

Safety through Perception: Multi-modal Traversability Analysis in rough outdoor environments

M. Breitfuß M. Schöberl J. Fottner

Technical University of Munich, Germany

{matthias.breitfuss, max.schoeberl, j.fottner}@tum.de

Abstract: One of the most crucial functionalities of mobile indoor and outdoor robots is the ability to understand the local environment in a reasonable manner. Based on an intelligent perception of the systems' surroundings, a secure input for further autonomous functionalities like path planning or navigation can be guaranteed. Especially in the domain of off-road vehicles and machinery difficulties concerning a meaningful comprehension of the environment increase, as within this field of application the surroundings are rough and unstructured, implicating unpredictable situations for perception algorithms. Therefore, procedures to analyse the traversability of off-road terrain have to be both, robust in order to ensure a reasonable understanding of the environment under different circumstances and flexible to be adaptable to unknown application areas. In this paper we propose a method satisfying these requirements. Our approach is based on an efficient way of combining geometric and visual features of the surroundings, leading to a real-time capable traversability analysis for autonomous off-road vehicles. The evaluation of the developed system has been conducted in virtual tests within a simulation as well as in real-world situations on a prototypical vehicle.

Copyright © 2021 The Authors. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0>)

Keywords: autonomous systems, safe environmental perception, traversability analysis, semantic segmentation, decision making

1. INTRODUCTION

Analysing the *traversability* of an environment is an essential functionality of every autonomous robot. In fact, the ability to acquire knowledge about the surroundings and locate areas of drivable terrain within the gathered information is fundamental to guarantee an intended operation of such systems. In other words, a stable and meaningful environmental perception is one of the key features concerning safety of unmanned ground vehicles (UGVs). In this context, the term *safety* relates to the prevention of vehicle damage through the avoidance of improper situations like tilting or stalling. In general these issues apply to UGVs from small floor-borne vehicles in automated warehouses to self-operating machines in the agricultural sector or the building industry.

With the aim of reducing the aforementioned source of danger, our work proposes a novel procedure to derive reasonable *traversability estimations* based on data acquired by state-of-the-art sensor systems like Stereo-Cameras and Light Detection and Ranging (LiDAR). Even though the analysis of surroundings with regard to the traversability has been an active field of research for many years, the majority of works like the one of Kim et al. (2006) deal with compact prototypical robots, whereas the application area of heavy vehicles and machinery has been addressed insufficiently. Hence, our work attends to the need for a universally valid and vehicle independent solution of a traversability analysis, resulting in safe and robust estimations in various areas of application.

The main contribution of our method is the consideration of traversability as a continuous value describing the quality of the terrain instead of simply differing between drivable and non-drivable ground in a binary manner. Therefore we build upon a multi-modal and real-time capable principle for traversability analysis. The method is accurate in terms of reducing hazards through erroneous environmental perception and adaptable to various UGVs through the integration of vehicle parameters. Thereby we guarantee meaningful analysis results specific to the system's capabilities. Our approach is based on the simultaneous execution of two complementary analysis steps:

- Examination of 3D point clouds with respect to geometric features of the terrain surface
- Semantic segmentation on image data to identify predominant ground types

To demonstrate the accuracy and robustness of our approach, the general usability of our system is investigated on a prototypical machine. Additionally, particularly challenging and hazardous situations are tested within a simulated environment for safety reasons.

This paper is structured as follows: In Section 2, representative research work regarding environmental perception and traversability analysis is discussed. Section 3 presents the overall structure of the proposed method and elaborates on the core components of our algorithm. Section 4 provides traversability assessments gained from virtual and real-world experiments. Finally, a conclusion of the paper is presented in Section 5.

2. RELATED WORK

The challenge of traversability analysis in different environments is actively researched in the field mobile autonomous systems. Regarding the subclass of indoor vehicles, surroundings are usually made up of a flat floor bounded by walls that potentially has critical obstacles distributed on it. As a result, the main perceptual task is to localise areas showing significant risings in comparison to the flat ground. As the problem statement for indoor applications is restricted to a two-valued differentiation between drivable ground and untraversable objects, it is sufficient for most cases to use a binary occupancy grid as proposed by Elfes (1989).

When dealing with autonomous systems in unstructured and rough outdoor surroundings, the last-mentioned technique quickly reaches its limits. Especially the occurrence of environments with a considerably higher complexity attended by uncertainty in identifying an unambiguous ground plane justify the need for more accurate approaches. In this context the survey of Papadakis (2013) provides a good overview of possible solutions. The author especially emphasises, that a meaningful environmental perception is build on an appropriate set-up of sensors that gather information about the surroundings, and a suitable form of corresponding spacial representations. Furthermore the research shows that mainly *exteroceptive* perception systems like Stereo-Cameras and Time-of-flight sensors come into operation. In this context, point clouds, which are accumulations of 3-dimensional points, are preferred for environmental description. Basically each data point is described by a Cartesian coordinate and optionally contains further information like colors or intensities (Otepka et al. (2013)). In the case of insufficient computing power or low memory capacity, the storage and processing of every single detection point gets infeasible. As a workaround, digital elevation models (DEMs) can be alternatively used as a memory-efficient but less accurate form of spacial representation. Kweon and Kanade (1990) describe DEMs as 2D-arrays representing uniformly spaced patches of the surrounding terrain that contain elevation measurements. However, the constraint for DEMs, to keep one single height-value at every discrete map cell, implies disadvantages when dealing with overhanging objects like roofs or branches.

Based on these digital descriptions, different data processing methods are proposed in the literature to gain information about the traversability of the terrain. One popular strategy is the extraction of geometric features from the collected data describing terrain characteristics like terrain inclination and surface unevenness (Reddy and Pal (2016)). Bellone et al. (2013) propose a special variant of a local descriptor, which they call UPD (Unevenness Point Descriptor), in order to calculate gradients of the terrain surface based on PCA (Principal Component Analysis). A similar approach is presented by Gu et al. (2008). The authors use the mathematical principle of SVD (Singular Value Decomposition) to fit planes in local cut-outs of surface points and use resulting plane normals to describe the gradient of the terrain. Neuhaus et al. (2009) expand the approach of using PCA to calculate surface normals by gradually reducing the size of the local area considered in the analysis step, if no unambiguous

plane can be identified. In that way large traversable areas are efficiently determined while the recursive reduction of included points increases the accuracy in localising non-drivable obstacles.

In contrast to state-of-the-art approaches that often consider traversability as a pure geometric aspect of an environment, our method builds on an assessment of terrain features from several perspectives. The proposed combination of visual and geometric terrain information increases the robustness and accuracy of the estimations under various circumstances like changing weather conditions. Furthermore the integration of properties such as the vehicle's length, height or the ground clearance makes our analysis technique adaptable to a wide range of UGVs and hence proposes a generalised solution for the broad domain of indoor and outdoor autonomous vehicles.

3. PROPOSED METHOD

Our proposed approach to traversability analysis for mobile robots and autonomous machines in unstructured outdoor environments synergises a multi-modal procedure that embraces visual and geometric features of the surroundings with an efficient data processing strategy. Therefore we developed a robust and real-time capable perception system paired with a meaningful way to estimate the traversability of the environment. As the resulting outputs of our analysis procedure are subsequently used by UGVs in high-level algorithms like path planning, a safe and intended operation of the vehicles is inevitably tied to the plausibility of the terrain assessments made. In order to guarantee a preferably complete description of terrain features and thus ensure the required logic behind the data, we assume traversability to be a function of a multidimensional decision space spanned by the following three influencing factors (Fig. 1):

- Topography of environment (Geometric Features)
- Type of ground (Visual Features)
- Characteristics of UGV (Vehicle Specification)

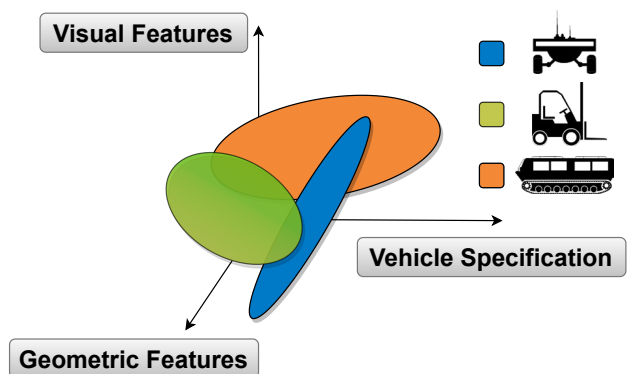


Fig. 1. Decision space spanned by geometric features, visual features and vehicle specifications, mapping to a 1D sample space of continuous traversability values. Coloured ellipsoids mark vehicle specific traversability tolerance zones for exemplary UGVs.

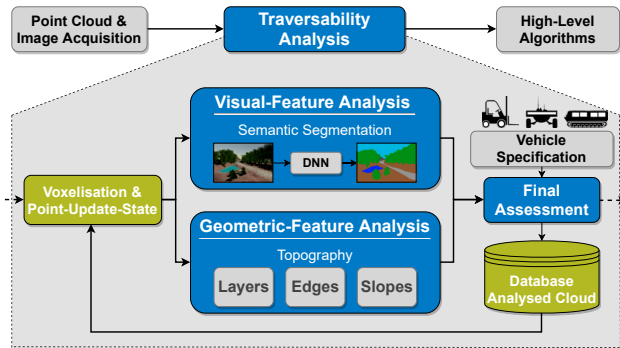


Fig. 2. Overview of the proposed framework consisting of analysis steps for traversability estimation (blue) and a concept for iterative point cloud fusion (green).

3.1 Overview

Our proposed algorithmic concept is schematically illustrated in Fig. 2. Basically our method represents the logical interface between raw input data of the real environment acquired by sensors and high-level algorithms realising and controlling specific functionalities of the UGV. Going more into detail, our method is made up of two interconnected sub-processes. One core component is the *traversability estimation* being implemented as a parallel analysis of visual and geometric features followed by the final assessment including vehicle-specific parameters (Fig. 2 - blue). The second part of the algorithm represents our concept of *iterative data-fusion* that combines raw incoming measurements with a persistent database containing an analysed cloud that consists of traversability estimations made at preceding time steps (Fig. 2 - green). Within the following sections all single components of the proposed framework are explicitly discussed.

3.2 Voxelisation and Point-Update-State Principle

Real-time capability is one of the fundamental characteristics our analysis method has to hold in order to satisfy general safety requirements. This specification is challenging to fulfil when huge point clouds have to be processed, since each step of our analysis procedure has to be executed on every single data point. As with every new sensor measurement the aggregated point cloud increases in size, real-time capability gradually suffers from the iteratively rising computing time. To counteract this problem, we propose a logical way to fuse raw sensor data with a feedback of already analysed point clouds. Paired with a specific downsampling method, referred to as voxelisation, our approach results in significantly smaller input data for the downstream assessments.

Basically, during each iteration the raw input cloud from the sensor system is overlaid with an equally spaced 3D-grid. It can be imagined as a set of cubes (*Voxels*) with defined dimensions being strung and stacked so to enframe the whole point cloud. During the *voxelisation step*, all data points being situated within a common cube are merged in terms of the arithmetic mean. Afterwards the downsampled cloud is joined with the content of the *persistent database* containing the analysed point cloud that has been acquired in previous iterations (Fig. 2 - Database Analysed Cloud). For every voxel this overlap results in

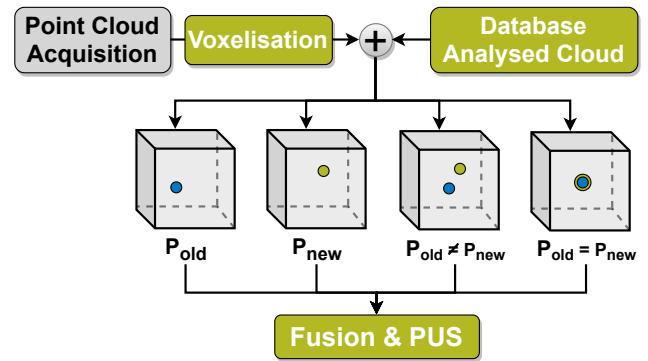


Fig. 3. Data fusion procedure: overlapping new inputs with the previously analysed cloud leads to 1 of 4 cases.

one of four unambiguous cases depicted in Fig. 3. As it can be seen, there are single or duplicate points possible within each voxel. Single ones occur either due to sensor measurements of previously unknown parts of the environment (P_{new}), or already exist in the persistent database (P_{old}). In contrast, repeatedly investigating a familiar area in consecutive scans or at later moments result in two data points. They either have almost the same coordinates when the sensor perceives an unaltered terrain patch or are noticeably unequal if the environment has changed between the scans. By using an adjustable threshold defining the minimum euclidean distance necessary for two points to be considered unequal, we are able to fine-tune the method's sensitivity with respect to terrain changes.

Depending on the predominant case (Fig. 3) only the most current point is preserved within each voxel, which finally gets marked by a binary value, referred to as *PUS* (*Point-Update-State*). This value indicates, whether a specific point has to be considered during the subsequent analysis process or can be neglected, if it is a familiar one. In the last case the underlying traversability estimation is already up-to-date. Tagging every point with its PUS at this early phase of the analysis process guarantees to keep the number of single computations as low as possible during the subsequent parts of the algorithm and thus ensures the maintenance of a real-time capable operation.

In summary, by voxelising the incoming data and using the novel concept of PUS we implement an efficient system for data cleansing. Thereby we not only ensure a viable fusion of old and new sensor information, but always guarantee the use of the latest data by a conscious removal of redundant or out-dated points.

3.3 Geometric-Feature Analysis

One major factor influencing the terrain traversability for an UGV is the appearance of the surface with respect to characteristic geometric features. Focusing on safety, it is essential to make the autonomous vehicle aware of areas in the environment containing impassable obstacles or terrain that is too steep to ascend. As a result, fundamental information like inclination and object locations have to be made available to the UGV. For this purpose we propose a processing sequence shown in Fig. 4, where geometric terrain features are calculated in a three-step analysis. As the points have to fulfil certain geometric criteria at every step, subsets of them can be immediately marked

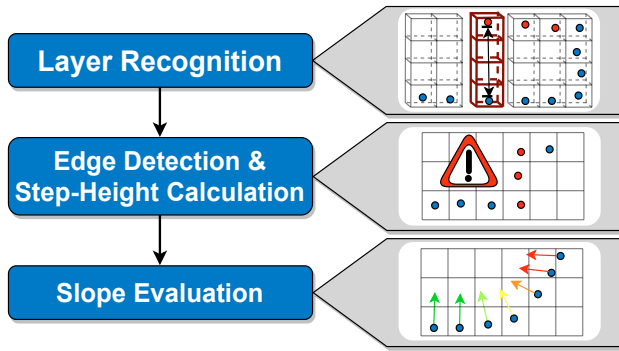


Fig. 4. Sequence of geometric-feature evaluation.

as untraversable leading to a reduction of the calculation effort for the subsequent parts of the analysis sequence.

Layer Recognition

In general, terrain traversability not only refers to the actual surface the UGV drives on, but also has to take the potential hazard of overhanging objects into account. As a result, it is crucial to consider the vehicle's dimensions when deciding about the risk of driving below such an element of the environment. For instance a drooping branch is uninteresting for a robot with a low height, whereas potentially depicting a serious collision obstacle for taller vehicles. In the first case, measurement points corresponding to the exemplary branch can be interpreted as part of a second layer within the input cloud and can be neglected during the subsequent traversability assessment. In contrast to that, the same points have to be associated with the principal layer in the second case, as the object mentioned poses a practical risk of collision (Fig. 5).

Thanks to the previously described process of voxelisation, the point cloud has a structure implicitly predetermined by the 3D-grid used during the downsampling step. As every voxel has a unique ID that is based on its position within the grid it is possible to withdraw columns of stacked voxels at arbitrary x- and y-coordinates. Built on the subset of points corresponding to these voxels, a comparison of the vertical distance between every pair of consecutive points is executed in a bottom-up manner.

Using 1, we allocate two sequential points to the same layer as long as they clamp a gap smaller than the vehicles maximum height. Otherwise the point with the higher z-value marks the beginning of a new layer within that voxel-column.

$$L(p_i) = \begin{cases} 1 & \text{if } i = 1 \\ L(p_{i-1}) & \text{if } i > 1, \Delta z < z_{min} \\ L(p_{i-1}) & \text{if } i > 1, z_{min} < \Delta z < z_{max} \\ L(p_{i-1}) + 1 & \text{if } i > 1, \Delta z > z_{max} \end{cases} \quad (1)$$

where p_i , with $i = 1, \dots, N$, indicate the set of consecutive points in a voxel-column, $L(p_i)$ the points' evaluated associative layer and z_{min}, z_{max} describe vehicle specific thresholds shown in Fig. 5.

As some UGVs like an autonomous forklift with an adjustable lift pole are characterised by a variable total height, it is additionally necessary to check for the case of point gaps within the vehicle specific height range.

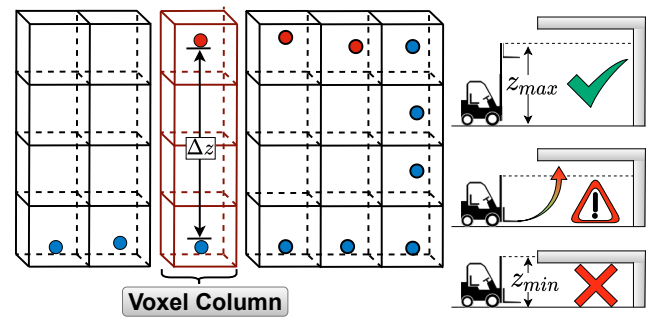


Fig. 5. Layer recognition within voxel column (blue: principal layer, red: second layer). Different cases have to be considered for UGVs with variable maximum heights.

In this case, the gap height is stored as another point feature serving subsequent algorithms as a control input to selectively limit the vehicle height at the corresponding areas. Thereby a safe transit below overhanging obstacles like doors of factory buildings or a canopy is guaranteed.

Edge Detection and Step-Height Calculation

After decreasing the number of relevant points within the input cloud by eliminating those not belonging to the principal layer, a novel technique for *obstacle detection* is executed on the remaining cloud. In this context the term *obstacle* refers either to objects above ground-level like buildings and other vehicles (positive obstacles), or to oppositional elements like potholes (negative obstacles). To identify both types in the course of one single calculation step, we make use of an efficient edge detection principle usually applied in the field of image processing. First, we transfer our problem statement from the 3-dimensional space to a 2-dimensional representation as shown in Fig. 6. As illustrated, we project the point cloud onto a 2D image plane by depositing the z-coordinate of the highest point in every voxel column in its corresponding pixel. In that way, a grey scale image representing the surface of the terrain, also referred to as height-map, is created. Using a well-tried edge-detection algorithm proposed by Canny (1986), the image is examined with respect to distinct gradients in the grey scale values corresponding to boundaries between drivable ground and positive or negative obstacles in the environment. The major steps of this procedure are the convolution of the image with two specific filter matrices, referred to as vertical and horizontal *sobel operators*¹,

¹ www.wikipedia.org/wiki/Sobel_operator

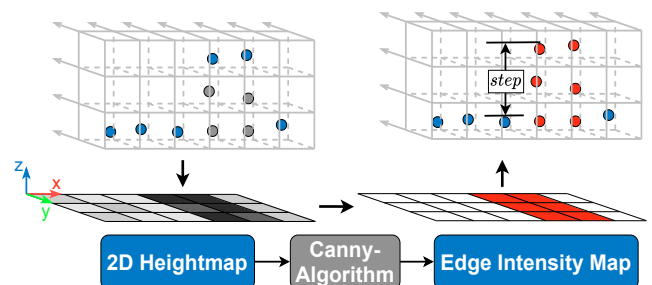


Fig. 6. Edge detection principle. left: maximum voxel point to heightmap projection (blue points) right: back-projection of resulting edge pixels (red points)

calculating the absolute intensities of the resulting edges and identifying the orientations of the edge lines. Based on the procedure's results the height of each located edge is calculated by taking the pixel value on top of each edge line and a second one perpendicular to it at its bottom. The vertical offset in between finally represents the sought-after step-height (Fig. 6 - marked step).

Depending on the vehicle's capabilities, each obstacle identified by means of this step height is classified as traversable or hazardous. This information finally gets back-projected on the 3D cloud, marking all points within common voxel columns corresponding to critical edges as not traversable. As a logical consequence, these points are insignificant for subsequent analysis steps, as their traversability is already certain at this phase of the procedure.

Slope Evaluation

As last analysis step concerning geometric terrain features, *local surface inclination* is investigated at every remaining point in the cloud that is part of the principal layer, but does not belong to any obstacle identified during the previous edge-detection process. In this context, the slope of the ground at a specific point in the cloud can be described as the angle between the vertical axis of the global coordinate system and the normal vector of a plane approximating the terrain surface in the local area around the query point. To obtain this information we apply the mathematical tool of PCA² (principal component analysis) on a neighbourhood of points (excluding edges) within a sphere of defined radius surrounding the central query point. The method makes use of the fact that the best fitting plane to the subset of extracted data is the one defined on the highest variance of the points. We therefore compute the covariance matrix of the data and determine their eigenvalues and in succession the corresponding eigenvectors. When the subset of points approximately form a plane-like surface, two of these eigenvalues are significantly higher than the third one, in which the eigenvector of the last-mentioned eigenvalue represents the sought-after normal vector at the query point.

3.4 Visual-Feature Analysis

To guarantee a safe navigation of an UGV in an outdoor environment it is not sufficient enough to exclusively calculate traversability on the basis of geometric terrain features, but rather also consider the type of the ground as an additional influencing factor. Different ground materials like grass, gravel and soil can be predominantly distinguished by means of their visual characteristics. We therefore propose the use of a convolutional neural network following the specific *UNet* architecture proposed by Ronneberger et al. (2015) to perform semantic segmentation on single camera images.

The choice of this specific network architecture was made because of its high precision concerning the localisation of the semantic classes which makes it possible to directly associate the gained information with the corresponding 3-dimensional data point in the point cloud. Basically our network uses the same structure and hyperparameters as

the original one, but differs in the size of the input image and the number of output layers, as we are distinguishing between 25 different semantic classes. This specific number results from the dataset we used for training the neural network. Specifically the training was performed on the open-source *RUGD-dataset* by Wigness et al. (2019), consisting of images showing rough and highly unstructured outdoor environments. Additionally the dataset contains the corresponding segmentation masks subdividing the image content in up to 25 different object- and ground-types like grass, soil, buildings or trees.

3.5 Final Assessment

As the final step during each iteration of our proposed analysis technique, all the information determined in the previous investigations is fused with the aim to calculate the conclusive traversability estimation T (Fig. 7). For that purpose vehicle specific parameters are integrated in the final analysis step. These parameters specify the system's limits with respect to the maximum inclination (θ_{max}) and step height (h_{max}) and offer weights to adjust the influence of each single factor (w_θ, w_h). For consistency the weights always have to sum up to 1. Furthermore, a mapping of punctual ground types to vehicle specific factors g is added, indicating how unfavourable or risky different materials are for traversal ($1 = \text{bad ground}$, $0 = \text{perfect ground}$).

Using 2, we check every point p with respect to these parameters and immediately mark one as undrivable, if either slope or step exceed the limits of the particular UGV. Otherwise we realise the sought-after dimensionality reduction from the multidimensional decision space (visual, geometry and vehicle specs) to the 1D sample space of a continuous traversability indicator. In other words, every point p in the cloud takes any traversability value T between 0 ("unacceptable") and 1 ("perfect"), implicitly standing for the vehicle characteristic *terrain quality*.

$$T(p) = \begin{cases} 0 & \text{if } \theta(p) > \theta_{max}, h(p) > h_{max} \\ 1 - (1 - g(p))(w_\theta \frac{\theta(p)}{\theta_{max}} + w_h \frac{h(p)}{h_{max}}) & \text{else} \end{cases} \quad (2)$$

Subsequent algorithms can now take the evaluated point clouds as input for various calculations. Especially *optimisation problems* like finding the most efficient path through an environment benefit from our proposed way of indicating traversability as a continuous quality value.

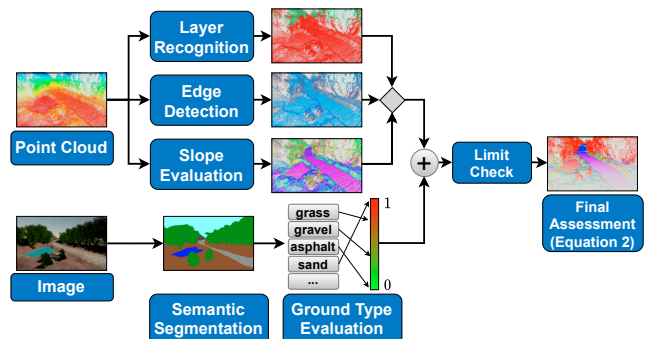


Fig. 7. Final traversability assessment based on the fusion of geometric and visual features of the environment.

² www.wikipedia.org/wiki/Principal_component_analysis

Besides, our results significantly contribute to a *safe operation* of UGVs when being used to appropriately set control outputs in order to adjust process variables like the vehicle's speed to the predominant terrain.

4. EXPERIMENTS AND RESULTS

To evaluate our proposed system for traversability analysis, we combine digital tests within a virtual environment with examinations of the application in real-world situations on a prototypical vehicle. We explicitly used simulation during the major development phase, as it allows to model specific environments particularly challenging for the analysis of terrain features. In contrast to that, testing the system on a real prototype is accompanied and massively influenced by aspects like sensor noise and weather conditions and therefore an ideal basis for increasing the methods robustness and stability. Fig. 8 illustrates the outputs of our method for different UGVs with corresponding vehicle parameters in a challenging outdoor situation. Additionally, Fig. 9 shows a representative abstract of a real-world application. The examples given substantiate the logic behind our multi-modal analysis concept and show the general adaptability to UGVs of various kinds.

5. CONCLUSION

Our proposed concept for traversability analysis shows good results in simulated tests as well as in real-world situations. The extensive evaluation procedure confirms the plausibility and robustness of our multi-modal approach in different surroundings, making the system applicable for different kinds of mobile robots and autonomous vehicles in outdoor environments. The proposed analysis technique increases safety for UGVs, as traversability is not considered as a binary state (drivable or not), but as a quality value describing the danger and effort of navigating on the terrain examined. It has to be considered critical, that our method heavily depends on the quality of incoming sensor measurements. For this reason, future research will focus on further development of our data-fusion strategy to better deal with noisy input data and intercept outliers affecting the analyses procedure adversely. Nevertheless, our proposed approach of regarding traversability as a vehicle-specific aspect influenced by geometric features of the environment and predominant ground materials produces meaningful results that can be used unob-

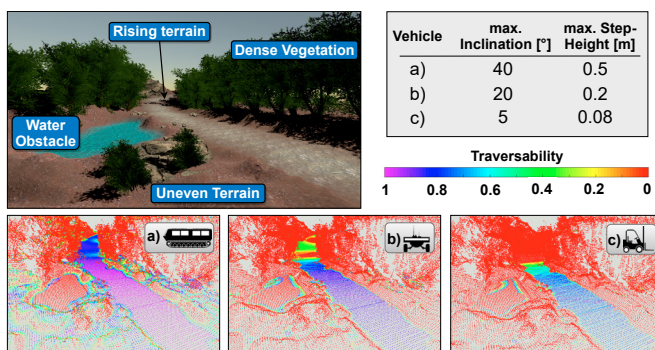


Fig. 8. Results of a simulated outdoor environment for different vehicle specifications depicted in the table.

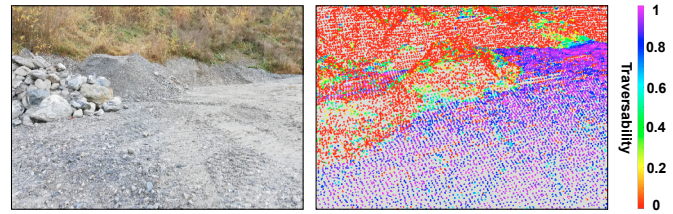


Fig. 9. Results of real-world application. Vehicle parameters are set according to our prototypical machine.

jectionable by any high-level algorithm further controlling an autonomous system.

REFERENCES

- Bellone, M., Messina, A., and Reina, G. (2013). A new approach for terrain analysis in mobile robot applications. In *IEEE International Conference on Mechatronics*, 225–230.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.
- Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *Computer*, 46–57.
- Gu, J., Cao, Q., and Huang, Y. (2008). Rapid traversability assessment in 2.5d grid-based map on rough terrain. In *International Journal of Advanced Robotic Systems*, 389–394.
- Kim, D., Sun, J., Oh, S.M., Rehg, J.M., and Bobick, A.F. (2006). Traversability classification using unsupervised on-line visual learning for outdoor robot navigation. In *Proceedings 2006 IEEE International Conference on Robotics and Automation*, 518–525.
- Kweon, S. and Kanade, T. (1990). High resolution terrain map from multiple sensor data. In *IEEE International Workshop on Intelligent Robots and Systems, Towards a New Frontier of Applications*, 127–134.
- Neuhaus, F., Dillenberger, D., Pellenz, J., and Paulus, D. (2009). Terrain drivability analysis in 3d laser range data for autonomous robot navigation in unstructured environments. In *Conference on Emerging Technologies & Factory Automation*, 1–4.
- Otepka, J., Ghuffar, S., Waldhauser, C., Hochreiter, R., and Pfeifer, N. (2013). Georeferenced point clouds: A survey of features and point cloud management. *ISPRS International Journal of Geo-Information*, 2, 1038–1065.
- Papadakis, P. (2013). Terrain traversability analysis methods for unmanned ground vehicles: A survey. *Engineering Applications of Artificial Intelligence*, 26(4), 1373–1385.
- Reddy, S.K. and Pal, P.K. (2016). Detection of traversable region around a mobile robot by computing terrain unevenness from the range data of a 3d laser scanner. *International Journal of Intelligent Unmanned Systems*, 4(2), 107–128.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *LNCS*, 9351, 234–241.
- Wigness, M., Eum, S., Rogers, J.G., Han, D., and Kwon, H. (2019). A rugd dataset for autonomous navigation and visual perception in unstructured outdoor environments. In *International Conference on Intelligent Robots and Systems*.