

## Generating high-level knowledge to enrich maps for autonomous ground vehicle

著者	Thomas Westfechtel
number	63
学位授与機関	Tohoku University
学位授与番号	情博第692号
URL	<a href="http://hdl.handle.net/10097/00127265">http://hdl.handle.net/10097/00127265</a>

氏名	トーマス ウェストフェクトル Thomas Westfechtel
学位の種類	博士 (情報科学)
学位記番号	情博第 692 号
学位授与年月日	平成 31 年 3 月 27 日
学位授与の要件	学位規則第 4 条第 1 項該当
研究科、専攻	東北大学大学院情報科学研究科 (博士課程) 応用情報科学専攻
学位論文題目	Generating High-Level Knowledge to Enrich Maps for Autonomous Ground Vehicles (自律移動ロボット用の地図を豊かにする高レベルの知識生成に関する研究)
論文審査委員	(主査) 東北大学教授 田所 諭 東北大学教授 岡谷 貴之 東北大学准教授 鏡 慎吾 東北大学准教授 大野 和則

## 論文内容の要旨

### 第 1 章

In view of the aging population of Japan, robotic solutions will play an integral role to overcome the resulting problems of a shrinking workforce as well as the problem of an increasingly large portion of population in need of special care-taking services. However, currently it is still challenging for robots to easily interact with humans or even operate within an environment that is shared with humans due to the sheer complexity. The robot has to simultaneously localize itself in the environment, perceive and understand the surrounding, and finally plan its action according to its environment, nearby agents and its task. To alleviate this problem, a common way is to provide the robot with maps of the environment containing a-priori high-level knowledge about it. These maps can vastly help the robot in the perceiving and understanding task. This could be in form of static objects of high interest like the position of traffic lights in a street scene. Another form of high-level knowledge could be location specific rules about the environment like which lanes are used for turning at a crossroad. In this work, the problem of generating high-level knowledge to enrich maps for autonomous ground vehicles from different standpoints for three different tasks is investigated.

### 第 2 章

The first task has its background in a crawler-type robot that is supposed to patrol a manufacturing plant on a daily basis for inspections tasks to alleviate the burden of human workers. During the route, the robot has to climb several stairways. This study presents a method for accurately detecting, localizing, and estimating the characteristics of stairways using point cloud data.

The main challenge is the wide variety of different structures and shapes of stairways. This challenge is often aggravated by an unfavorable position of the sensor, which leaves large parts of the stairway occluded. This can be further aggravated by sparse point data.

These difficulties are overcome by introducing a three-dimensional graph-based stairway detection method combined with competing initializations. The stairway

graph characterizes the general structural design of stairways in a generic way that can be used to describe a large variety of different stairways. By using multiple ways to initialize the graph, stairways can be robustly detected even if parts of the stairway are occluded. Furthermore, by letting the initializations compete against each other, the best initialization that accurately describes the measured stairway can be found. The detection algorithm utilizes a plane-based approach. Different planar segmentation algorithms were also investigated and experimentally compared in an application-orientated manner.

The system accurately detects and estimates the stairway parameters with an average error of only 2.5mm for a variety of stairways including ascending, descending and spiral stairways. The method robustly works with different depth sensors for either small or large-scale environments and for dense and sparse point cloud data. Despite this generality, the system's accuracy is higher than most state-of-the-art stairway detection methods.

### 第 3 章

The second task deals with self-driving vehicles. Self-driving vehicles will provide mobility for elderly people especially in rural regions with few public transport options. To operate, the self-driving vehicles rely on detailed semantic maps of the environment. In this work, a method to autonomously generate a semantic map enriched with knowledge of parking spot locations is proposed. The method detects and uses parked vehicles in the surroundings to estimate parking lot topology and infer vacant parking spots via a graph-based approach. The method works for parking lot structures in different environments, such as structured parking lots, unstructured/unmarked parking lots, and typical suburban environments. Using the proposed graph-based approach to infer the parking lot structure, it is possible to extend the estimated parking spots by 57%, averaged over six different areas with ten trials each. The accuracy of the algorithm is increased when combining multiple trials over multiple days. With ten trials combined, the whole parking lot structure was estimated, and all parking spots were detected in four out of the six evaluated areas.

### 第 4 章

In the third task, the generation of semantic maps for self-driving vehicles was investigated. The main focus is to generate large scale semantic 3D maps of the environment where each point is classified depending on their semantic class (i.e. road, sidewalk, building, ...). In this work, the focus was on the fusion of RGB data of a camera and depth data of a LIDAR.

With the recent rise of deep convolutional neural networks, the semantic segmentation of pictures has made a great boost in terms of performance. To build upon this performance boost and the color and semantic data gathered from a round view camera system with depth data gathered from a LIDAR sensor were fused. In the framework, each LIDAR scan was projected on the pictures to extract the color and semantic information while at the same time creating a large-scale 3D maps of the environment using a LIDAR-based SLAM algorithm. A random forest with combined geometrical features of the LIDAR and color and semantic features from the camera system was trained. The features complement each other and the combination of them yields an increase in IoU of around 20

percent points compared to only using geometrical features and around 13 percent points compared to using mainly the semantic information gathered from the pictures. The system achieves an average IoU-score of around 62% over 14 different semantic classes, including three different classes for ground types. Through this sensor fusion state-of-the-art results were achieved in comparison to other large scale outdoor semantic labeling methods, despite having more semantic classes and considerably less training data. Some examples on how the extracted high-level knowledge can be applied for further applications is shown, in particular the ability of the technique to retrieve trajectories of pedestrian and vehicles within the environment. In addition, using the acquired knowledge points of the specific classes can be removed from the overall point cloud, i.e. removing the traffic from the environment or removing the vegetation to get views on buildings from angles that are normally hidden by trees or other objects.

## 第 5 章

In this study, different techniques for generating high-level knowledge to enrich maps for autonomous ground vehicles are shown. Firstly, object detection in cluttered environment techniques are demonstrated for the particular case of stairway detection (Chapter 2) as well as vehicle detection (Chapter 3). Secondly, a way to learn high-level knowledge through observation of other traffic participant is introduced for the specific case of parking spot detection (Chapter 4). Thirdly, a method for detailed semantic scene segmentation is shown for the case of urban street scenes (Chapter 5). These techniques can help overcome the current difficulties and limits of autonomous ground vehicles. By adding the acquired high-level knowledge to a map of the environment, the robot can rely on a broader a-priori environment knowledge. This does not only help to increase the robot's scene understanding but is an essential part for further action planning of the robot.

## 論文審査結果の要旨

我が国において、少子高齢化に伴う労働力の不足や高齢の人々の生活を支えるためにロボットによる自動化が必要である。しかし工場プラント点検や自動運転といったロボットが実世界を自律的に動き回るアプリケーションを実現するには、複雑な環境の中から対象（階段など）の構造を認識し、その環境が持つ空間の意味やルール（空間配置など）を解釈する能力の開発が課題となっている。すなわち、ロボットに搭載した LIDAR やカメラなどのセンサデータから複雑で不確実な対象を認識し、安定した自動走行を実現することが求められている。そのためには、異なる種類の階段にも対応できる頑健性、明確に書かれていない暗黙のルールをセンサデータから読み解き、それを地図上に明示的に表現する技術が必要である。

本論文は、LIDAR で計測した 3 次元点群から異なる種類の階段の構造や、駐車場の明確に書かれていない暗黙のルールを読み解く新たな認識方法を実現することを目的としている。すなわち、従来の階段認識の課題を解決する新たなグラフを利用した 3 次元点群の認識方法、自動車搭載センサの長期のデータから周囲の車を認識し駐車スペースを他人の行動から類推する方法、少ない労力で機械学習のための正解データを作り地図のセマンティクスを高精度化する方法を明らかにしている。本論文はこれらの研究成果を取りまとめたものであり、全編 6 章からなる。

第 1 章は序論であり、本研究の背景、目的および構成を述べている。

第 2 章では、LIDAR で計測した 3 次元点群から、異なる種類（上り、下り、螺旋）や大きさの階段を認識する方法について述べている。異なる点群の密度や距離にある階段を頑健に認識する方法を提案し、データセットを利用して認識性能を従来法と比較している。また、グラフ表現と平面検出アルゴリズムと初期化アルゴリズムを組み合わせることで、一部が隠れている上り階段や、階段の踊り場から下り階段を頑健に認識できることを実験結果から示している。同一手法で異なる形状の階段を認識できること、異なる大きさの階段のパラメータを推定できることを述べている。

第 3 章では、深層学習を利用して LIDAR で計測した 3 次元点群から周囲の車の位置と向きを認識する方法について述べている。3 次元点群をトップビューに変換することで、画像でプレトレーニングした深層学習のネットワークを利用した認識ができ、画像と LIDAR を合わせることで認識精度を向上させている。また、トップビューに変換したデータを利用することで、認識結果から車の位置と向きを計算している。KIT (独) が公開している KITTI データセットを利用した認識精度の評価で、平均 87.9% の認識率、0.2m の位置精度で車の位置を検出できることを示している。

第 4 章では、自動車搭載センサの長期のデータから自動車の形状を抽出し、駐車スペースの空間構造を推定する方法について述べている。駐車スロットが明示的に指定されていない場所であっても、駐車スペースの構造を認識できることを示している。グラフベースのレイアウト解析により、車が停車していない場所でも平均 57% を駐車場として認識できること、10 日間のデータを統合することで 6 箇所のうち 5 箇所ですべての駐車場とその利用率を計測できることを示している。

第 5 章では、少ない労力で機械学習のための正解データを作り、地図のセマンティクスを高精度化する方法について述べている。画像と形状データのセマンティクスから特徴的な小領域を見つけ、その部分の正解データを人が作る方法を提案している。選択する小領域の数を減らすことができるため、従来よりも少ない労力で正解データを作成できることを述べている。適用の結果、60% 以上の精度で 3 次元点群を認識できることを示している。

第 6 章は結論である。

以上要するに本論文は、LIDAR で計測した 3 次元点群から異なる種類の階段の構造や、駐車場の空間が持つ暗黙のルールを読み解く新たな認識方法を詳細な研究により明らかにしたもので、応用情報科学およびロボット工学の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。