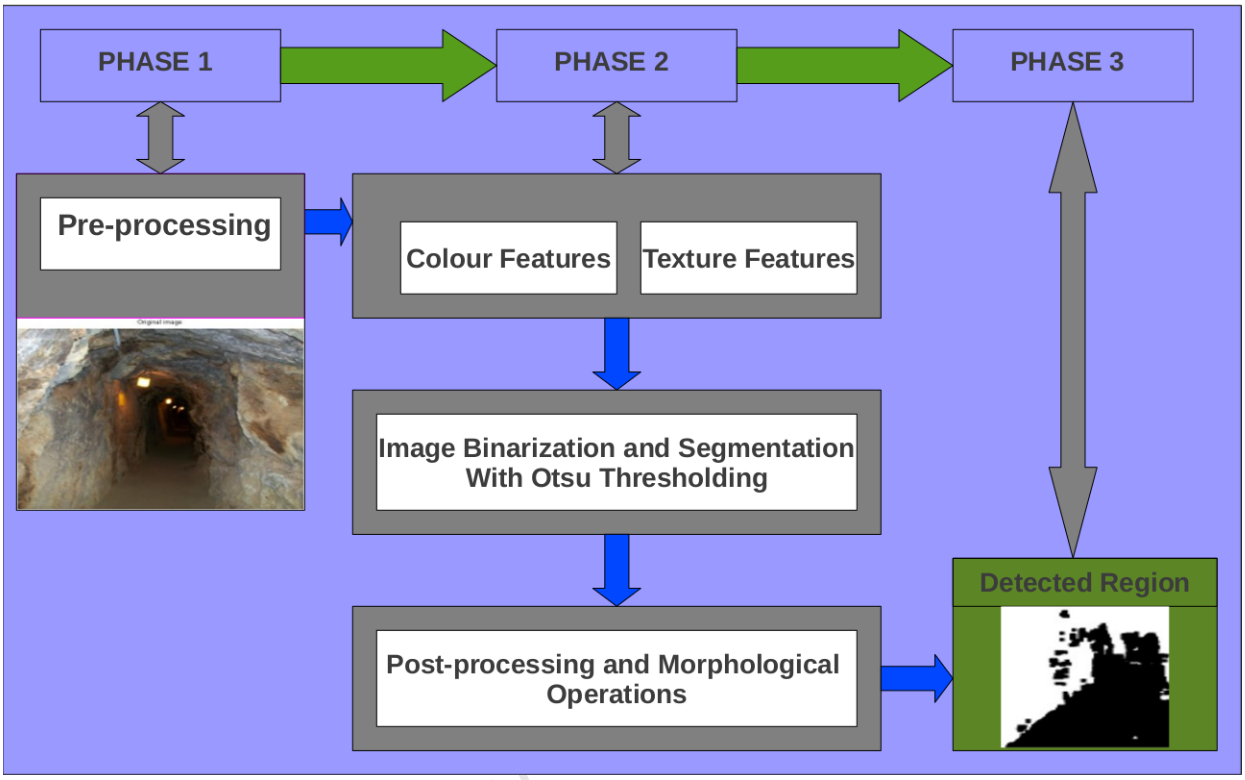



Figure 3.2: Block Diagram of the Entropy Model for Segmentation

University

The first consideration is image pre-processing, which aims to improve the image quality and reduce the size for computational efficiency and quality results. The main computations are done in the second phase, which involves extracting relevant features and then segmenting the image. Details of the processes are presented in the subsequent Sections. It is worth metioning that most of the implementations in this research have been aided with pre-existing imaging libraries.

#### Image Initialisation and Preprocessing

The only step in phase 1 (Figure 3.2) is image pre-processing, which involves two major operations, image downsampling and contrast enhancement. Image downsampling [90], also known as dimensionality reduction, is a common pre-processing step to facilitate computational efficiency.

An automatic learning method for dimensionality reduction is presented by Angelova *et al.* [3]. Their assumption is that different terrains represent different classes. Intuitively, it was assumed that in the lower dimensional space two visually similar terrains, mapped to the same cluster, which are not normally discriminated in the visual space, might be discriminated sufficiently after introducing the automatic dimensionality reduction. However, conventional image downsampling methods do not accurately represent the appearance of the original image, and the perceived appearance of an image is altered when the resolution is lowered [89].

In this research, an image downsampling filter that preserves the appearance of blurriness in the lower resolution image is needed. Several downsampling options exist, such as bicubic, bilinear, antialias and nearest neighbour [23, 57]. However, the choice of downsampling varies for different applications. An appearance-preserving downsampling filter called spatial antialias is adopted in this research. Spatial anti-aliasing is the technique of minimizing the distortion artifacts known as aliasing when representing a high-resolution image at a lower resolution. Anti-aliasing is used in digital photography, computer graphics, digital audio, and many other applications. The choice of resolution for an image depends on the application at hand. The mine images used in this work were down-sampled to 300 $\times$ 225 resolution as part of the pre-processing stage. This simple downscaling method is used for computational efficiency.

Initial processing, contrast enhancement, is usually carried out on raw data prior to data analysis. This is necessary to correct any distortion due to the characteristics of the imaging conditions and imaging system. The greyscale image used as the input is obtained by averaging the three RGB colour channels for each pixel $p$ in image $I$. In order to aid visual interpretation, the image contrast is enhanced with a histogram equalization, as shown in Figure 3.3, and features were computed at each pixel location, $P_i$. Figure 3.3 shows the graphical representation of the number of pixels in an image as a function of their intensity. The x-axis shows the pixel intensity levels, while the y-axis represents the number of pixels corresponding to each intensity level.

Let I be a given image represented as a $P_r$ by $q_r$ matrix of integer pixel intensities ranging from 0 to $L-1$, where $L$ is the number of possible intensity values, often $L = 256$. Let $k$ denote the

University of Cape Town

normalised histogram of $I$. Then

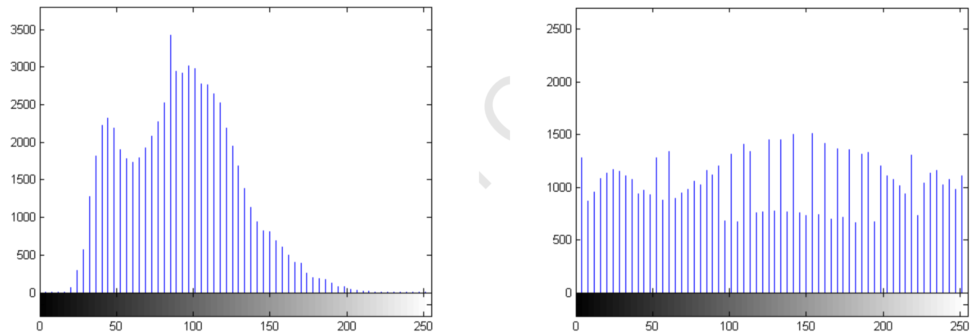
$$k_n = \frac{\text{number of pixels with intensity } n}{\text{total number of pixels}} \quad n = 0, 1, ..., L-1. \tag{3.1}$$

The histogram equalised image, say $k'$, will be defined as

$$k'_{i,j} = \text{floor}((L-1)\sum_{n=0}^{f_{i,j}} k_n). \tag{3.2}$$

The floor of $x$ depicted $\lfloor x \rfloor$ is defined as the nearest integer $\leq x$. Equation (3.2) is equivalent to transforming the pixel intensities, $p$, of $I$ by the function

$$T(p) = \text{floor}((L-1)\sum_{n=0}^{p} k_n). \tag{3.3}$$




(a) Image histogram (b) Transformed histogram

Figure 3.3: Image Histogram and the Corresponding Transformed Histogram

University of Cape Town

### Colour and Texture Features

The visual features used in this research are colour features and texture features. The colour features are readily available as the RGB colour representation of the image $I$ with intensity values for each pixel in the RGB channels ranging from $0 - 255$. The greyscale image $I'$ is obtained by averaging the three RGB colour channels for each pixel $p$ in an image $I$ as demonstrated in Equation (3.4) and Figure 3.4.

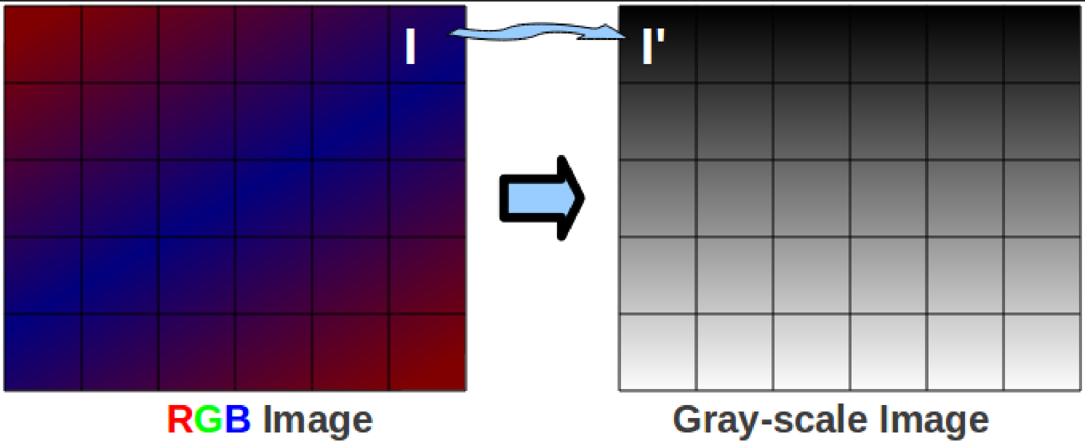



Figure 3.4: Depiction of the Greyscale Image Extraction

University of Cape Town

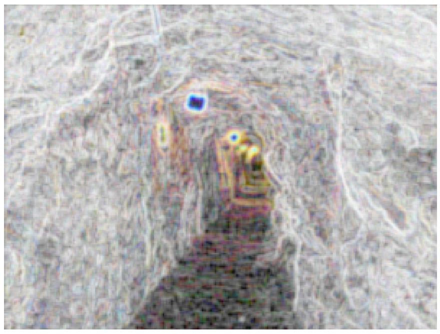
$$I' = \frac{R+G+B}{3}. \tag{3.4}$$

On the other hand, the texture features are computed using a probabilistic approach based on the local entropy. Entropy is a statistical measure of randomness that can be used to characterise the texture of the input image. High entropy indicates a high variance in the pixel values while low entropy is associated with fairly uniform pixel values [8]. Thus, the entropy helps in identifying pixel regions sharing similar textural properties. The entropy for each pixel of the image $I$ is computed with a 9 $\times$ 9 window, which accounts for a reasonable percentage of the textural distribution of each pixel region. The entropy filtered image (EFI) returns an array where each output pixel contains the entropy value of the 9 $\times$ 9 neighborhood around the corresponding pixel in the image $I$. Figure 3.5 shows a mine frame and the corresponding EFI obtained.

**Image Segmentation**

This section describes the second step in phase 2 of Figure 3.2. In this research, the purpose of segmentation is to identify the drivable areas on the mine image frames, that is, the image $I$, into two types of objects: drivable and non-drivable areas. After pre-processing the image and extracting the colour and texture features, the search for an ideal threshold begins using a segmentation technique proposed by Otsu [66]. The Otsu method finds the threshold that

(a) Original Mine Frame



(b) Entropy Filtered Frame

Figure 3.5: Original Mine Image and the Corresponding Entropy Filtered Image

minimises the weighted within-class variance of the image and invariably maximises the between-class variance of the image. The algorithm follows in Section 3.2.4. Otsu calculates the optimum threshold separating those two classes so that their combined spread, also known as the intra-class variance, is minimal, maximising the separation between object and background. The process of finding the optimal threshold is iterative, starting from an arbitrary value. The iterative threshold is set to stop when the change in threshold value is insignificant. This indicates that the threshold is very close to optimum. The mean values are computed by summing the product of each intensity and the corresponding histogram proportion. The sum is then divided by the total sum of the pixel intensity values to get the average.

The thresholding process is seen as the partitioning of pixels of an image in two classes: $P_1$ (object) and $P_2$ (background). This method is recursive and searches the maximisation for the cases: $P_1$ (0,1,..., $T$) and $P_2$ $(T+1, T+2, ..., L-1)$, where $T$ is the chosen optimal threshold and $L$ the number of intensity levels of the image [73, 91]. The Otsu thresholding method exhaustively searches for the threshold that minimises the intra-class variance $\sigma^2_\omega(t)$ defined in Equation (3.5) as a weighted sum of variances of the two classes.

$$\sigma^2_\omega(t) = \omega_1(t)\sigma^2_1(t) + \omega_2(t)\sigma^2_2(t). \tag{3.5}$$

University of Cape Town

### 3.2.4 Otsu Thresholding Method

**Input**: Entropy filtered image $I$

**Output:** Segmented image $I'$

**Step 1:** Estimate histogram and probabilities of each intensity level

**Step 2:** Set up initial $\omega_i(0)$ and $\mu_i(0)$

**Step 3:** Search all possible thresholds $t = 1 \dots L$;

1. Update class probability ($\omega_i$) and class mean ($\mu_i$).

2. Compute $\sigma_b^2(t)$.

**Step 4:** The optimal threshold $T$ corresponds to $\max(\sigma_b^2(t))$

University of Cape Town

#### Morphological Operations

The last step in phase 2 of Figure 3.2 involves the morphological operations. Morphological operations are often used to understand the structure of an image [82]. In this research, the main morphological operations can be likened to flood-filling and are referred to as erosion and dilation. The two operations are explained below:

- Dilation: In dilation, the output value of an image pixel is the maximum value of all the pixels in the input pixel's neighborhood. Thus, dilation broadens the boundaries of regions of white pixels. This is usually done with the support of a structuring element $S_e$ with a specified neighbourhood, which in this case is a 9×9 square structuring element (neighbourhood). The 9×9 neighbourhood is used to perform a sliding window operation on the image for the step-wise decision-making processes.

- Erosion: This is one of the two basic operators in the area of mathematical morphology, the other being dilation. The decision-making process also uses the 9×9 neighbourhood where the pixel value of the output pixel is the minimum value of all the pixels in the input pixel's neighborhood.

  The structuring element is superimposed on top of the input image so that the origin of the structuring element coincides with the input pixel coordinates. If for every pixel in the structuring element, the corresponding pixel in the image underneath is a foreground pixel, then the input pixel is left as it is. If any of the corresponding pixels in the image are background, however, the input pixel is also set to background value. However, the structuring element, used in the process, may have to be supplied as a small binary image, or in a special matrix format, or it may simply be hardwired into the implementation, and not require specifying at all. The structuring element determines the precise details of the effect of the operator on the image, its choice usually depends on the application at hand.

The initial assumption is that the entropy method would return a similar probabilistic distribution for pixel regions sharing the same textural properties (i.e. drivable area) within the mine frame. However, this cannot be guaranteed in its entirety, as there could be some interference (noisy pixels) and thin brightness discontinuities in the processed mine frame. Morphological operations help in reducing such interference by removing isolated blocks within the mine image and thereafter revealing large areas of connected pixels [95]. The erode/dilate filter helps to remove small wrong areas (areas with some noise). Figure 3.6 shows an example of the effect of morphological operation on a mine frame. It is worth mentioning that the method adopted in this research is not to present algorithms with perfect results, but to find a larger amount of pixels (considered as a drivable region in this research) in a safe proportion. A safe proportion corresponds to proposed pixels classification that falls in the true positives (TP) and true negatives (TN) range as defined in this research work.

University of Cape Town

(a) Thresholding result of a mine frame

(b) Erosion/Dilation effect on corresponding mine frame

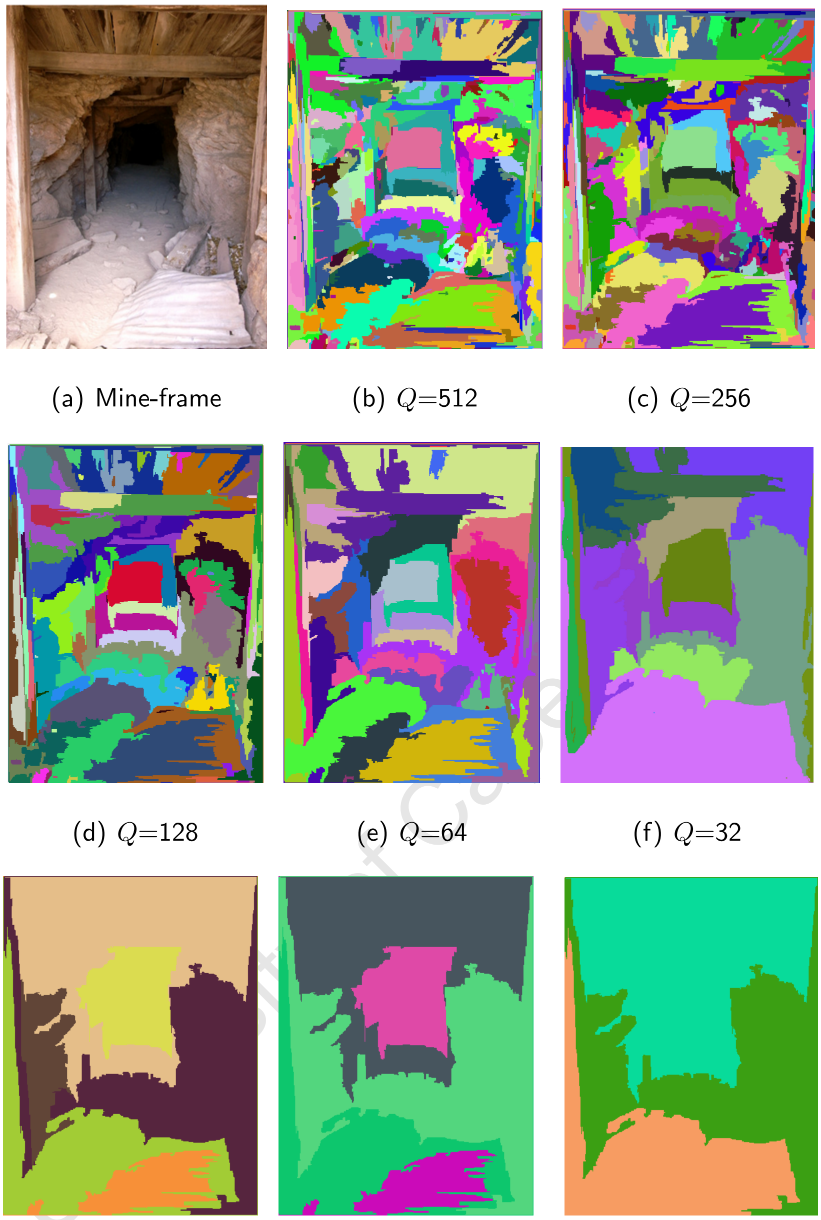
Figure 3.6: Result of Morphological Operations on a Mine Image Classification

### 3.2.5 SRM Approach to Drivability Analysis

The SRM algorithm has two important criteria: the merging predicate and specified cluster $Q$ for the input image [64]. SRM is noted for its computational efficiency, simplicity and good performance, as seen in section 4.1.3. The flexibility of $Q$ is a major advantage as a trade-off parameter that is adjusted to obtain a compromise between the observed results and the strength of the model. In the present experiment, after testing with different values of $Q$, the value $Q = 32$ gave the optimal result for the image classification. Q is a parameter that controls the coarseness and busyness of the classification. Figure 3.7 shows stages of region segmentation on a mine frame at different $Q$ levels. Q is a parameter that needs to be set, controlling the statistical complexity of the optimally segmented image [15]. It was observed that a small value of $Q$ results in few segments (sometimes over-merging, i.e Figure 3.7(i)) and a high $Q$ value may result in too many segments/regions (sometimes under-merging, i.e Figure 3.7(b)). However, the nature of the application at hand would determine the value of the $Q$ parameter. In this experiment, after testing with different values of $Q$, the value $Q = 32$ gave the optimal result for the image classification where the region of interest is clearly distinguished.

Figure 3.8 and Section 3.2.6 present the flow of the algorithm. The algorithm uses a four-

University of Cape Town



(a) Mine-frame (b) $Q$=512 (c) $Q$=256

(d) $Q$=128 (e) $Q$=64 (f) $Q$=32

(g) $Q$=16 (h) $Q$=8 (i) $Q$=4

Figure 3.7: Stages of region merging/segmentation on a mine frame at different $Q$ levels.

connectivity scheme to determine adjacent pixels relative to the centre pixel (in green) as shown in Figure 3.9. The pixels are sorted in ascending order based on the sort function in Equation (2.7). Thereafter, the algorithm considers every pair of pixels $(p, p')$ of the set $D_I$ and performs the statistical test based on the merging predicate [54]. If the regions of the pixels differ and the mean intensity is sufficiently similar to be merged, the two regions are merged. The SRM method presents the list of pixels belonging to each segmented region with their average mean intensities. The anticipated drivable region is composed of pixel region which forms clusters at the base of each observed image $I$ towards the midpoint, when scanning from the left. The underlying assumption is that a pixel region closer to the robot's view tends to form the drivable part, as can be seen in the test cases presented.

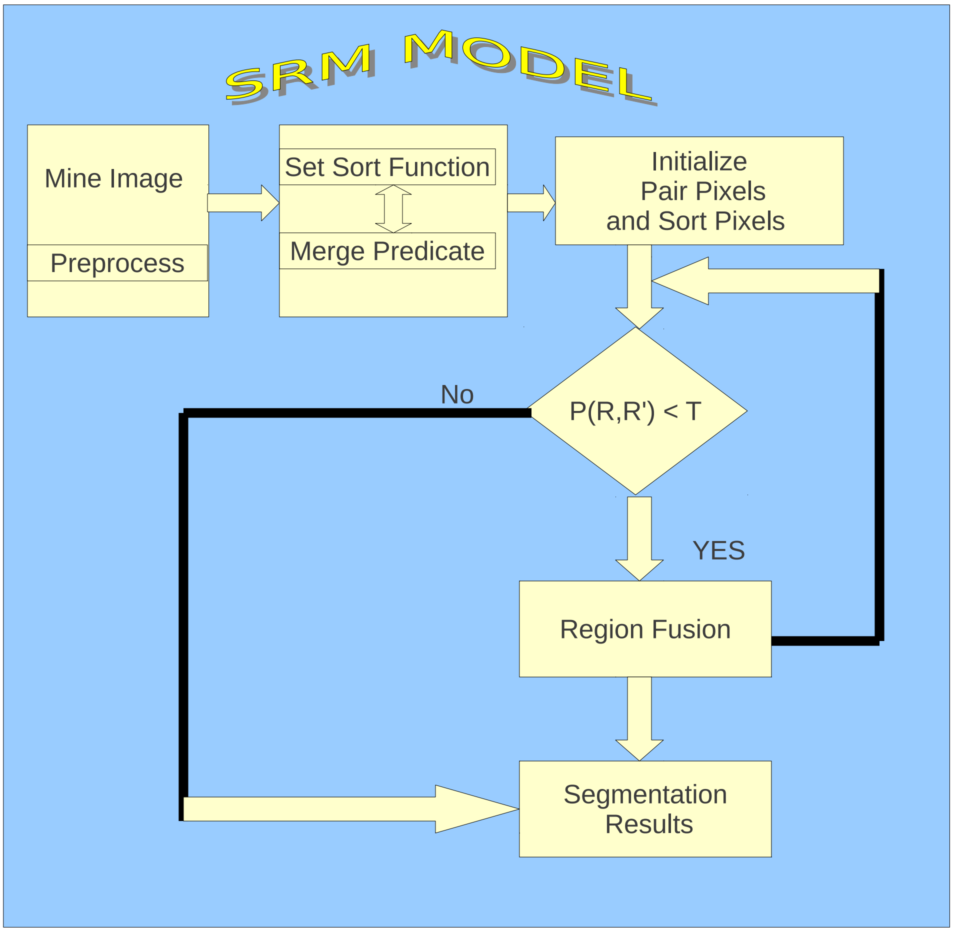



Figure 3.8: Flow of SRM Model

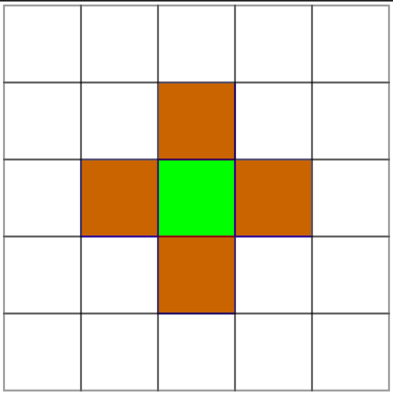


Figure 3.9: Depiction of the Four-Connectivity Scheme

### 3.2.6 Pseudocode for SRM Method

**Input**: Image $I$ and number of segments $Q$

**Output:** Image region segmentation $I'$

---

**Step 1:** Initialise input image $I$ and $Q$

**Step 2:** $D_I = \{$the 4-connectivity adjacent pixels$\}$

**Step 3:** $\bar{D}_I = sort(D_I, f)$

While $\bar{D}_I \neq \emptyset$

for i = 1 to $|\bar{D}_I|$ do

**Step 4:** if $(((P(R_{(p'_i)}, R_{(p_i)}) == true) \; and \; (R_{(p'_i)} \neq R_{(p_i)}))$

**then merge regions** $(R_{(p'_i)}, R_{(p_i)})$

---

### 3.2.7 Evaluation Mechanism

Qualitative evaluation and quantitative (confusion matrix) methods were used in the experiment. The qualitative evaluation, which is the visual comparison, is presented in Sections 4.1 and 5.3.

University of Cape Town

The quantitative evaluation is presented in Section 5.5. In this research, the confusion matrix technique is considered [26]. The confusion matrix procedure is repeated $n$ times for the purpose of cross-validation, with $n \in \{3, 5\}$, where each n subsample is used exactly once as the validation data. The idea is to evaluate the accuracy level (hit rate) of the algorithms (entropy and SRM) in the following context.

- True negatives (TN): The non-matches pixels that were correctly rejected.
- True positives (TP): The number of drivable pixels correctly detected (correct matches).
- False negatives (FN): The proposed pixel matches that were not correctly detected.
- False positives (FP): The proposed pixel matches that are incorrect.

Thus, the accuracy (acc %) is given as

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\%. \tag{3.6}$$

A confusion matrix is a validating tool which contains information about actual and predicted classifications done by a classification system. The data in the matrix is commonly used in evaluating the performance of such a system. An example of a confusion matrix is depicted in Table 3.1.

| | | **Actual** | |
|---|---|---|---|
| | | *Parameter x* | *Parameter y* |
| **Predicted** | *Parameter x* | A | B |
| | *Parameter y* | C | D |

Table 3.1: Schematic of a confusion matrix

The performance accuracy (AC) in the confusion matrix is the proportion of the total number of predictions that were correct. This is generally expressed as :

$$AC = \frac{\sum(\text{left-diagonal-entries})}{\sum(\text{All-entries})} \times 100\%. \tag{3.7}$$

University of Cape Town

Thus, the performance accuracy (AC) of the confusion matrix depicted in Table 3.1 is given below:

$$AC = \frac{A + D}{A + B + C + D} \times 100\%. \tag{3.8}$$

## 3.3 Chapter Summary

In this chapter, the research rationale is re-established and the methodology is presented in detail. This research focuses on assisting robots' mission in mines by enhancing robots' capability of identifying drivable regions in underground terrains. A detailed approach to the entropy and SRM models in drivability analysis is also documented in order to present the underlying principles of these models as applied to current research. The entropy method is based on a probabilistic approach, which computes textural features of the mine images for further processing. The Otsu approach is used to segment the mine images, after which morphological erosion and dilation are used to minimise interference by removing isolated blocks within the mine image and revealing large areas of connected pixels. On the other hand, SRM, which is based on two important components, merging predicate and specified cluster, reconstructs the structural components and retains clusters of the mine images that are closer to the robots view. The evaluation mechanism is presented in this chapter.

University of Cape Town

# Chapter 4

# 2D RESULTS AND EXPERIMENTAL EVALUATION

University of Cape Town

The main focus in this research is to enhance robots' capability of identifying drivable regions within a mine as they traverse a course. This chapter presents some results of the investigation conducted in order to provide useful qualitative conclusions. Furthermore, the entropy and SRM methods adopted in this research are benchmarked with publicly available image frames and methods.

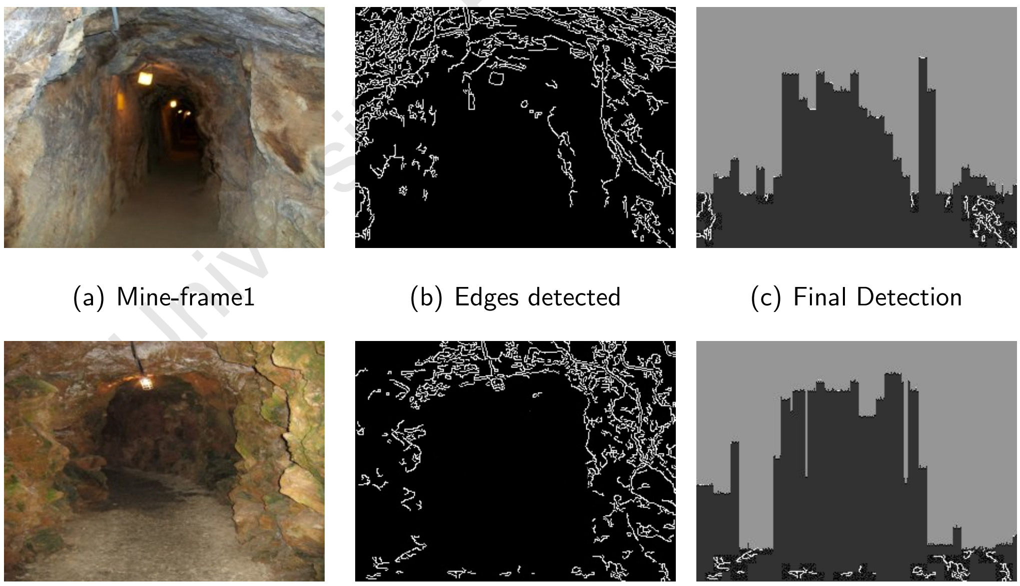
## 4.1 Experimental Evaluation

For the test images used in this work, different test cases of mine frames were carefully chosen from publicly available mine frames [11, 43, 56, 58] and on a stream of images captured from a local mine using common photo cameras. The test cases are representative of different regions, such as the shaft, stope and gallery, in an underground mine environment. The images are chosen to represent a wide variety of objects and contexts. Some of the images contain a single complete object (occlusion) in a portion of the image and some have no objects. The entropy and SRM methods are tested on a stream of rough mine frames and performance evaluation is carried out to provide useful qualitative and quantitative conclusions.

An earlier attempt to utilise an edge detection based method, as adopted in early research on structured terrain classification failed [25]. Edge detection is the process of identifying areas of sharp discontinuities in an image. In the experiment performed, the Canny edge detection method [14] is used for edge features extraction, then pixel regions considered as non-drivable are minimised. Details of the detection method can be found in [25]. Figure 4.1 presents the classification results on some tested mine frames. It is clear from Figures 4.1(b) and 4.1(e) that the results of the edge detection method do not adequately identify the true image boundaries for underground mines terrain classification.

Clearly, the results presented in Figures 4.1(c) and 4.1(f) show that an entirely edge detection based method is not promising. This is partly due to the fact that underground mines are inherently unstructured and many rough edges may be obtained, as seen in Figures 4.1(b) and 4.1(e). Also, simple representations, using the edge detection method, may only be sufficient to detect terrain classes that are easily separable from the rest, i.e the sky, grass and kerb regions in an image.

University of Cape Town



(a) Mine-frame1 (b) Edges detected (c) Final Detection

(d) Mine-frame2 (e) Edges detected (f) Final Detection

Figure 4.1: Test Cases: Mine Frames and the Corresponding Drivable Region Detected.

### 4.1.1 Experiment 1: Observation on the Entropy Approach

In this section, some results obtained on a stream of mine frames using the entropy approach are presented. Figure 4.2 shows the visual assessment of the intermediate and final detection results on some mine frames using the entropy method. The first row of Figure 4.2 consists of original mine frames, the second row presents the entropy filtered images while the third row is the result of Otsu thresholding. The last row presents the final detection with the RGB representation of the corresponding mine frames, the lower (green) regions form the drivable part while the upper (red) regions form the non-drivable part. It was observed that the results produced by the entropy approach tend to be affected by the effect of strong shadow regions on the tested frames. This can be seen in the qualitative results, as most of the connected regions produced in some of the frames are in the shadow areas.

Town

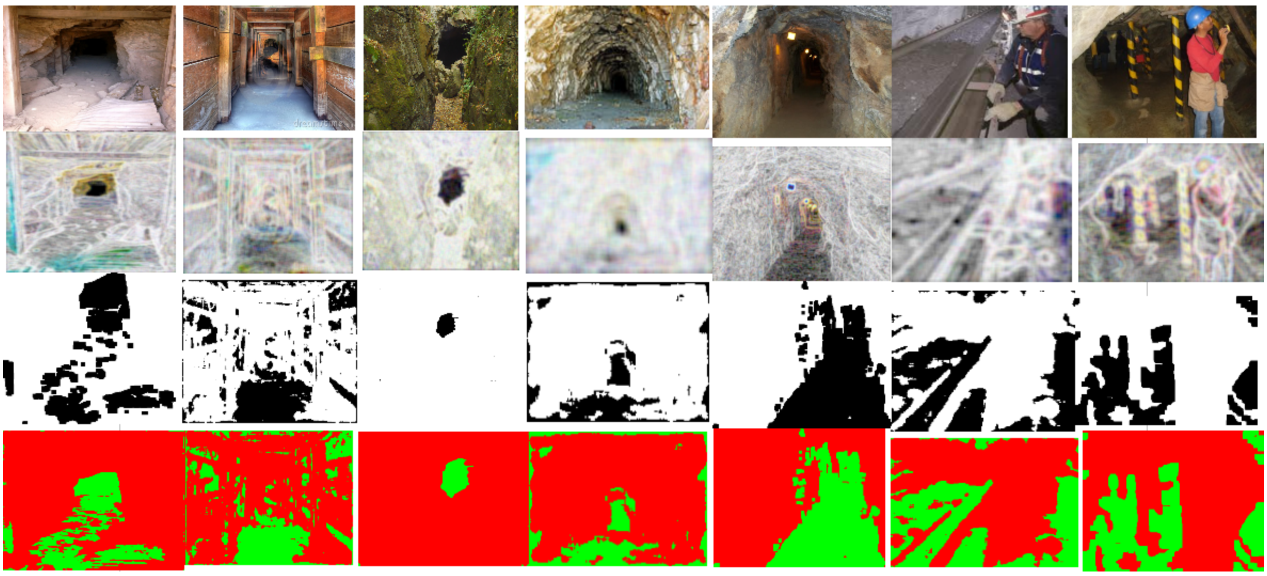


Figure 4.2: Qualitative Results of Entropy Approach on Some Mine Frames

### 4.1.2 Experiment 2: Observation on the SRM Approach

In this section, some of the results obtained using the SRM method are demonstrated in Figure 4.3. The first row is the original mine frames. The second row represents the results of the clusters generated for regions with homogeneity. The third row gives the RGB representation of the drivable regions extracted for corresponding frames. The base (green colour) region indicates the drivable region while the upper (red colour) region represents the non-drivable region. It is evident from these results that SRM has the ability to reconstruct the structural components and retain clusters of the mine images that are closer to the robot's view. Pixel regions closer to the robot's view tend to form most of the drivable region. It is worth mentioning that the choice of $Q$ in this research is only experimentally determined. However, an adaptive means can be explored on choosing the $Q$ parameter in general. This may however also depend on the application at hand.

Figure 4.3: Qualitative Results of SRM Method on Some Mine Frames

University

### 4.1.3 Experiment 3: Qualitative Comparison of Entropy and SRM on Mine Frames

The visual comparison of the entropy and SRM methods on some mine frames is presented in Figure 4.4. In the experiment, the images in the first row are the original mine images, the images in the second row represent drivable regions detected using the entropy approach while the third row represents the drivable regions detected for corresponding frames using the SRM approach. The lower (green colour) regions indicate the drivable regions while the upper (red colour) regions represent the non-drivable regions. It can be seen from the results that the two regions (drivable and non-drivable) were clearly distinguished using mainly the SRM approach in

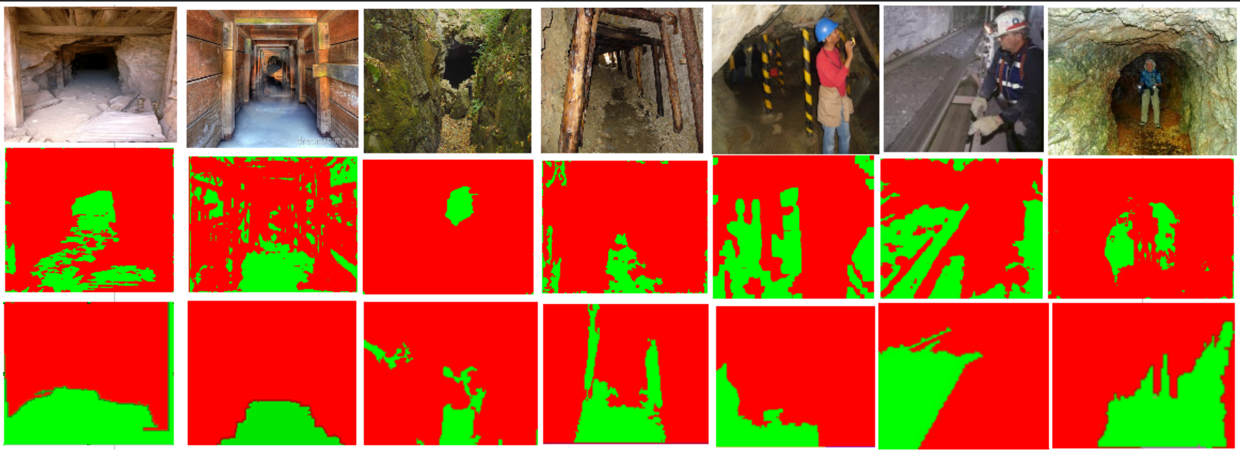


Figure 4.4: Visual Comparison of Entropy and SRM on Some Mine Frames

almost all the tested scenarios. Thus, considering Figure 4.4, one could arguably conclude that qualitatively, the SRM method outperforms the entropy method.

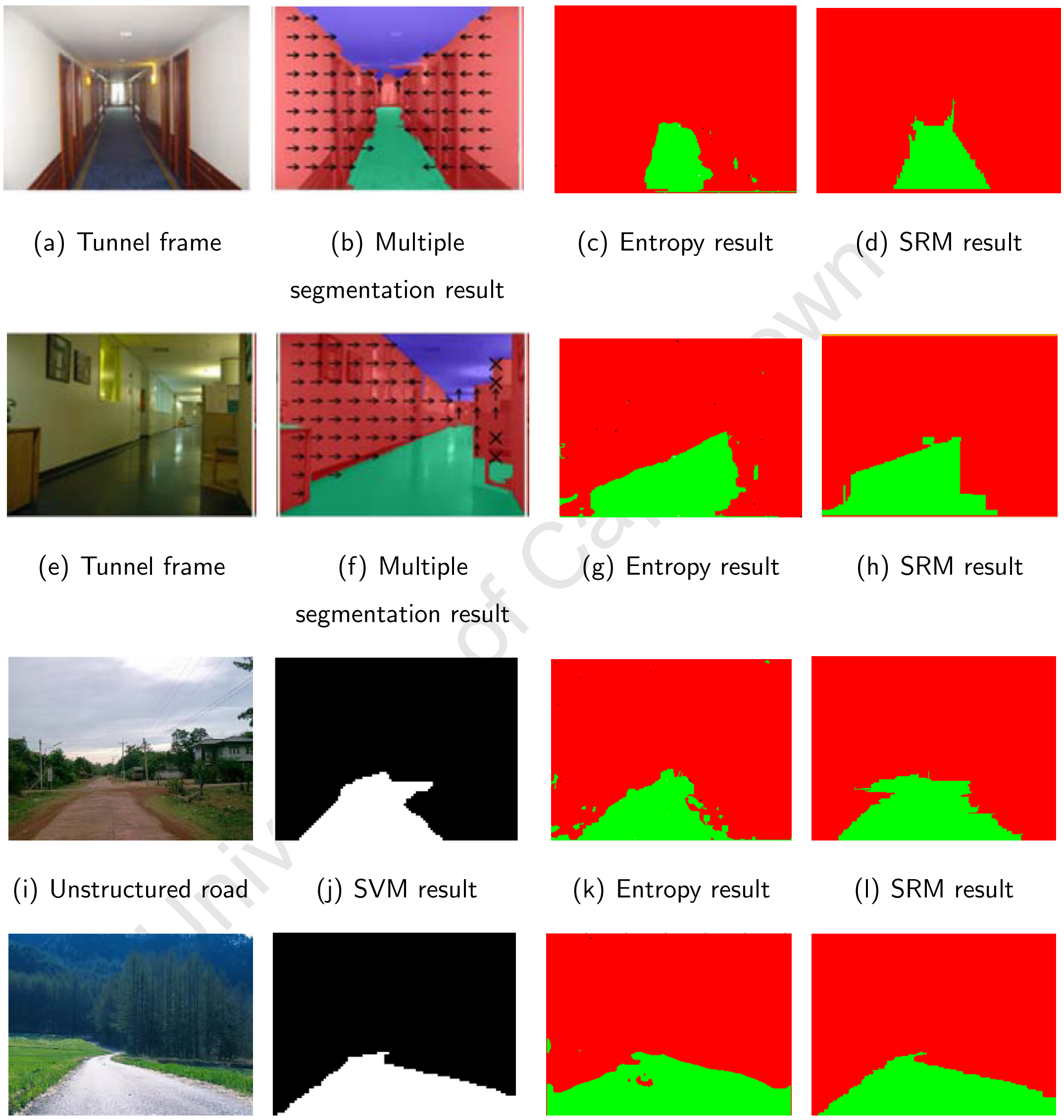
## 4.2 Experiment 4: Benchmarking the Proposed Model

To validate the performance of the detection algorithms, publicly available images and methods are used as a benchmark [22, 46]. The methods used as benchmark are the multiple segmentation approach and the Support Vector Machines (SVM). The multiple segmentation approach [22] focuses on the surface layout extraction from an input image by learning appearance based models of the geometric classes, which coarsely define the 3D scene orientation of the image. Their idea is to classify each image into its respective region, as seen in the world (3D) model, unlike the 2D view, which pretends that the world is flat. On the other hand, the SVM approach [46] attempts to classify unstructured road images using hybrid features. The features used are colour, edge and texture features. SVM attempts to separate two classes into distinctive regions with an optimal decision hyper-plane that has a maximum margin using the training sample. The road region is labelled as +1 while the non-road region is labelled as -1 based on some statistical analyses. Their aim is to distinguish the road region from the background.

Figure 4.5 shows the input images (tunnel frames and ground based unstructured road frames) and their corresponding detection results using the entropy and SRM methods. It is worth mentioning

University of Cape Town

that the results for the input images, using the entropy and SRM algorithms, provide an alternate method for drivable region detection on the frames. Furthermore, there are no 3D results for the tested frames, since it was not possible to gain access to depth maps of the images that provide critical information on the 3D image analysis.



(a) Tunnel frame (b) Multiple segmentation result (c) Entropy result (d) SRM result

(e) Tunnel frame (f) Multiple segmentation result (g) Entropy result (h) SRM result

(i) Unstructured road (j) SVM result (k) Entropy result (l) SRM result

(m) Unstructured road (n) SVM result (o) Entropy result (p) SRM result

Figure 4.5: Validating the Detection Models with Publicly Available Image Frames and Methods

## 4.3 Chapter Summary

In this chapter, a failed attempt at using the edge detection based method for drivable region detection in mines is discussed. The attempt is based on the Canny edge detection method, which is expected to define the true boundaries of the mine image. The inherently unstructured nature of the mine images makes it difficult to identify the true edges using edge detection methods. Thus, an entirely edge detection based approach is not promising in this domain of interest, as observed from the results in Figure 4.1. This chapter also presents the 2D qualitative results of a particular group of mine frames representing different regions of the mines, such as the shaft, stope and gallery (with and without occlusion). The entropy and the SRM models were used to assess the qualitative performance of these mine frames. By comparing the performance of the entropy and SRM models, it becomes clear that the SRM approach is more promising, since it reveals a larger number of connected drivable region pixels in almost all scenarios. The investigation indicates that the robot's visual capability is enhanced as it navigates a mine autonomously. This chapter also validates the performance of the detection algorithms; using publicly available images and methods, results indicate that the detection algorithms provide an alternate means for surface detection.

University of Cape Town

# Chapter 5

# 3D RESULTS AND EXPERIMENTAL EVALUATION

University of Cape Town

This chapter presents the process of the underground tunnel image acquisition and details the procedure by which the XBOX kinect 3D sensors device operates. This chapter presents the 2D and 3D qualitative results generated using the detection algorithms and depth information from the device, as well as the quantitative evaluation results of both terrains (mine and tunnel).

## 5.1 Underground Terrain Image Acquisition

In this research, the tested mine frames are a combination of publicly available mine frames with a stream of mine frames captured in a local mine using common photo cameras. The underground tunnel image frames are captured using the XBOX kinect 3D sensor device. Figure 5.1 shows the general layout of the capturing cycle. After capturing the video stream, consisting of RGB and depth information, the images are separated into frames. This is necessary in order to facilitate processing. Each image frame is then processed using the SRM and entropy algorithms to detect the drivable regions. The corresponding depth information for each image frame is integrated to render the 3D visualisation.

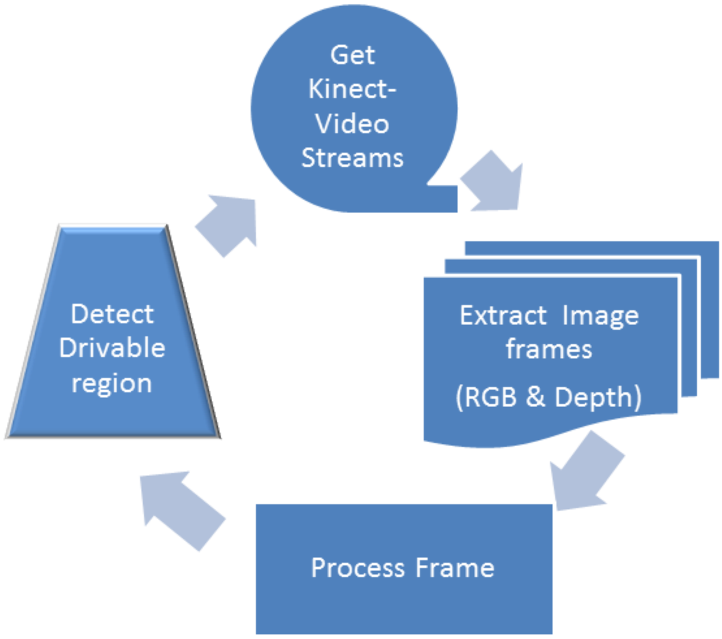



Figure 5.1: Overview of the Kinect Frames Capturing Cycle.

## 5.2 3D Experiment

As noted earlier, purely 2D approaches to robots' vision do not adequately address the great challenge of geometrical image structure in the world model, since it does not fully explore the 3D scene relationship. In this work, a stream of underground tunnel images is captured using the XBOX kinect 3D sensor device in order to explore the ground truth of the image frames and take advantage of the depth map, which enhances one's knowledge of the image scene.

### 5.2.1 Operation of the XBOX Kinect-Motion 3D Sensing Device

The 3D Kinect sensor device with the XBOX 360 console is a relatively new motion-sensing device released in November 2010 [50]. It has wide application in computer gaming and can also be used in speech and voice recognition applications, as it comes with a multi-array mic.

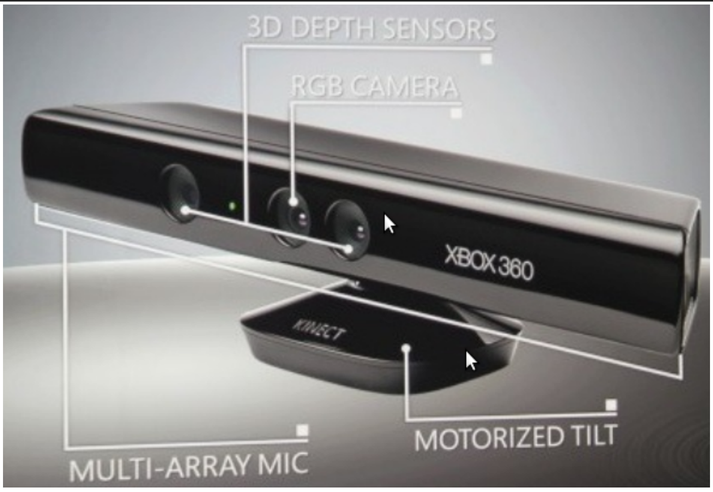



Figure 5.2: XBOX Kinect 3D Sensors

The Kinect device, as shown in Figure 5.2, consists of two major sensors: the RGB sensor and the depth sensor [30]. The RGB sensor produces RGB images, while the depth sensors produce the corresponding depth information, as shown in the first and second columns of Figure 5.3 respectively. The depth sensor consists of the infrared laser projector and an infrared sensor. The laser projector emits infrared light, which illuminates a scene and allows the depth of each pixel to be measured by the infrared sensor, based on the arrival time of the modulated light.

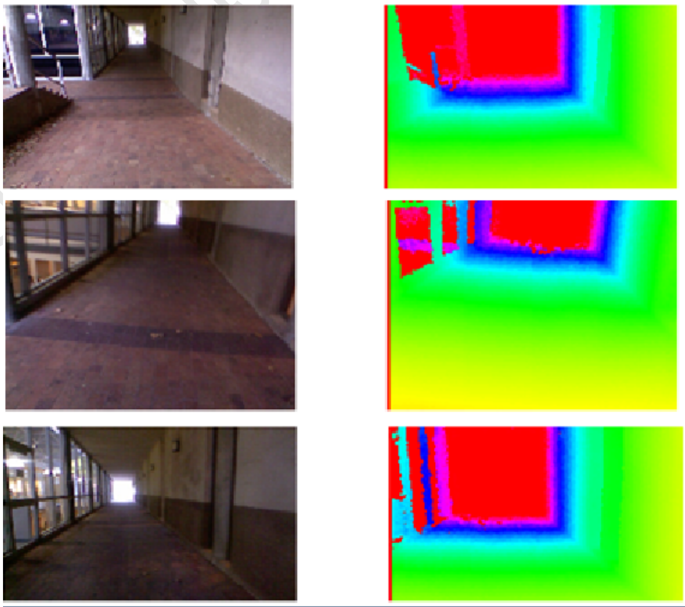


Figure 5.3: Tunnel-Frames in Succession and the Corresponding Depth Images.

University of Cape Town

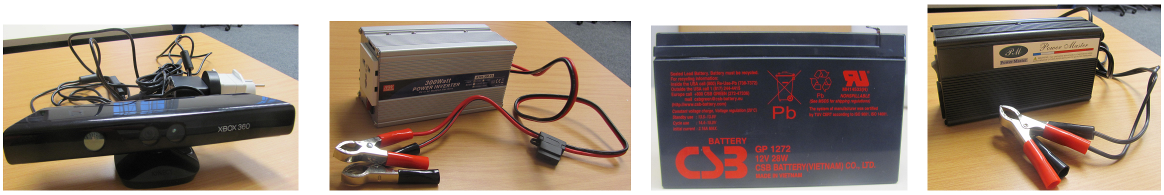
The XBOX Kinect sensor, simultaneously captures depth and RGB images of 640 x 480 resolution at 30 frames per second (fps), as shown in Figure 5.3. The depth data from the device calculates, in millimeter (mm), the distance of each pixel's location relative to the sensor device. The depth images indicate how far or near each pixel's region is perceived by the sensor in the target underground environment. Observe from the second column of Figure 5.3 that a pixel's region of relatively equal depths share the same colour, such as the RGB and yellow regions [2], i.e depth discontinuities in the scene often correspond to colour or brightness changes in the sensor device. Unknown depth pixels might result from the device, especially if the laser rays from the sensor hit a shadow or window. The depth data returns zero in such a situation [50]. Furthermore, unknown depths (vanishing points) are partly due to the limitation in the precision or accuracy of the depth sensor.

In this research, the kinect device is supported by a special power system in the underground environment. The power system is a battery system comprising an inverter and a 12-volt lead acid battery. Figure 5.4 shows the acquisition hardware subsystem (devices) used with the 3D Kinect sensor in the data-capturing process with the factory specifications listed in Appendix A. The Kinect device comes with two major cables, a USB cable, which goes directly to the USB port of the computer for image capture and display, and a power cable, which goes into the battery system.

The battery charger (Figure 5.4(d)) is an alternating current (AC) to direct current (DC) multi-step battery charger, which is ideal for lead acid battery (Figure 5.4(c)) technology. The battery was charged for a minimum of 18 hours in order to ensure its ability to supply high input surge currents to the sensor device through the DC/AC power inverter (Figure 5.4(b)). The inverter, with modified sine wave technology, changes DC to AC power. The converted AC sources can be at the required voltage using an appropriate control switch. Figure 5.5 shows the device components set up with the captured images (RGB and depth) displayed side by side on the laptop screen. The experimental setup is moved along the tunnel pathway for image capture as perceived by an autonomous robot. To capture the underground surface environment, the sensor is mounted at a small tilt angle $\theta$.

Using the images from the Kinect device, depth information is fused with the RGB image in

University of Cape Town




(a) Kinect sensor (b) DC/AC inverter (c) Lead battery (d) Battery charger

Figure 5.4: Device Component of the Kinect Sensor Placement in the Data Capture Process

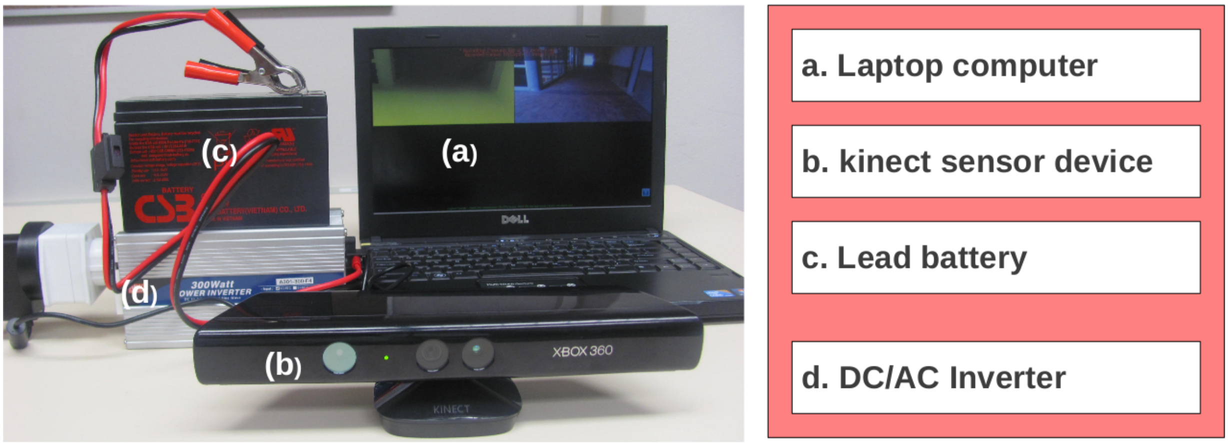
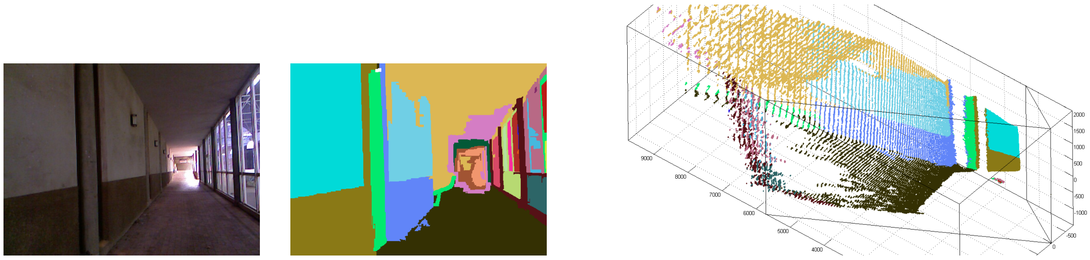



Figure 5.5: Robot Sensor Data Capturing Platform

order to gain better understanding of the imagery and target environment. The fusion of the multi-sensory data (RGB and depth) enhances one's knowledge of the image structure and allows the system to obtain additional information on the vanishing points which might have become available as a result of the perspective effect [33]. Figure 5.6 shows an example of the 2D SRM view in a 3D point cloud representation. This reveals the ground truth of each pixel's position in the image scene as regards the floor, wall and roof region of the tunnel frame.

University of



(a) Tunnel frame (b) 2D SRM view (c) 3D Point cloud of SRM view

Figure 5.6: 2D and 3D Point Cloud View of Tunnel Frame Regions

## 5.3 Experiment 1: Qualitative Observations on Underground Tunnel Terrains

The experiment is conducted on a stream of underground tunnel images captured with the aid of the 3D XBOX Kinect sensor. For the sake of brevity, only a few frames are presented. Figure 5.8 shows the 2D and 3D qualitative results of some tunnel images using the entropy and SRM methods.

It is important to note that when constructing 3D imagery of a scene, the 2D information provides a valuable starting point. The cartesian coordinate in 3D helps to specify each pixel point uniquely and reveals the ground truth of the image classification in reality. The ground truth refers to the information collected in an environment, which allows image data to be compared to real features and materials on the ground. On a 3D cartesian coordinate system, the $x$ and $y$ axes give the pixel value information while the third, $z$, axis depicts the depth information. The depth cue provides useful information about the 3D scene of the image classification as regards the floor, wall and roof region of the tunnel frames. Thus, it creates an accurate understanding of where an autonomous robot should navigate in real time.

To express the 3D coordinates of the pixel points, a depth coordinate system with its origin at the perspective centre of the infrared sensor is considered. Creating a meshgrid structure, as demonstrated in Figure 5.7, which corresponds to the resolution of the image, 640 by 480, and then interpolating the corresponding depth information at those points results in a point

University of Cape Town

cloud representation of the image frames. The meshgrid structure provides an easy platform to integrate the depth information. By integrating the multisensory input data, the 3D view of the classified terrain as indicated in Figures 5.8(c),5.8(e) and 5.8(h) is obtained.

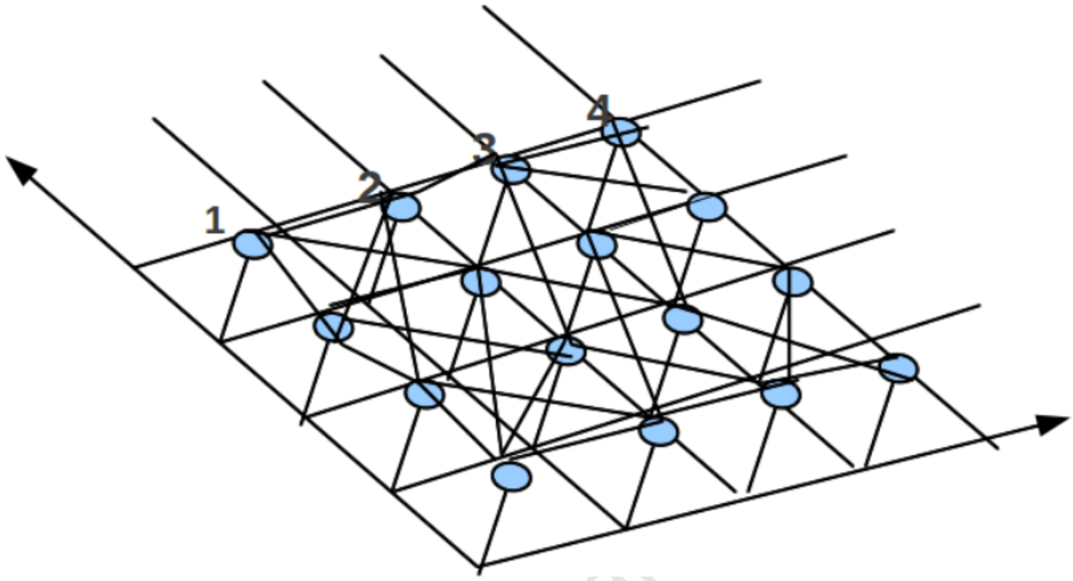
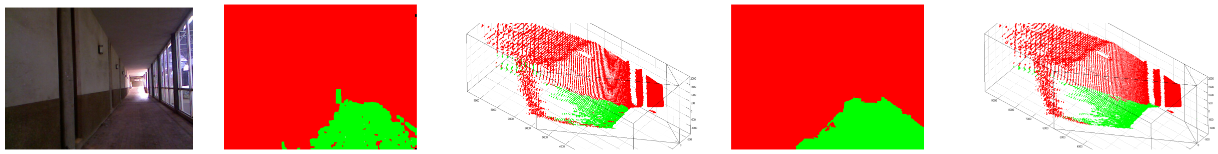
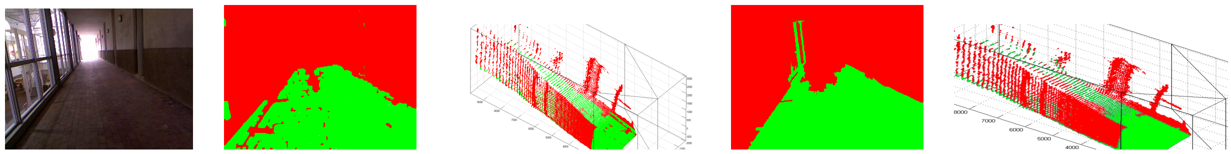



Figure 5.7: Depiction of the Meshgrid Structure for Connecting Pixel Points for 3D Visualisation

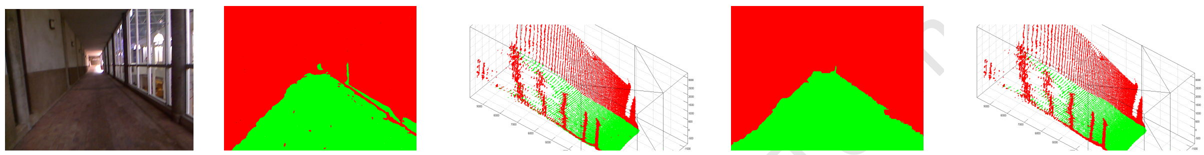
University of Cape Town



(a) Tunnel frame (b) 2D entropy (c) 3D entropy (d) 2D SRM (e) 3D SRM



(f) Tunnel frame (g) 2D entropy (h) 3D entropy (i) 2D SRM (j) 3D SRM



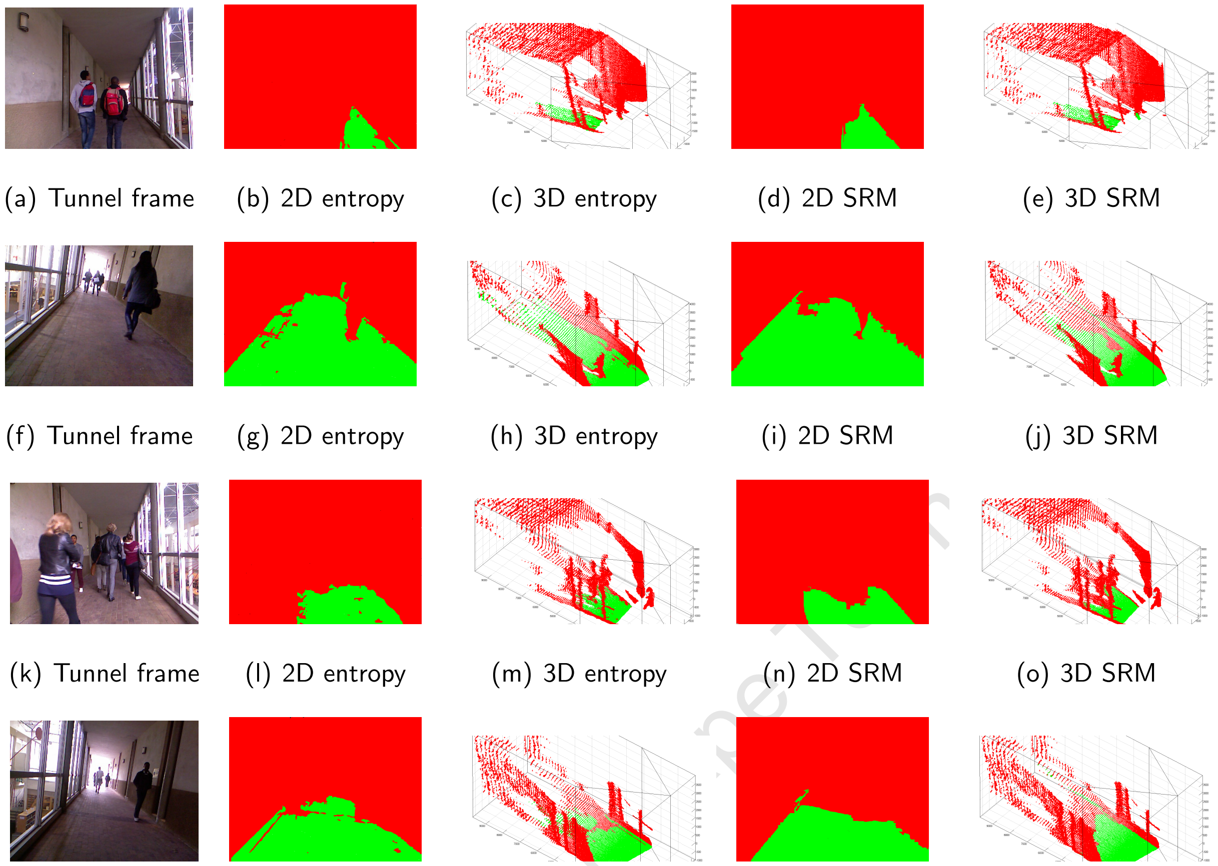
(k) Tunnel frame (l) 2D entropy (m) 3D entropy (n) 2D SRM (o) 3D SRM

Figure 5.8: 2D and 3D Comparison of the Entropy and SRM Models on Tunnel Frames

## 5.4 Experiment 2: Qualitative Observations on Underground Tunnel Terrains

Experiment is also conducted on a stream of underground tunnel images with occlusion. Some of the frames have a single object (occlusion) in the portion of the image, while others have multiple objects. It is clear, from Figure 5.9, that the presence of obstacles in the images reduces the drivable space for autonomous navigation. However, it is important for an autonomous robot to navigate the safe drivable (green) regions in the image frames to minimise accidents.

University of Cape Town



(a) Tunnel frame (b) 2D entropy (c) 3D entropy (d) 2D SRM (e) 3D SRM

(f) Tunnel frame (g) 2D entropy (h) 3D entropy (i) 2D SRM (j) 3D SRM

(k) Tunnel frame (l) 2D entropy (m) 3D entropy (n) 2D SRM (o) 3D SRM

(p) Tunnel frame (q) 2D entropy (r) 3D entropy (s) 2D SRM (t) 3D SRM

Figure 5.9: Occlusion Cases: 2D and 3D Comparison of the Entropy and SRM Models on Tunnel Frames

## 5.5 Experiment 3: Quantitative Evaluation of the Underground Terrains

This section presents the quantitative results of the two terrains (mine and tunnel), based on the applied algorithms. The quantitative performance of both entropy and SRM approaches to drivability is evaluated using the confusion matrix validation process n times ($n \in \{3, 5\}$). In the evaluation, pixel positions are randomly hand-labelled with the aid of an automated code (10 pixels per time for n-fold validation, making 30 pixels per frame [30 frames = 900 pixels] for 3-fold and 50 pixels per frame [30 frames = 1500 pixels] for 5-fold ). The correctness of the pixel $(i, j)$ is evaluated based on its current classification position$(x, y)$ in the detected frame relative to its position $(x, y)$ in the original frame. The estimated confusion matrix validation accuracy is the overall number of correct classifications divided by the number of instances in the image-data $I_d$. Table 5.1 shows the quantitative performance of the algorithms with $n = 3$ and $n = 5$. One can see that the entropy method has a higher accuracy in underground tunnel classification than in underground mine classification. This is partly due to the fact that underground mines are very unstructured compared to tunnels, which are relatively smoother. However, one can conclude that the SRM method outperforms the entropy method in almost all scenarios.

Table 5.1: Comparing Drivability Analysis of Underground Terrains Using Entropy and SRM algorithms

| Terrains | Algorithms | n-fold confusion matrix validation | Correctly classified pixels (TP, TN) | Incorrectly classified pixels (FP, FN) | Accuracy of detection (%) |
|---|---|---|---|---|---|
| Underground mines (30 image frames) | Entropy | 3 | 520 | 380 | 57.78 |
| | | 5 | 815 | 685 | 54.33 |
| | SRM | 3 | 745 | 155 | 82.78 |
| | | 5 | 1200 | 300 | 80.00 |
| Underground tunnels (100 image frames) | Entropy | 3 | 2250 | 750 | 75.00 |
| | | 5 | 3970 | 1030 | 79.4 |
| | SRM | 3 | 2750 | 250 | 91.67 |
| | | 5 | 4480 | 520 | 89.60 |

University of Cape Town

## 5.6 Chapter Summary

This chapter presents the quantitative performance of the models on mine and tunnel images and details the 3D results. By comparing the performance of the entropy and SRM models, it becomes clear that the SRM approach is more promising, and outperforms the entropy method in almost all scenarios. The details of the sensor device used in capturing the underground tunnel image frames are also documented. The XBOX Kinect 3D sensors produce streams of RGB and depth images and facilitate the analysis of detection results in 3D view, unlike a common photo camera that only allows 2D presentation. Thus, integrating the depth information allows the system to obtain better understanding of the classification result. The 3D results make clear the roof, floor and wall regions of the image frames and thus provide a more useful analysis.

University of Cape Town

# Chapter 6

# CONCLUSION, RECOMMENDATION AND FUTURE WORK

University of Cape Town

This chapter summarises and concludes this research work. Possible future work is also suggested in this chapter.

## 6.1 Research Summary and Conclusion

Computer vision is a challenging area and much research is still being carried out in this area, although less attention is paid to underground terrains, probably because of their roughness. The motivation for this research is the dual needs for safety and efficient productivity constantly faced by the mining industry. While robots exhibit a great deal of potential to deal with safety issues in mines, an effective vision model is critical to their success. The use of robot vision in underground terrains has received little attention in research. This work aims to achieve autonomous drivable region detection in underground mine environments. This research has demonstrated the feasibility of enhancing autonomous robots' capability of identifying drivable regions in underground terrains. Different regions of mines, representing a wide variety of terrains, ranging from the stope, shaft and gallery, are investigated. Investigation is also conducted on underground tunnel images (with and without occlusion). The augmentation of the 2D underground tunnel results with 3D drivability maps for autonomous robots further reveals the ground truth of the classification results. The results show that it is possible to detect a good drivable region in underground terrains using the proposed approach.

The entropy approach and the SRM algorithm are adopted as means of identifying drivable regions within a mine frame. The performance of both the entropy and SRM methods in drivability analysis is evaluated. Investigating the performance of both methods shows that SRM outperforms the entropy approach in almost all the scenarios investigated and determination of the drivability of an underground terrain can be achieved. However, the entropy approach can be improved upon for better detection results. Using the entropy approach, the computed local entropy gives useful textural information about the pixels' distribution. The entropy returns probabilities of the randomness of the pixel, $p_i$, grey tone within a fixed window. The probabilistic textural information was used in the mine image classification together with Otsu thresholding. The SRM algorithm, on the other hand, is able to reconstruct the main structural components of the underground mine imagery by a simple but effective statistical analysis. The SRM method worked well on a variety of mine frames tested, as shown in Figures 4.3 and 4.4 and Table 5.1. It can be seen from the results that the two regions (drivable and non-drivable) were most clearly distinguished with the SRM method. Furthermore, integrating the depth maps from the XBOX Kinect 3D sensors in the 3D perception representation also reveals the level of correctness of the image classification, as well as the ground truth of the classification results, as shown in Figures 5.8 and 5.9.

University of Cape Town

The field of 3D analysis and multisensory data fusion is gaining increasing attention from researchers. This is partly due to emerging trends in the commercial availability of 3D scanning systems or devices that produce a high information accuracy level for a variety of applications. This research has successfully taken advantage of the multisensory information from the Kinect device to analyse the region of interest in an imagery. The utilisation of 3D coordinates perception revolutionised computer vision, image analysis and autonomous processes by providing a better understanding of the image 3D structure.

## 6.2 Recommendation

The major focus in this work is feature extraction and classification for front view mine frame detection and detecting drivable and non-drivable regions of underground terrains in 2D by augumentation with 3D results. This, in turn, enhances autonomous robots' visual capability to identify drivable regions in the mine environment. The idea is that the robot should go as far as it can see; when it gets there it would see further. It is important to note that how far the robot would see is dependent on the nature of the sensors device deployed, as well as the precision of the device. The result of this work is a useful application that would accelerate further motion and path planning (control and mechanical decisions) for autonomous robots' navigation in underground terrains. The solution presented in this work would enhance the development of a platform for performing autonomous safety inspections and any other autonomous task in underground terrains. This research has demonstrated the feasibility of enhancing robots' capability of identifying drivable regions in a mine environment and has significant application potential in South Africa. However, the solution presented in this research is more than a case study and applicable to underground terrains in general.

University of Cape Town

## 6.3 Future Work

The current classification results can be improved upon by utilising more machine learning algorithms or hybrid methods in mine frame detection for better performance and future adoption. Furthermore, a natural extension is to use alternate methods of performing drivability analysis in underground terrains. For example, performing drivability analysis can be explored using other distance sensors, such as the laser, in acquiring depth maps for 3D drivability analysis. In addition, drivability analysis can be explored to handle environmental noise such as shadow.

# Appendix A

# Devices Factory Specifications

University of Cape Town

Table A.1: Devices Factory Specifications

| Device | Specifications |
|---|---|
| kinect Sensor (XBOX 360) | Sensor: Colour & depth-sensing lenses<br>Sampling frequency: 30 frames/sec<br>**Data streams**<br>• 640 × 480 @ 30 Hz (32-bit colour)<br>• 640 × 480 @ 30 Hz (16-bit depth (mm))<br>Depth sensor range: 1.2-3.5m (4-11 ft)<br>Physical tilt range: $27^0$<br>Angular view : $43^0$ vertically and $57^0$ horizontal<br>Connectivity: USB 2.0 |
| Lead Battery (GP 1272) | Voltage per unit: 12V<br>Capacity: 7.2 Ah @ 20hr-rate to 1.75V per cell @ $25^oC(77^oF)$<br>Resistance: $\approx$ 23m$\Omega$<br>Maximum discharge current: 100A/130A(5sec)<br>Maximum charging current limit: 2.16A<br>Weight: $\approx$ 2.4kg(5.28lbs)<br>Nominal operating temperature range: $25^oC \pm 3^oC(77^oF \pm 5^oF)$ |

Table A.2: Devices Factory Specifications

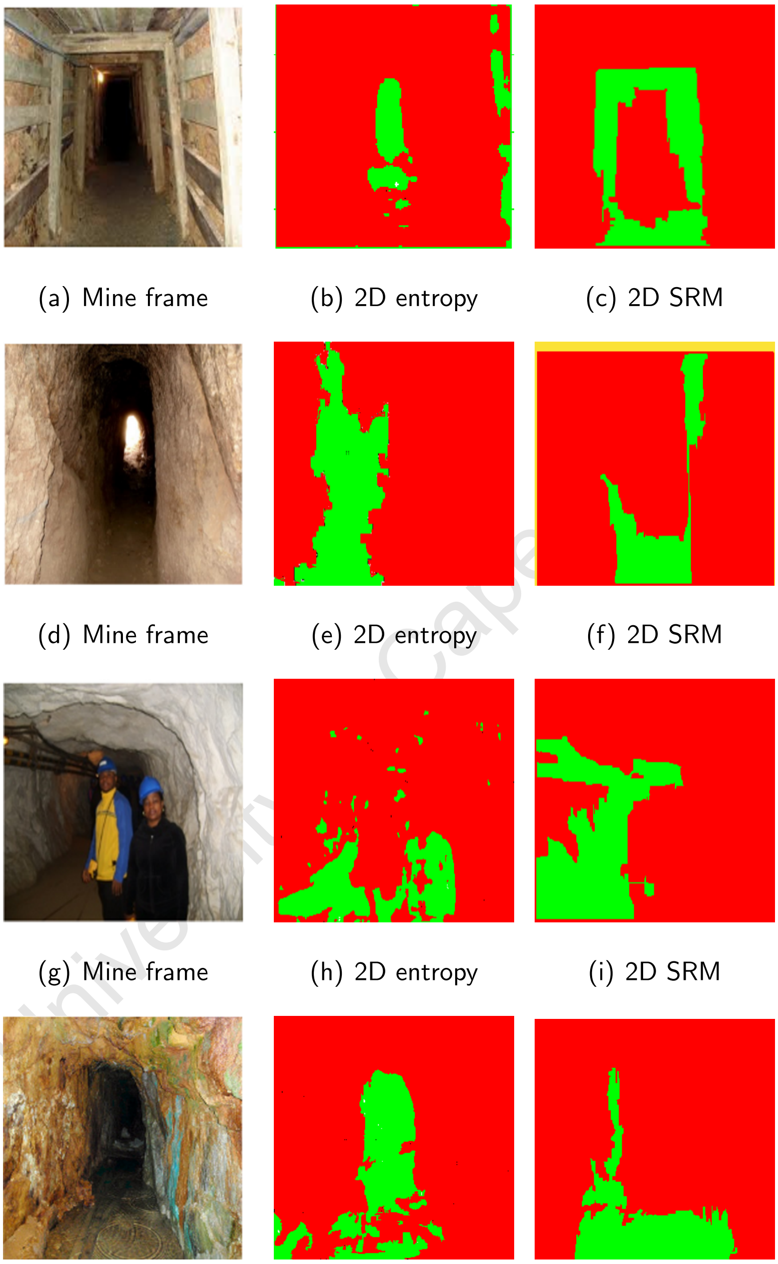
| **Device** | **Specifications** |
|---|---|
| Power inverter (DC/AC) | Rated output power: 300W<br>Output surge power: 600W<br>Output Voltage: 230Vac<br>Frequency: 50Hz $\check{A} \pm 1\%$<br>Cooling: $>$100W or $>$40(NTC) fan on<br>Waveform: Modified Sine wave<br>Working Humidity: 20 % - 90 % RH non-condensing<br>Efficiency : 85% |
| Battery charger (PM-0312F-3L) | Input range: 90-240 Vac<br>Frequency: 50-60 Hz<br>Weight: 1.1kg<br>O/P: 12 Vdc/ 3 Amp GER. TYPE<br>Dimensions: 180 $\times$ 90 $\times$ 50 mm<br>Features: Multi steps charger setup $\&$ auto shutdown<br>Accesory: AC wire $\&$ DC alligator clip |

University of Cape Town

# Appendix B

# 2D Qualitative Comparison on Mine Frames

University of Cape Town



(a) Mine frame (b) 2D entropy (c) 2D SRM

(d) Mine frame (e) 2D entropy (f) 2D SRM

(g) Mine frame (h) 2D entropy (i) 2D SRM

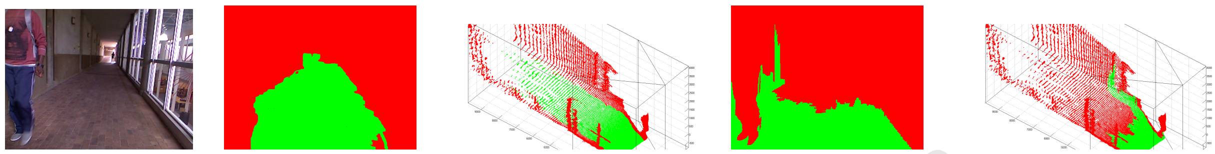
(j) MIne frame (k) 2D entropy (l) 2D SRM

Figure B.1: 2D Visual Comparison on Mine Frames

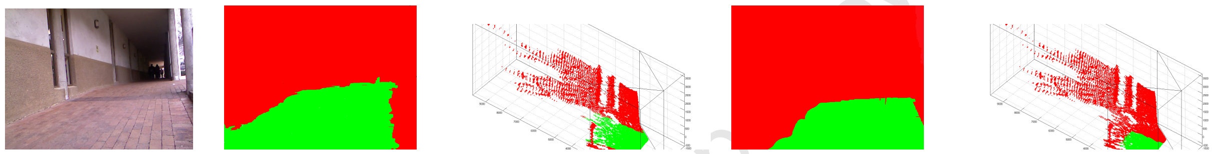
# Appendix C

# 2D and 3D Qualitative Comparison on Tunnel Frames

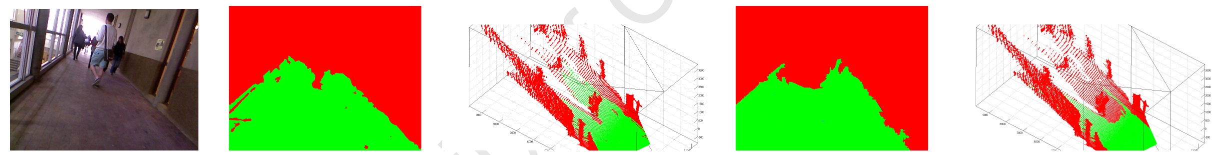
University of Cape Town



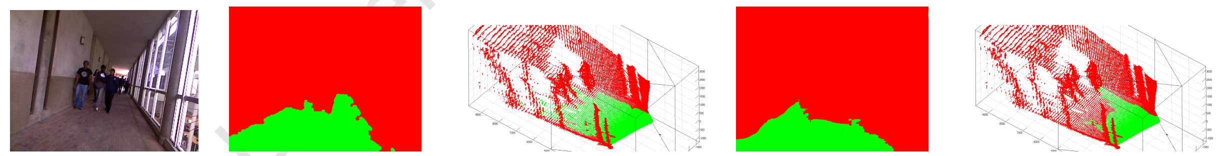
(a) Tunnel frame (b) 2D entropy (c) 3D entropy (d) 2D SRM (e) 3D SRM



(f) Tunnel frame (g) 2D entropy (h) 3D entropy (i) 2D SRM (j) 3D SRM



(k) Tunnel frame (l) 2D entropy (m) 3D entropy (n) 2D SRM (o) 3D SRM



(p) Tunnel frame (q) 2D entropy (r) 3D entropy (s) 2D SRM (t) 3D SRM

Figure C.1: 2D and 3D Comparison of the Entropy and SRM Models on Tunnel Frames

University of Cape Town

# References


[1] S. Aminossadati and K. Hooman. Numerical simulation of ventilation air flow in underground mine workings. In *12th U.S./North American Mine Ventilation Symposium Wallace (ed)*, pages 253–259. The University of Nevada, Reno, June 2008.

[2] H. Andreasson, R. Triebel, and A. Lilienthal. Vision-based interpolation of 3D laser scans. In *Proceedings of the IEEE International Conference on Autonomous Robots and Agents*, pages 455–460. ICARA, August 2006.

[3] A. Angelova, L. Matthies, D. Helmick, and P. Perona. Dimensionality reduction using automatic supervision for vision-based terrain learning. In *IEEE Conference on Robotics: Science and Systems (RSS)*, pages 1–8. RSS, March 2007.

[4] A. Angelova, L. Matthies, D. Helmick, and P. Perona. Fast terrain classification using variable-length representation for autonomous navigation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. CVPR, March 2007.

[5] A. Angelova, L. Matthies, D. Helmick, and P. Perona. Learning slip behaviour using automatic mechanical supervision. In *IEEE International Conference on Robotics and Automation*, pages 1741 – 1748. ICRA, April 2007.

[6] J. Nsasi Bakambu and V. Polotski. Autonomous system for navigation and surveying in underground mines. *Journal of Field Robotics*, 10(DOI: 10.1002/rob.20213):829–847, July 2007.


University of Cape Town

[7] L. K. Bandyopadhyay, S. K. Chaulya, P. K. Mishra, A. Choure, and B.M. Baveja. Wireless information and safety system for mines. *Journal of Scientific and Industrial Research*, 68(1):107–117, 2009.

[8] S. Battiato. Entropy and gibbs distribution in image processing: An historical perspective. *University of Catania, Department of Mathematics and Informatics*, pages 1–21, 2001.

[9] S. Battiato, G.M. Farinella, and G. Puglisi. SVG vectorization by statistical region merging. In *Proceedings of 4th Conference Eurographics Italian Chapter*, pages 1–7. The Eurographics Association, July 2006.

[10] J. Benjamin, Y. Papelis, R. Pillat, G. Stein, and D. Harper. A practical approach to robotic design for the DARPA urban challenge. *Journal of Field Robotics*, 25(8):528–566, July 2008.

[11] Bretterk. Ironminers. Available online: http://brettberk.com (accessed on 10 May 2012).

[12] A. Broggi, C. Caraffi, S. Ghidoni, P. Paolo Porta, and P. Zani. Terramax vision at the urban challenge 2007. *IEEE Transactions on Intelligent Transportation Systems*, 11(1):195–205, March 2010.

[13] F. Calderero and F. Marques. Region merging techniques using information theory statistical measures. *IEEE Transactions on Image Processing*, 19(16):1567–1586, June 2010.

[14] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, 1986.

[15] M. E. Celebi, H. A. Kingravi, J. Lee, W. Stoecker, J. M. Malters, H. Iyatomi, Y. A. Aslandogan, R. Moss, and A. A. Marghoob. Fast and accurate border detection in dermoscopy images using statistical region merging. *TexasWorkforce Commission*, James A. Schlipmann Melanoma Cancer Foundation and NIH (SBIR 2R44 CA-101639-02A2)(16):1–10, 2006.

[16] Z. Chunsun, B. Emmanuel, and G. Armin. Knowledge-based image analysis for 3D road reconstruction. *Asian Journal of Geoinformatics*, 1(1):3–14, June 2001.

University of Cape Town

[17] G. Chunzhao and M. Seiichi. Drivable road region detection based on homography estimation with road appearance and driving state models. In *Proceedings of the 4th IEEE International Conference on Autonomous Robots and Agents*, pages 204–209, February 2009.

[18] Cienciaviva. Introduction to mining. Available online: http://www.cienciaviva.pt (accessed on 20 August 2012).

[19] C. Clark, C. Olstad, K. Buhagiar, and T. Gambin. Archaeology via underwater robots: Mapping and localization within maltese cistern systems. In *10th International Conference on Control, Automation, Robotics and Vision*, pages 662–667. ICARCV, December 2008.

[20] C. Conner, A. Rizzi, and H. Choset. Composition of local potential functions for global robot control and navigation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3546– 3551, October 2003.

[21] A. Damarupurshad. African (mining-related) government departments, mineral/mining-related organisations and mining companies with interests in africa. *Department of Minerals and Energy (Republic of South Africa)*, D13(ISBN: 0-9584376-7-X):1–58, June 2004.

[22] H. Derek, A. Alexei, and H. Martial. Recovering surface layout from an image. *International Journal of Computer Vision*, 75(1):151–172, February 2007.

[23] R. Dugad and N. Ahuja. A fast scheme for image size change in the compressed domain. *IEEE Transactions on Circuits and System for Video Technology*, 11(4):461–474, 2001.

[24] J. Statutes Editor. Juta's statutes of South Africa 2009/10, 2010. ISBN 978 0 7021 8472 7.

[25] O. Falola, I.O. Osunmakinde, and A. Bagula. Supporting drivable region detection by minimising salient pixels generated through robot sensors. In *Proceedings of the Twenty-First Annual Conference of the Pattern Recognition Association of South Africa (PRASA)*, pages 87–92. MIAPR ISBN 978-0-7992-2470-2, 2010.

[26] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(1):861–874, December 2005.

University of Cape Town

[27] D. Ferguson, A. Morris, D. Hàhnel, C. Baker, Z. Omohundro, C. Reverte, S. Thayer, W. Whittaker, W. Burgard, and S. Thrun. An autonomous robotic system for mapping abandoned mines. In *Proceedings of Conference on Advances in Neural Information Processing Systems*, pages 1–8. MIT Press, 2003.

[28] J. P. Godbaza, M. J. Cree, and A. A. Dorrington. Mixed pixel return separation for a full-field ranger. In *Proceedings of the 23rd International Conference on Image and Vision Computing, New Zealand (IVCNZ)*, pages 1–6. IEEE, November 2008.

[29] J. C. Gonzàlez, S. Martìnez, A. Jardòn, and C. Balaguer. Robot-aided tunnel inspection and maintenance system. In *Proceedings of the 26th International Symposium on Automation and Robotics in Construction*, pages 420–426. Elsevier Automation in Construction, August 2009.

[30] J. J. Green, K. Hlophe, J. Dickens, R. Teleka, and M. Price. Mining robotics sensors. *International Journal of Engineering and Advanced Technology (IJEAT)*, 1(4):8–15, April 2012.

[31] S. Greve, Marcin Grzegorzek, Carsten Saathoff, and Dietrich Paulus. Classification of image regions using the wavelet standard deviation descriptor. In *Proceedings of the International Multiconference on Computer Science and Information Technology*, pages 703–708, 2010.

[32] D. W. Hainsworth. Teleoperation user interfaces for mining robotics. *Journal of Autonomous Robots*, 11(1):19–28, July 2001.

[33] R. Hartley and F. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.

[34] C.J.H. Hartnady. South Africa's gold production and reserves. *South African Journal of Science*, 105(1):328–330, September 2009.

[35] M. Haselich, M. Arends, D. Lang, and D. Paulus. Terrain classification with markov random fields on fused camera and 3D laser range data. *Active Vision Group, AGAS Robotics, University of Koblenz-Landau, 56070 Koblenz, Germany*, pages 1–6, June 2011.

University of Cape Town

[36] A. Hemami and F. Hassani. An overview of autonomous loading of bulk material. In *26th International Symposium on Automation and Robotics in Construction*. ISARC, 2009.

[37] M. A. Hermanus. Occupational health and safety in mining- status, new developments and concerns. *Journal of the Southern African Institute of Mining and Metallurgy*, 107(1):531–538, May 2007.

[38] K. Hlophe. GPS-deprived localisation for underground mines. In *CSIR 3rd Biannual Conference, Science Real and Relevant,*, page 1. CSIR International Convention Centre, Pretoria, South Africa, August 2010.

[39] T. Hong, C. Rasmussen, T. Chang, and M. Shneier. Road detection and tracking for autonomous mobile robots. In *Proceedings of the SPIE 16th Annual International Symposium on Aerospace/Defence, Sensing, Simulation and Controls, Orlando, FL*, pages 1–9. Intelligent Systems Division, April 2002.

[40] C. Hsing-Chung, W. Hua, R. Chris, and M. Tony. Radar interferometry for 3-D mining deformation monitoring. In *Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium*, pages 2090 – 2092. IGARSS, July 2007.

[41] C. Hsu, F. Lian, C. Huang, and Y. Chang. Detecting drivable space in traffic scene understanding. In *Proceedings of the IEEE International Conference on Systems Science and Engineering*, pages 79–84, June 2012.

[42] Industrial Development Corporation (IDC). SA mineral industry: A strategic overview of the South African mining industry. *Minerals and Energy*, C08 0031 9059(1):1–80, September 2011.

[43] Istock. Istock photos. Available online: http://www.istockphoto.com (accessed on 10 May 2012).

[44] P. Jan, M. Jun, and T. Russ. Robot learning. *IEEE Robotics and Automation Magazine, Technical Committee Spotlight*, 2002.

University of Cape Town

[45] E. T. Jaynes. A mathematical theory of communication. *The Bell System Technical Journal*, 106(4):620–630, 1957.

[46] W. Jian, J. Zhong, and S. Yu-Ting. Unstructured road detection using hybrid features. In *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding*, pages 482–486. ISBN: 978-1-4244-3703-0/09, 2009.

[47] T. Joaquin and N. Patricio. A new approach to visual-based sensory system for navigation into orange groves. *Journal of Sensors*, 11(4):4086–4103, April 2011.

[48] S. John and F. Stephen. Bayesian maximum entropy image reconstruction. *Lecture Notes-Monograph Series, Spatial Statistics and Imaging*, 20:341–367, 1991.

[49] B. Kartikeyan, A. Sarkar, and K. L. Majumder. A segmentation approach to classification of remote sensing imagery. *International Journal of Remote Sensing*, 19(1):1695–1709, 1998.

[50] K. Khoshelham and S. Oude Elberink. Accuracy and resolution of Kinect depth data for indoor mapping applications. *MDPI Open Access Sensors*, 12(ISSN 1424-8220):1437–1454, February 2012.

[51] N. J. Lavigne, J. A. Marshall, and U. Artan. Towards underground mine drift mapping with RFID. In *Proceedings of the 23rd IEEE Canadian Conference on Electrical and Computer Engineering,*, pages 1–6. IEEE, May 2010.

[52] W. le Roux. When is a workplace safe or unsafe?: The safety criterion in terms of the occupational health and safety act and the mine health and safety act. *Journal of the Southern African Institute of Mining and Metallurgy*, 111(1):1–5, 2011.

[53] J. P. Leger. Trends and causes of fatalities in South African mines. *Journal of Safety Science*, 14(3-4):169–185, 1991.

[54] H. Li, H. Gu, Y. Han, and J. Yang. An efficient multiscale SRMMHR (statistical region merging and minimum heterogeneity rule) segmentation method for high resolution remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(2):67–73, June 2009.

University of Cape Town

[55] L. Maohai, S. Lining, H. Qingcheng, C. Zesu, and P. Songhao. Robust omnidirectional vision based mobile robot hierarchical localization and autonomous navigation. *Information Technology Journal*, 10(1):29–39, 2011.

[56] Megalife. Megalife resources. Available online: http://megalife.com.ua (accessed on 12 May 2012).

[57] N. Merhav and V. Bhaskaran. Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation. *IEEE Transactions on Circuits and System for Video Technology*, 7(3):468–476, 1997.

[58] Iron Miners. Gold mining and historical documentation. Available online: http://www.ironminers.com (accessed on 15 May 2012).

[59] J. Minguez and L. Montano. Sensor-based robot motion generation in unknown, dynamic and troublesome scenarios. *Journal of Robotics and Autonomous Systems*, 52(4):290–311, July 2005.

[60] J. Minguez and L. Montano. Extending collision avoidance methods to consider the vehicle shape, the kinematics, and dynamics of a mobile robot. *IEEE Transaction on Robotics*, 25(2):367–381, April 2009.

[61] A. Miranda, N. Neto, and L. Rittner. A simple and efficient road detection algorithm for real time autonomous navigation based on monocular vision. In *Proceedings of the IEEE 3rd Latin American Robotics Symposium LARS '06*, pages 92–99. 1-4244-0537-8/06, August 2006.

[62] V. Narunas and B. Andreas. Towards pathplanning for unmanned ground vehicles (UGV) in 3D plane-maps of unstructured environments. *Kùnstl Intell (Springer-Verlag)*, 25(DOI 10.1007/s13218-011-0098-7):141–144, February 2011.

[63] F. Neuhaus, D. Dillenberger, J. Pellenz, and D. Paulus. Terrain drivability analysis in 3D laser range data for autonomous robot navigation in unstructured environments. In *Proceedings of the IEEE international conference on Emerging Technologies Factory Automation*, pages 1–4. Active Vision Group, University of Koblenz-Landau, Germany, September 2009.

University of Cape Town

[64] R. Nock and F. Nielsen. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1–7, November 2004.

[65] I.O. Osunmakinde and T. Ndhlovu. Ground plane detection for autonomous robots in complex environments inclined with flexed far-field terrains. In *Proceedings of the 14th IASTED International conference on robotics and applications (RA 2009)*, pages 664–095. Cambridge, Massachusetts, USA, ACTA Press ISBN: 978-0-88986-813-7, November 2009.

[66] N. Otsu. A threshold selection method from gray-level histogram. *IEEE Transactions on Systems, Man, and Cybernetics*, 9:62–66, 1978.

[67] J. Peters, J. Morimoto, and R. Tedrake. Robot learning. *IEEE Robotics and Automation Magazine: Technical Committee Spotlight September*, 2009.

[68] T. Pingbo, D. Huber, and B. Akinci. A comparative analysis of depth-discontinuity and mixed-pixel detection algorithms. In *Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling, 3DIM '07.*, pages 29–38, August 2007.

[69] PricewaterhouseCoopers("PwC"). South African Mine : Review of trends in the South African mining industry. *SA Mine*, 6(1):1–48, November 2011.

[70] Mining H. S. and E. Mining accidents in April: South African April mine fatalities, 2006. http://www.klasslooch.com (accessed on 20 March 2012).

[71] SA. Mining in South Africa. Available online: http://www.esnet.oneworld.net (accessed on 10 May 2011).

[72] J. Sachs. *Digital Image Basics*. Digital Light and Color, Cambridge Massachusetts, 1999.

[73] S. A. Salem, N.V. Kalyankar, and S.D. Khamitkar. Image segmentation using edge detection. *International Journal on Computer Science and Engineering (IJCSE) ISSN : 0975-3397*, 02(03):804–807, June 2010.

[74] P. D. Sathya. Image segmentation using minimum cross entropy and bacterial foraging optimization algorithm. In *IEEE International Conference on Emerging Trends in Electrical and Computer Technology*, pages 500–506. ICETECT, March 2011.

University of Cape Town

[75] S. Scheding, G. Dissanayake, E. Nebot, and H. Durrant-Whyte. An experiment in autonomous navigation of an underground mining vehicle. *IEEE Transactions on Robotics and Automation*, 15(1):85–95, 1999.

[76] N. Senthilkumaran and R. Rajesh. Edge detection techniques for image segmentation - a survey of soft computing approaches. *International Journal of Recent Trends in Engineering*, 1(2), 2009.

[77] Y. Seung-Nam, J. Jae-Ho, and H. Chang-Soo. Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel. *Journal of Automation in Construction (ELSEVIER)*, 16(1):255–261, May 2007.

[78] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27, 1948.

[79] Y. Shin, C. Jung, and W. Chung. Drivable road region detection using laser range finder for outdoor patrol robots. In *IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA*, pages 877–882. ISBN: 978-1-4244-7868-2/10, June 2010.

[80] H. J. Simons. Deaths in South African mines. *Chamber of Mines report of governmnet mining*, 1(1):41–55, 1960.

[81] W. Slawo and J. Ed. Colour edge detection in RGB using jointly euclidean distance and vector angle. *Journal of Vision Interface*, 99:19–21, 1999.

[82] P. Soille. *Morphological Image Analysis: Principles and Applications.* Springer-Verlag New York, Inc., 2003.

[83] SouthAfrica-info. Mining and minerals in South Africa. http://www.southafrica.info (accessed on 15 April 2012).

[84] R. Spowart. Mine rescue at durban deep in joburg suspended. Available online: http://www.salabournews.co.za (accessed on 20 May 2012).

University of Cape Town

[85] S. Stacy, G. Laxmi, A. Esa, and W. Jack. Creation of robot for subsurface void detection. In *Proceedings of the IEEE International Conference on Technologies for Homeland Security*, pages 669–676. HST, May 2009.

[86] S. R. Teleka, J. J. Green, S. Brink, J. Sheer, and K. Hlophe. The automation of the 'making safe' process in South African hard-rock underground mines. *International Journal of Engineering and Advanced Technology*, 1(4):1–7, April 2012.

[87] S. Teppei and K. Yoji. Online road surface analysis using laser remission in structured outdoor environments. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 15–21. IROS, October 2010.

[88] S. Thrun, D. Hahnel, D. Ferguson, M. Montemerlo, R. Triebel, W. Burgard, C. Baker, Z. Omohundro, S. Thayer, and W. Whittaker. A system for volumetric robotic mapping of abandoned mines. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 4270–4275. ICRA Taipei, Taiwan, September 2003.

[89] M. Trentacoste, R. Mantiuk, and W. Heidrich. Blur-aware image downsampling. *Journal of Eurographics*, 30(2):1–10, 2011.

[90] Y. Young Uk and O. Se-Young. Three-feature based automatic lane detection algorithm (TFALDA) for autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 4(4):219–225, 2003.

[91] P. Valarmathi and M. Aruna. A novel approach for image segmentation using SOM. *International Journal of Computer Science and Technology*, 2(1):84–85, December 2011.

[92] J.N. van der Merwe. Future of the South African mining industry and the roles of the SAIMM and the universities. *Journal of the Southern African Institute of Mining and Metallurgy*, 111(9):581–592, 2011. ISSN 2225-6253.

[93] S. Wicks. Separating people and risk. Submitted as partial requirement for the course PSC 410, University of Pretoria, 2008.

University of Cape Town

[94] S. Young-Woo, C. Urmson, D. Wettergreen, and L. Jin-Woo. Building lane-graphs for autonomous parking. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan*, pages 6052–6057. 978-1-4244-6676-4/10, October 2010.

[95] S. Zhou, J. Gong, G. Xiong, H. Chen, and K. Iagnemma. Road detection using support vector machine based on online learning and evaluation. In *IEEE Intelligent Vehicles Symposium*, pages 256 – 262. San Diego USA, 2010.

University of Cape Town