Autonomous Robots' Visual Perception in Underground Terrains using Statistical Region Merging

Omowunmi Isafiade, Isaac Osunmakinde, *Member, IEEE,* and Antoine Bagula, *Member, IEEE,*

Abstract—Robots' visual perception is a field that 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. In the history of mining, the mortality rate of mine workers has been alarming and robots exhibit a great deal of potentials to tackle safety issues in mines. However, an effective vision system is crucial to safe autonomous navigation in underground terrains. This work investigates robots' perception in underground terrains (mines and tunnels) using statistical region merging (SRM) model. SRM reconstructs the main structural components of an imagery by a simple but effective statistical analysis. 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 dynamic underground terrains.

Keywords—Drivable Region Detection, Kinect Sensor, Robots' Perception, SRM, Underground Terrains.

I. Introduction

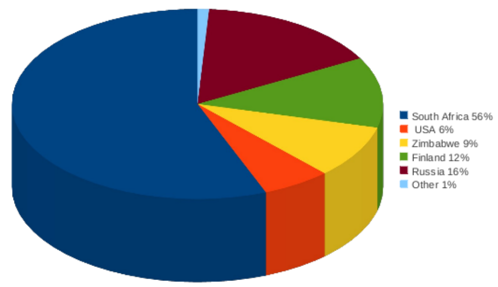
THE mining industry is a world-class one and a cornerstone of the South African economy [2]. According to Figure 1, South Africa supplies over 80% of the world's platinum group metals (PGMs) need. The industry makes a significant contribution to the overall economic activity [3] which includes, but is 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) [4]. Thus mining remains crucial for the country's economy.

While these benefits are derived from the mining industry, many risks are associated with the business of the extraction of an ore body[5] and the environment tend to degrade fairly rapidly. According to a report in 2011 [3], fatalities were still unacceptably high, in line with the trend experienced in the recent past.

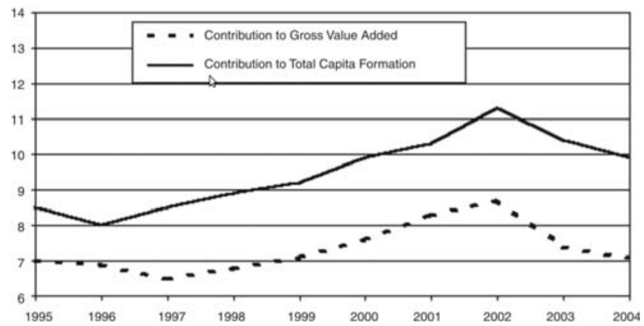
O. Isafiade and A. Bagula are with the Department of Computer Science, University of CapeTown, South Africa, e-mail: wunmi.isafiade@uct.ac.za, bagula@cs.uct.ac.za

I. Osunmakinde is with the School of Computing, College of Science, Engineering and Technology, University of South Africa, South Africa, e-mail: osunmio@unisa.ac.za

Manuscript received November 30, 2012; revised February 21, 2013.



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



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

Fig. 1. South African Mining's Contribution to the Economy (Department of Minerals and Energy)

In the quest to address safety issues in mines, it is widely recognised that autonomous robots could play a key role. Robots can be used for checkpoint and safety inspection tasks [22] in a mine. However, an effective vision model (robots' perception) is critical for safe autonomous navigation within an underground terrain.

A. Robots' Perception

A good perception of robots is part of the critical ingredients that often formulates safe autonomous navigation. How the robot perceives and interpret its immediate environment is crucial [7]. However, several research efforts have been directed towards autonomous navigation in underground environments and research continues in this area. The use of sensor fusion in

3D visualisation of underground terrains is currently gaining much attention from researchers [6]. Figure 2 depicts a robot navigating a mine. This research focuses on the perception module (which comprises sensors with high-quality visual capabilities), a critical component in autonomous navigation, while 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.

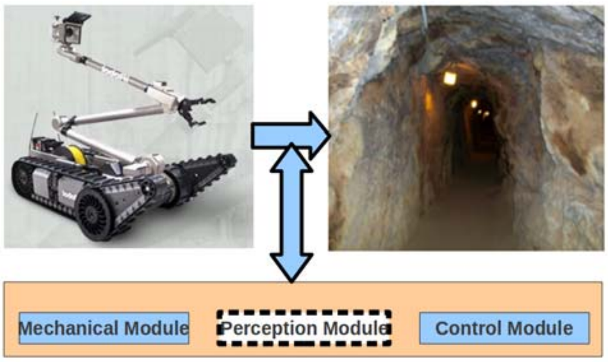


Fig. 2. Overview of Autonomous Robot Navigation in a Mine

B. Contributions and Outline

Effective and safe autonomous navigation of a robot is subject to its ability to perceive and interpret its immediate environment adequately. However, drivability analysis of underground environments, with visualisation results in 3 dimensional view is still an ongoing research. In this work, we conducted an experiment using publicly available underground mine images with images captured in a local mine and on a stream of dynamic and rough underground tunnel images captured with the aid of the 3D XBOX kinect sensor device. Using the SRM model together with data fusion from the two kinect (laser and infrared) sensors, we obtained promising results where drivable region and non-drivable region are clearly distinguished. Evaluation is also carried out for useful qualitative and quantitative conclusions and future adoption. The major goals of this paper are as follows:

1) Modelling and application of the SRM in drivability analysis of underground terrains using publicly available mine images with a stream of images captured locally in an underground mine. Experiment is also conducted on a stream of dynamic underground tunnel images captured with an XBOX 3D Kinect sensor.
2) The augmentation of 2D results with 3D drivability maps for autonomous robots, which would need to climb steps in the mines, and benchmarking the proposed model with related methods and images.

To the best of the researchers knowledge, though being widely used in computer vision, the SRM model has never been applied to an underground environment.

The rest of the paper is organised as follows: Section II presents some related work. In section III, the methodology and framework is presented in detail. Section IV follows with experimental results and a review of the outcome measures for qualitative and quantitative performance evaluation. Section V concludes the paper and future work are also presented.

II. Background Studies

A. Related Work

The problem of improving the vision of robots for autonomous navigation has gained significant attention over the years [1]. Notably is the DARPA grand challenge [7] which is intended to spur innovation in unmanned ground vehicle navigation. The goal of the challenge was to develop an autonomous robot capable of traversing unrehearsed off-road terrain. Several approaches have been adopted to address the aforementioned problem, most of which are domain-specific. However, research on autonomous navigation in underground environments continues. Underground mines, which present unique and terrain-specific human hazards, still call for the serious attention of researchers.

Joaquin *et al.* [8] 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 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 that could assist autonomous navigation.

Derek *et al.* [9] 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 aim at transforming 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.* [20] 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.* [21], the SVM is defined as a technique motivated by the statistical learning theory, which has shown its 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.

From the literature, it is observed that much research on improving robots' vision in different scenarios/environment has been conducted. However, Underground terrains have received little attention, probably owing to their roughness and environmental/technological constraints, compared to structured and unstructured surface terrains.

In a recent study, Teleka *et al.* [22] investigated the "Automation of the 'Making Safe' Process in South African Hard-Rock Underground Mines". The work was motivated by the fact that 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 mines. The goal is to reduce miners' exposure to the dangerous underground mine environment and consequently improve productivity. However, robots need to overcome specific vision tasks in order to carry out their operations more effectively. Hence, an effective vision model is critical to the success of robots in mines.

While autonomous navigation in an underground mine environment has been studied for more than twelve years [10], a robust algorithm that is applicable in different terrains is yet to emerge. Thus, it remains an ongoing key challenge. Several segmentation algorithms exist, such as the k-means, entropy and edge detection based methods, but in this research a statistical approach based on the SRM model is adopted. Our choice of the SRM model is prompted by the existence of interesting statistical properties, such as separability and homogeneity, in the model. The aforementioned properties appear promising for the segmentation task in this research.

B. Statistical Region Merging (SRM) Model

A region 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 to interpret parts of the image as objects by transforming the pixels into visually meaningful partition of regions and object. The object is everything 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 [13]:

a. $H_p(R_g) = true \; \forall g$
b. $H_p(R_g \cup R_h) = false \; \forall \; adjacent(R_g, R_h)$
c. $\bigcup_{g=1}^{G} R_g = I \; with \; g \neq h \; and \; R_g \cap R_h = \emptyset$

Statistical region merging (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 [14], [15], [16], [17]. Richard *et al.* [14] present an elaborate theoretical analysis of the SRM algorithm in order to analyse the underlying principles. SRM is applied to skin imaging technology in [16] 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 regions or pixels. SRM uses a union-find data structure or merge-find set that 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 an accuracy in segmentation close to optimum [14]. 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 (1) with an extension in Equations (2) and (3).

$$P(R, R') = \begin{cases} \text{true} & \textit{if} \quad \forall c \in \text{(R,G,B)}, |\bar{R}'_c - \bar{R}_c| \leq T \\ \text{false} & \text{otherwise} \end{cases} \tag{1}$$

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

$$k(R) = g \sqrt{\frac{1}{2Q\,|R|} \ln(6|I|^2 R_{|R|})}. \tag{3}$$

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 $\{0, 1, \cdots, g - 1 \text{ pixels}\}$ where $g = 256$. The observed image I' is generated by sampling each statistical pixel for the three RGB channels. 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 [17]. 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 differ in at least one colour channel

Equation (4) defines the sort function [14], where p'_a, p_a represent pixel values of a pair of adjacent pixels of the colour channel.

$$f(p, p') = \max_{a \in R,G,B} |p_a - p'_a|. \tag{4}$$

III. THE PROPOSED METHODOLOGY

A. Proposed System Model for Drivability Analysis

This work conducts a drivability analysis of underground terrains for autonomous robots by devising means through which drivable regions can be identified in an underground terrain. This consequently improves the vision of a robot, allowing it navigate on only drivable regions in a mine frame, thereby minimising accidents while executing its tasks. Figure 3 depicts the proposed system model.

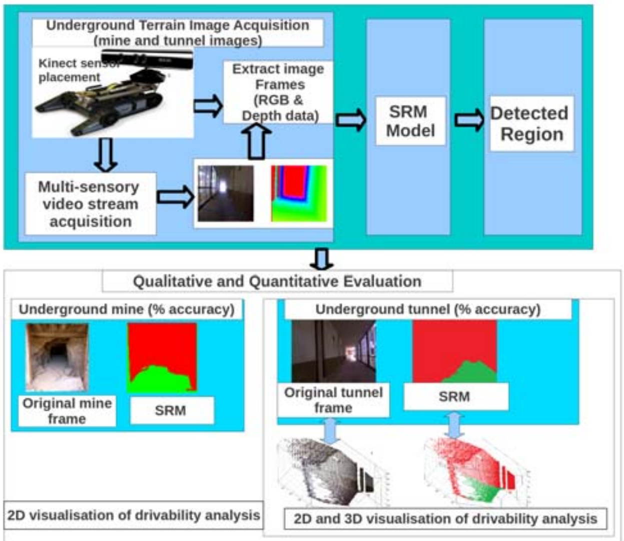


Fig. 3. Schematic of the Proposed System Model (Perception Module)

B. Underground Terrain Image Acquisition

In this work, the tested mine frames combine a selection of publicly available mine frames with a stream of mine frames captured in a local mine. The underground tunnel image frames are captured with the 3D XBOX kinect sensor device as shown at the top left of Figure 5. Figure 4 shows the general layout of the capturing cycle.

1) Operation of the 3D Kinect Motion Sensing Device: The 3D kinect sensor device, with the XBOX 360 console, consists of two major sensors, which are the RGB sensor and the Depth sensor. The RGB sensor produces RGB images while the depth sensor produces the corresponding depth information/images as shown in Figure 5. The depth sensor consists of the infrared laser projector and an infrared sensor. The laser projector projects the data while the infrared sensor calculates the time taken for the laser rays to hit the target environment. The experimental setup is moved along the tunnel pathway for image capture as perceived by an autonomous robot. To capture underground surface environment, the sensor is mounted with a small tilt angle θ.

The XBOX kinect sensor produces, simultaneously, depth and RGB images of 640 x 480 resolution at 30 frames per second (fps). The depth data from the device calculates, in millimeter (mm), the distance of each pixel's location relative to the sensor device. We might also get unknown depth pixels especially if the rays from the sensor are hitting a shadow, window etc, the depth data returns zero under such situation. Furthermore, unknown depths may be parlty due to the limitation in the precision or accuracy of the depth sensor. The depth images indicate how far or near each pixel's region is perceived by the sensor in the target underground environment. The fusion of the multi-sensory data (RGB & Depth) enhances our knowledge of the image structure and allows our system to obtain an additional information about the vanishing points which might have occured as a result of the perspective effect [11].

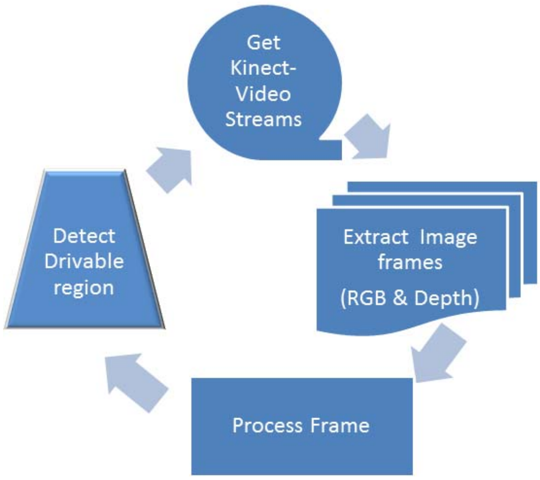


Fig. 4. Overview of the Kinect Frames Capturing Cycle.

Fig. 5. Robot Sensor Data Capturing Platform

C. SRM Approach to Drivability Analysis

The SRM algorithm has two important criteria: the merging predicate and specified cluster Q, which determines the number of segments/regions, for the input image. SRM is noted for its computational efficiency, simplicity and good performance as seen in Section IV-A. 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 our experiment, after testing with different values of Q, the value $Q = 32$ gave the optimal result for the image classification. Figure 7 presents the segmentation results of a mine frame at different Q levels. Q is a parameter that controls the coarseness and busyness of the classification.

The algorithm uses a 4-connectivity scheme to determine adjacent pixels relative to the center pixel (in green) as shown in Figure 6. The pixels are sorted in ascending order based on the sort function in Equation (4). 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. If the regions of the pixels differ and the mean intensity are sufficiently similar enough to be merged, then the two regions are merged.

The SRM method presents the list of pixels belonging to each segmented region with their average mean intensities. We focus on the pixels region which forms clusters at the base of each observed image I towards the midpoint when scanning from the left. This forms the pixel region closer to the robots view and thus, the drivable part as can be seen in the test cases presented.

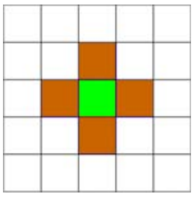


Fig. 6. Depiction of the Four-Connectivity Scheme

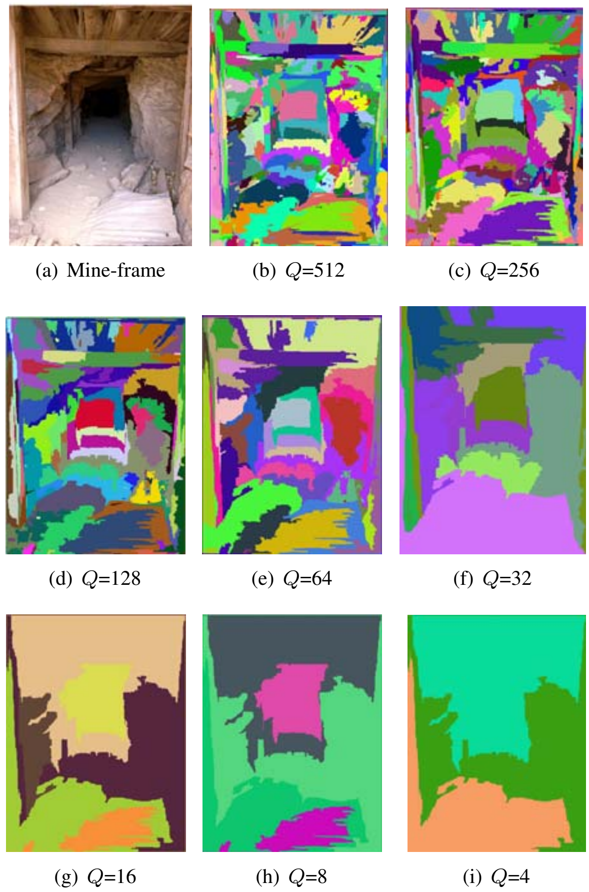
Pseudocode for SRM Algorithm

Step 1: Initialise image I and estimated segments 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 $(R_{(p'_i)}, R_{(p_i)})$

D. Scoring and Evaluation Scheme

Qualitative and quantitative (confusion matrix) evaluation approaches are considered. Both are used as a measure of performance of the method described in this work. The qualitative evaluation which is the visual representation is presented in Figures 8, 9 and 10. The quantitative evaluation is presented in Table I. In this work, we considered the confusion matrix validation technique [12]. We repeated the confusion matrix procedure n times, with $n \in \{3, 5\}$, where each n subsamples are used exactly once as the validation data. The idea is to evaluate the accuracy level (hit rate) of the algorithm in the following context.

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



(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

Fig. 7. Stages of Region Merging/Segmentation on a Mine Frame at Different Q Levels.

Thus, the accuracy (acc %) is given as;

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \tag{5}$$

IV. Experimental Evaluation

A. Experiment 1: 2D Qualitative Observations on Underground Mine Terrains

The major focus in this work is feature extraction and classification for front view drivable region detection in order to enhance the visual capability of autonomous robots in underground terrains. For the images experimented on in this work, different test cases of mine frames were carefully chosen from publicly available mine frames [18] 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.

Figure 8 presents the qualitative detection results on some mine frames. The first row presents the original mine frames. The second row presents the results of the clusters generated for regions with homogeneity. The third row presents 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 clear 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. One can see that pixel regions closer to the robot's view tend to form most of the drivable region.

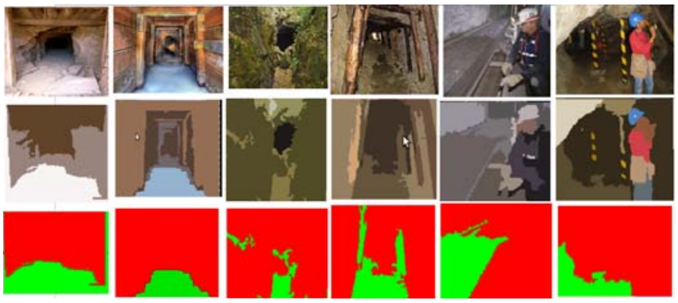


Fig. 8. 2D Qualitative Results of SRM Method on Mine Frames

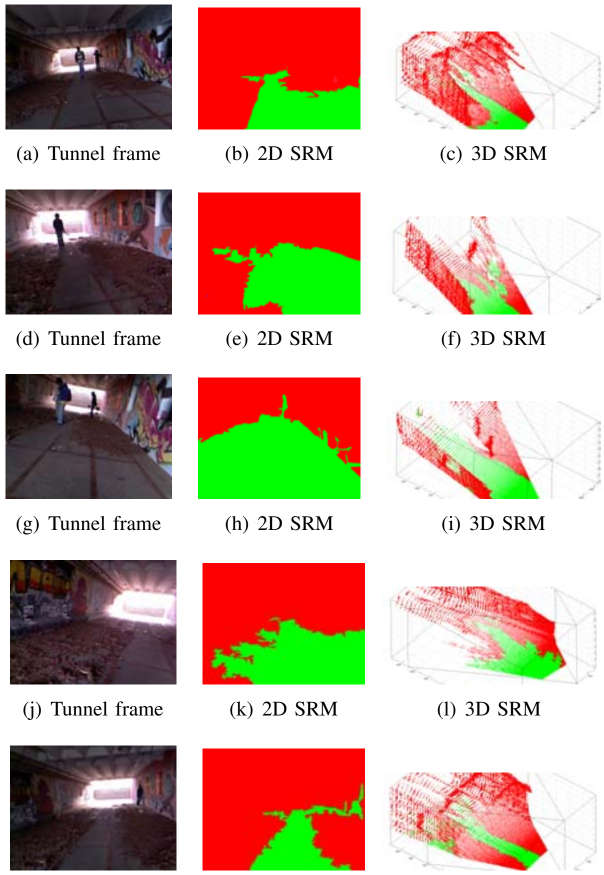
B. Experiment 2: 2D and 3D Qualitative Observations on Underground Tunnel Terrains

The experiment is conducted on a stream of underground tunnel images captured with the aid of an XBOX 3D Kinect sensor. The underground images comprises rough and dynamic tunnel images as presented in Figures 9 and 10 respectively. For space management, only a few frames are presented. Figures 9 and 10 show the 2D and 3D qualitative results of rough and dynamic tunnel images using the SRM model.

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. 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, one can see that Figures 9 and 10 creates an accurate understanding of where an autonomous robot should navigate in real time.

C. Experiment 3: Quantitative Evaluation of Ground Truth Obtained for the Underground Terrains

In this section, we present the quantitative results of the two terrains (mine and tunnel), based on the applied algorithm. We conducted an experiment to evaluate the quantitative performance of the SRM approach to drivability. We utilised the confusion matrix validation process n times ($n \in \{3, 5\}$). In the experiment conducted, we randomly hand-labelled pixel positions 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 classification divided by number



(a) Tunnel frame (b) 2D SRM (c) 3D SRM

(d) Tunnel frame (e) 2D SRM (f) 3D SRM

(g) Tunnel frame (h) 2D SRM (i) 3D SRM

(j) Tunnel frame (k) 2D SRM (l) 3D SRM

(m) Tunnel frame (n) 2D SRM (o) 3D SRM

Fig. 9. 2D and 3D Qualitative SRM Results on Rough Tunnel Frames

of instance in the image-data I_d. Table I shows the quantitative performance of the SRM algorithm on Mine and tunnel frames with $n = 3$ and $n = 5$. One can see that the detection accuracy is reliable with over 80% in all cases.

D. Experiment 4: Benchmarking the proposed model with publicly available images and methods

To validate the performance of our detection algorithm, publicly available images and methods are used as a benchmark [9], [21]. Figure 11 shows the input images (tunnel frames and ground based unstructured road frames) and their corresponding detection results using the SRM method. Table II shows the quantitative result of the validation with SRM detection accuracy of over 80%. It is worth mentioning that our results for the input images, using the SRM algorithm, provide an alternate method for drivable region detection on the frames. Furthermore, there are no 3D results for the tested frames as we did not access to depth maps of the images which provide critical information for the 3D image visualisation.

V. Summary and Conclusion

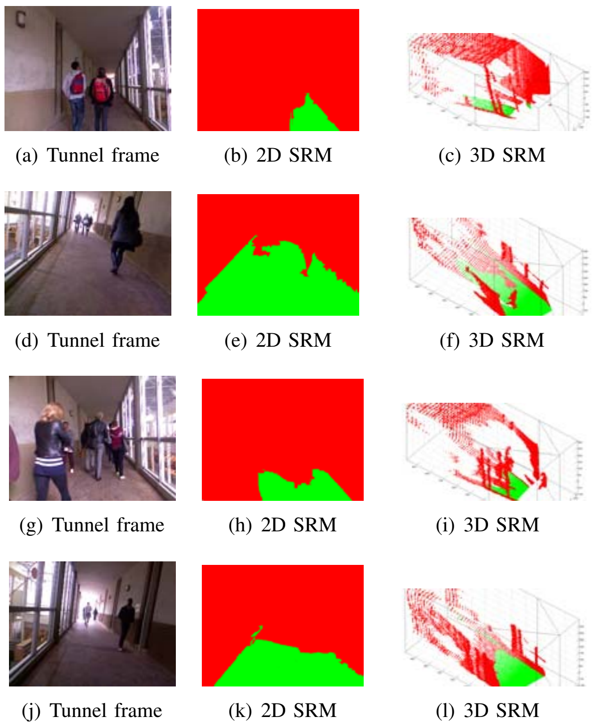
This work has demonstrated the feasibility of enhancing robots' capability of identifying drivable regions in underground terrains. SRM algorithm is investigated as a means of

TABLE I
QUANTITATIVE RESULTS OF UNDERGROUND TERRAINS USING THE SRM ALGORITHM

Algorithm	Terrain	n-fold confusion matrix validation	Correctly classified pixels (TP, TN)	Incorrectly classified pixels (FP, FN)	Accuracy of detection (%)
SRM	Underground mines (30 image frames)	3	745	155	82.78
		5	1200	300	80.00
	Underground tunnels (100 image frames)	3	2750	250	91.67
		5	4480	520	89.60

TABLE II
QUANTITATIVE RESULTS OF VALIDATING SRM WITH EXISTING APPROACHES

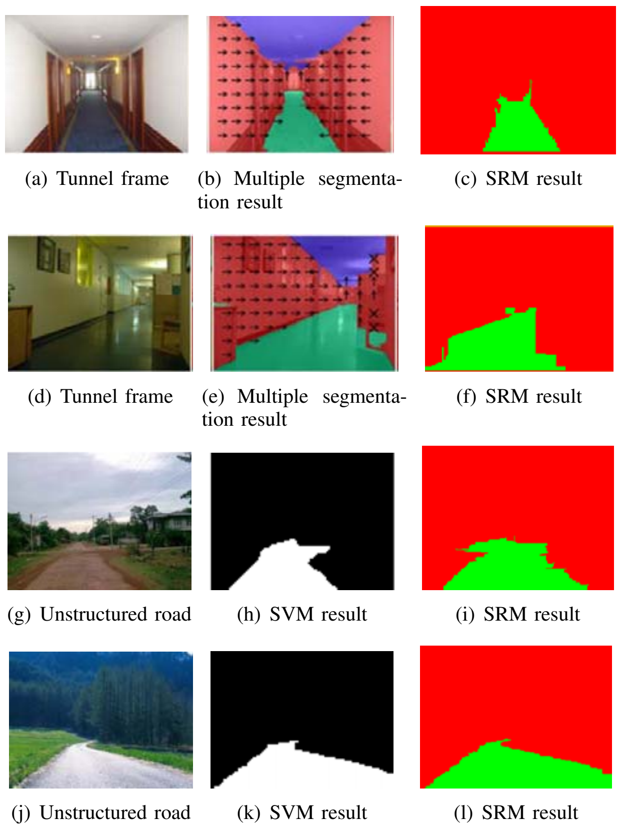
Terrain	Algorithm	n-fold validation	Right classification	Wrong classification	Accuracy (%)
Publicly Available Image Frames	SVM	3	96	24	80.00
		5	170	30	85.00
	SRM	3	97	23	80.83
		5	175	25	87.50
	Multiple Segmentation	3	99	21	82.50
		5	177	23	88.50



(a) Tunnel frame (b) 2D SRM (c) 3D SRM
(d) Tunnel frame (e) 2D SRM (f) 3D SRM
(g) Tunnel frame (h) 2D SRM (i) 3D SRM
(j) Tunnel frame (k) 2D SRM (l) 3D SRM

Fig. 10. 2D and 3D Qualitative SRM Results on Tunnel Frames

identifying drivable regions in underground terrains because it shows promise in its statistical feature property. Different regions of the mines representing a wide variety of terrains ranging from the stope, shaft and gallery were investigated. We also conducted an experiment on a stream of underground tunnel image frames captured using the XBOX Kinect sensor and further benchmarked our approach with publicly available images and methods.



(a) Tunnel frame (b) Multiple segmentation result (c) SRM result
(d) Tunnel frame (e) Multiple segmentation result (f) SRM result
(g) Unstructured road (h) SVM result (i) SRM result
(j) Unstructured road (k) SVM result (l) SRM result

Fig. 11. Validating our Approach with Publicly Available Image Frames and Methods

The SRM algorithm is able to re-construct the main structural components of the underground mine imagery by a simple but effective statistiscal analysis. The SRM method worked well on a variety of mine and tunnel frames tested as shown in Figures 8, 9 and 10. The detection accuracy of the SRM approach is reliable with over 80% accuracy as shown in Tables I and II. The 3D perception representation also reveals the level of correctness and clarity of the image classification as shown in Figures 9 and 10.

The major focus in this work is feature extraction and classification for front view drivable region detection in 2D by augmenting with 3D results. This would enhance autonomous robots' visual capability to identify drivable regions in underground environments.

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 environments. However, the current classification can still be improved upon by utilising more machine learning algorithms and more sophisticated cameras, for example a laser scanner, for better performance and future adoption.

Acknowledgment

The authors gratefully acknowledge resources made available by the University of CapeTown (UCT) and University of South Africa (UNISA), South Africa.

References

[1] S. Chen and Y. Li and N. Ming, *Active Vision in Robotic Systems: A Survey of Recent Developments*, The International Journal of Robotics Research, AID: 0278364911410755, Vol. 30, No. 11 ,pp. 1343 1377, September, 2011.

[2] A. Damarupurshad, *African (Mining-Related) Government Departments, Mineral/Mining-Related Organisations and Mining Companies with Interests in Africa*, Journal of Minerals and Energy (Republic of South Africa), Vol. D13, No. 1 ,pp. 1-58, June, 2004.

[3] PricewaterhouseCoopers. *SA Mine: Review of Trends in the South African Mining Industry*, SA Mine, Vol. 6 No. 1, pp. 1-48, November 2011.

[4] C.J.H. Hartnady, *South Africa's Gold Production and Reserves*, South African Journal of Science, Vol. 105, No. 1 ,pp. 328-330, September, 2009.

[5] M. A. Hermanus, *Occupational Health and Safety in Mining- Status, New Developments and Concerns*, Journal of the Southern African Institute of Mining and Metallurgy, Vol. 107, No. 1 ,pp. 531-538, May, 2007.

[6] J. J. Green and K. Hlophe and J. Dickens and R. Teleka and M. Price, *Mining Robotics Sensors*, International Journal of Engineering and Advanced Technology (IJEAT), Vol. 1, No. 4 ,pp. 8-15, April, 2012.

[7] J. Benjamin and Y. Papelis and R. Pillat and G. Stein and D. Harper,*A Practical Approach to Robotic Design for the DARPA Urban Challenge*, International Journal of Field Robotics, Vol. 25, No. 8 ,pp. 528-566, July, 2008.

[8] T. Joaquin and N. Patricio, *A New Approach to Visual-Based Sensory System for Navigation into Orange Groves*, Journal of MDPI Open Access Sensors, *Vol. 11*, pp. 4086-4103, April 2011.

[9] H. Derek and A. Alexei and H. Martial. *Recovering Surface Layout from an Image*, International Journal of Computer Vision, *Vol. 75*, pp.151-172, 2007.

[10] S. Scheding and G. Dissanayake and E. Nebot and H. Durrant-Whyte, *An Experiment in Autonomous Navigation of an Underground Mining Vehicle*, IEEE Transactions on Robotics and Automation, Vol. 15, No. 1 ,pp. 85-95, 1999.

[11] K. Khoshelham and S. Oude Elberink, *Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications*, Journal of MDPI Open Access Sensors, Vol. 12, No. 1 ,pp. 1437-1454, February, 2012.

[12] T. Fawcett, *An Introduction to ROC analysis*, Pattern Recognition Letters, Vol. 27, No. 1, pp. 861-874, December, 2005.

[13] S. Battiato and G.M. Farinella and G. Puglisi. *SVG Vectorization by Statistical Region Merging*. Proceedings of 4th Conference Eurographics Italian Chapter, pages 1-7, July, 2006.

[14] R. Nock and F. Nielsen, *Statistical Region Merging*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 11 ,pp. 1-7, November, 2004.

[15] F. Calderero and F. Marques, *Region Merging Techniques Using Information Theory Statistical Measures*, IEEE Transactions on Image Processing, Vol. 19, No. 16 ,pp. 1567-1586, June, 2010.

[16] M.E. Celebi and H. A. Kingravi and J. Lee and W. Stoecker and J. M. Malters and H. Iyatomi and Y.A. Aslandogan and 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), No. 16 ,pp. 1-10, 2006.

[17] H. Li and H. Gu and 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, Vol. 2, No. 2 ,pp. 67-73, June, 2009.

[18] Bretterk. *Irominers*, Available online: http://brettberk.com/wp-content/uploads/2009/05/death-valley-looking-inside-mine-shaft.jpg (accessed on 10 May 2012).

[19] C. Hsu and F. Lian and C. Huang and Y. Chang. *Detecting Drivable Space in Traffic Scene Understanding*. Proceedings of the IEEE International Conference on Systems Science and Engineering, pp 79-84, June, 2012.

[20] S. Zhou and J. Gong and G. Xiong and H. Chen and K. Iagnemma. *Road Detection using Support Vector Machine Based on Online Learning and Evaluation.* In IEEE Intelligent Vehicles Symposium, pp 256-262., San Diego USA, 2010.

[21] W. Jian and J. zhong and S. Yu-Ting. *Unstructured Road Detection using Hybrid Features.* Proceedings of the Eighth International Conference on Machine Learning and Cybernetics,, pp. 482-486., July 2009.

[22] S.R. Teleka and J. J. Green and S. Brink and 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 (IJEAT) Vol. 1, pp. 1-7., April 2012.

Omowunmi Isafiade is a research associate and Ph.D. student at the University of Cape Town (UCT), South Africa. Past affiliations were with the African Institute for Mathematical Sciences (AIMS), Cape Town, South Africa, where she obtained a post-graduate diploma (PGDs) degree in mathematical sciences in June 2010, and a Bachelor (BSc.) honors degree in computer science from Federal University of Agriculture Abeokuta (FUNAAB), Nigeria. She received an M.Sc. degree in computer science from University of CapeTown (UCT) in December 2012 and her research interest lies in image processing & computer vision, data mining & knowledge management and ubiquitous intelligence.

Isaac Osunmakinde received his M.Sc. and Ph.D. in Computer Science in 2006 and 2009 respectively from the University of Cape Town (UCT), South Africa. He was a Senior Research Scientist until the end of 2011 at the Council for Scientific and Industrial Research (CSIR). Currently, he is an Associate Professor at the University of South Africa. Professor Osunmakinde has authored prestigious articles in peer-reviewed journals, book chapters and conference publications, and his research interest includes Machine Learning, Robotics, Computer Vision, etc. He is a member of the IEEE Computational Intelligence (CI) society and distinguished speaker of the IASTED CIB 2011, Pittsburgh, USA.

Antoine Bagula received an M.Sc. degree in Computer Engineering from Catholic University of Louvain (UCL) in Belgium and a M.Sc. Degree in computer science from the University of Stellenbosch in South Africa. He obtained a Ph.D in communication systems from the KTH-Royal Institute of Technology in Sweden. The focus of his research is on the design, analysis and control of intelligent telecommunication systems and their integration into the emerging ubiquitous networks. Besides his strong involvement in research at the academic and industry levels, Dr Bagula is currently involved in training activities in developing countries as an ICTP, ISOC and WirelessU trainer/consultant.