Guidance of Mobile Robot Navigation in Urban Environment using Human-Centered Cloud Map

Jae-Yeong Lee, Sunglok Choi, Seunghwan Park, Jaeho Lim, Seungmin Choi, Seohyun Jeon, Yunseok Lee, Beomsu Seo, Wonpil Yu
Intelligent Robot Research Laboratory, ETRI
218 Gajeong-ro, Yuseong-gu, Daejeon, 34129, Republic of Korea
jylee@etri.re.kr

Abstract

Autonomous navigation in a city-scale environment brings several technical challenges that are difficult to solve by traditional approaches. In this paper, we briefly discuss the limitations of the conventional navigation methods based on robot-centered environment modeling and understanding, and present recent an ongoing developments of the DeepGuider Project. The DeepGuider Project aims to develop a navigation guidance system that enables robots to navigate in urban environment without pre-mapping of the environment. In the paper, the main concepts and overall system architecture is briefly presented.

1 Introduction

Autonomous navigation in the unconstrained environments is one of the most important functions in realizing robot services. In particular, there has been increasing demand on service robots in urban and everyday living environments such as courier robots and food delivery robots. In order to enable such robot services, a robot navigation technology capable of navigating to an arbitrary destination in the urban environment is required.

Autonomous navigation in a city-scale environment brings several technical challenges that are difficult to solve by traditional methods. First, traditional robot navigation requires a precise robot-centered map which is usually built in advance through SLAM(simultaneous localization and mapping) algorithm. However, robot-centered modeling or mapping of city-scale space including every streets and buildings is very costly and sometimes impractical. Second, even after initial mapping is successful, detecting changes (static changes like new buildings, shops, trees) and keeping the map up to date for consistency is another challenging issue. The environment is not fixed and the dynamic and static changes cause significant problems for existing map-based approaches. Third, the success of map-based navigation relies on accurate localization. However, in a complex and dynamic urban environment, accurate and reliable localization is easy to fail, leading to a navigation failure.

These problems are difficult to avoid with current approaches of robot-centric environment modeling and understanding. On the other hand, unlike robots, we humans have the ability to navigate even in unvisited cities and places. It is mainly because human is able to understand and utilize semantics of the surroundings (roads, buildings, the connection of roads, landmarks) with an optional aid of abstracted map information (paths and landmarks from electronic maps such as Google Map, Naver Map).

In this paper, we introduce main concepts of the DeepGuider Project which started recently as a national research project in Korea, and describe the technical issues and the proposed system architecture. The DeepGuider

Project aims to develop a navigation guidance system that enables robots to navigate in indoor and outdoor urban environments without pre-mapping of the environment nor any pre-built robot-centered map. Instead of robot-centered map, the guidance system utilizes existing human-centered digital maps such as Google Map or Naver Map (hereinafter, they are called cloud map) to get abstracted navigation information of the environment. The abstract navigation information includes road topology, path to destination, and POIs[1] along the path. Street-view or road-view images provided by the cloud map services and GPS information can also be optionally utilized.

Main advantages of the DeepGuider approach is as follows. Since the proposed system uses existing human-centered navigation maps, there is no need for additional mapping and it is possible to apply a robot navigation service instantly to any places and areas. Therefore, if the proposed system is realized, nationwide navigation service is possible, and various indoor and outdoor robot services such as delivering goods and guiding people to places can be realized. The DeepGuider Project is an open source software project, and its all results are released in public via a GitHub repository (https://github.com/deepguider).

2 Related Works

There have been many studies on minimizing mapping efforts or mapless navigation to overcome the limits of traditional SLAM-based navigation. Brubaker *et al.* [1] proposed a self-localization method which utilizes visual odometry and online road maps as the inputs. It localizes by matching the shape of trajectory of the vehicle obtained from visual odometry with the ones from free online OpenStreetMap. They adopt a probabilistic approach to cope with inherent ambiguities in the map (*e.g.,* in a Manhattan world). Recently, Mirowski *et al.* [2] presented an end-to-end deep reinforcement learning approach that can be applied on a city scale. They show that it is possible to learn navigation directions by using only Google StreetView without pre-given map. It demonstrates large-scale learning from real-world imagery, but training and testing is done on the same environment. Google also recently announced concept of experimental research of global localization, which combines Visual Positioning Service (VPS), StreetView, and machine learning to accurately identify position and orientation in urban environment[4]. It uses the smartphone camera as a sensor and Google StreetView images as references to match. The problem is that the imagery from the phone at the time of localization may differ from what the scene looked like when the Street View imagery was collected. As one way, they suggest to filter out temporary parts of the scene and focus on permanent structure that doesn't change over time by machine learning automatically.

Another branch of approach is topological representation of the space and localization. Milford *et al.* [3] proposed the RatSLAM method based on the rat's navigation mechanism. RatSLAM builds a local graph map of the nodes of spaces in online and localizes based on the topological connectivity of the spaces and feature matching of each space. Badino *et al.* [5] proposed a hybrid topometric localization method that combines topological localization using spatial connectivity of the places and metric localization method by Bayesian filtering. Recently, Bruce *et al.* [6] presented a reinforcement learning method that learns navigation controls to reach destination based on a topological representation of the space with omnidirectional images as nodes of the navigation graph.

Road structure or topology provide an important clue for a semantic understanding of the environment and localization. However, there have been only a limited number of studies on this branch. Brubaker *et al.* [1], as described already, utilizes shape of road for self-localization. Kumar *et al.* [7] presented a method to classify road types on street images into intersection and non-intersection based on deep network ensembles. They reported 72.1% accuracy on Mapillary images which consists of 300,000 street images. Amini *et al.* [8] suggested a deep learning method to output vehicle control from raw sensor data and high level of route map using a variational network. Researches on extracting or recognizing road topology have been conducted mainly on aerial photos [9] and research on frontal images on the ground is very rare.

3 System Architecture

Figure 1 shows overall architecture of the proposed guidance system. The overall system flow is as follows. Once a user requests a robot service through the DeepGuider system with a destination information, the guidance system retrieves paths and map information from the cloud map service. Then the guidance system tries to recognize road structure, POIs, and other semantic information from the raw sensory inputs from the robot.

[1]Points of Interest: a specific point location that someone may find useful or interesting.

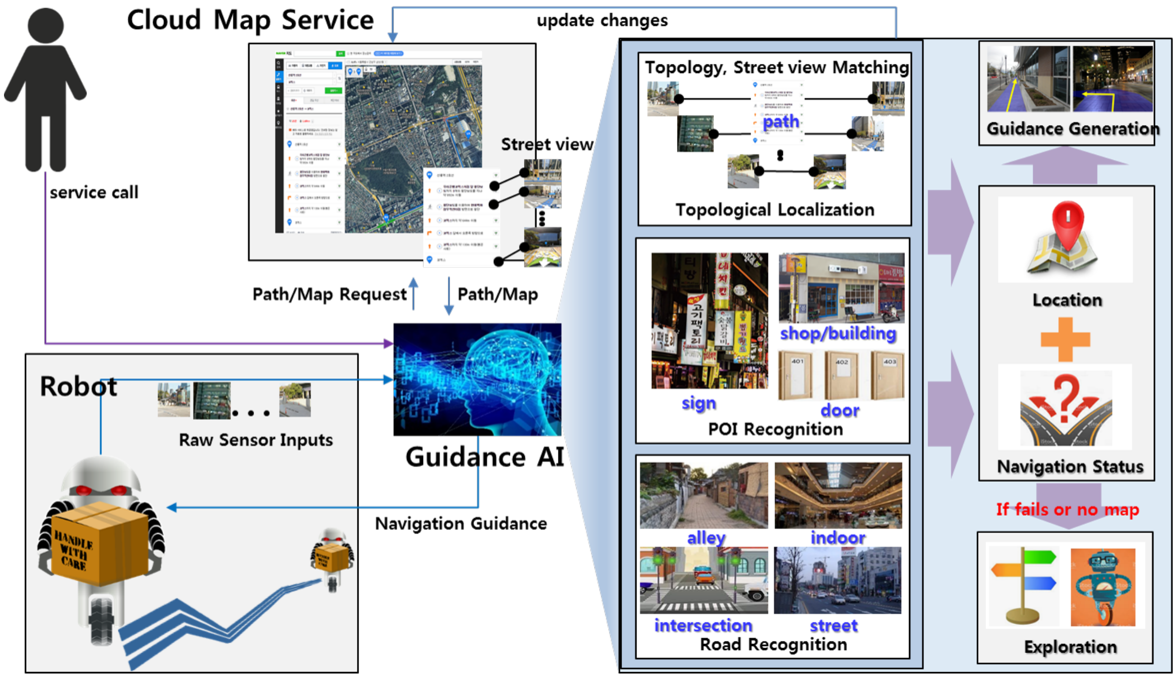

Figure 1: System architecture

The extracted information then is matched with the map information to locate robot position on the path. If the localization is successful, an online navigation guidance is generated and sent to robot. On the other hand, if the localization fails or it gets lost, the guidance system invokes an exploration module, which find ways until location is recovered.

4 Implementation

The DeepGuider system is currently in development. Therefore, only the guidance scenarios in normal and lost situation are described here.

4.1 Guidance Scenario in Normal Conditions

After a user orders a product for delivery via web, mobile or other means, the service provider checks the ordered goods, loads them on the robot, and specifies the destination of the delivery. After confirming that the delivery destination has been specified, the guidance system accesses the cloud map service and retrieves a routing path from the current position of the robot to the destination. Since the routing path obtained from the cloud map service is composed of a vehicle-centric or a pedestrian-centric path, it is difficult to directly use it for the robot navigation. The guidance system converts the routing path as a sequence of predefined robot guidance commands. The robot guidance commands consist of nodes and actions. The nodes are the important way points in the map that the robot have to pass through and the actions are the semantic motion commands to direct the robot to the next node. After that, the start command is transmitted to the robot. And during the navigation, a guidance command in every step is selected according to the position of the robot and is sent to the robot.

The robot captures the front, rear and side images and other sensor data such as GPS and odometer while navigating and send them to the guidance system. The robot also automatically avoids collisions by recognizing local obstacles. The guidance system localizes the robot on the map by comparing the image and sensor data transmitted by the robot with the map information such as street view images and POIs(Point of Interests) extracted from the cloud map service. The POIs here includes the store names and logos on the path.

Based on the estimated location of the robot, the guidance system selects and provides a guide command to transmit to the robot. If the robot's final destination is located indoor, the system guides the robot to find and access the building entrance, navigate the doorway, and reach the final destination such as a specified room or shop. If an indoor map is provided, the map information is used. If not, the destination location is estimated and searched through POI recognition and active exploration. In this case, the guidance system generates a exploring guidance command which is described in Subsection 4.2. When the destination is reached, the delivery is finished and the robot calls the user to pick up the goods.

4.2 Fail Recovery Scenario

When the robot passes a congested area or a point where it is difficult to extract feature points, the guidance system is easy to lost. For example, the robot can enter wrong alley in a complex city environments. In such cases, the guidance system recognizes that failures when a measure of reliability on the currently recognized location falls below a predefined threshold. The guidance system then propagates the context information to the internal active exploration module, and the active exploration module first attempts to return to the last successfully localized node, using the internal visual memory stored in the robot.

To return to the last successfully localized node, a guidance command utilizing visual memory is generated from the active exploration module and transferred to the robot. After the robot successfully returns to the recent node, the guidance system changes back its status to normal and resumes the normal guidance that was originally performed. If it is difficult to return to the previous node based on visual memory due to sensor uncertainty or changes in surrounding conditions, the active exploration module executes a full exploration mode. In this case, the robot tries to search in new surrounding environment until it recognizes a particular POI or node.

Even in the above two situations, the robot continuously transmits information to help the guidance system to locate the robot. And if the reliability of the current robot's position returns back to be high, the guidance system determines that the failure situation has been overcome, terminates the exploration mode, and proceeds with the normal guidance.

5 Conclusion

In this paper, we presented a new navigation framework to enable robots to navigate in urban environment without pre-mapping of the environment. The key idea is to make the robots to understand and utilize the human-centered maps or models of the environments. As the project has just started, only the concept and overall system architecture is presented in the paper. Its implementation and validation in real environment will be presented in future work.

Acknowledgement

This work was supported by the ICT R&D program of MSIT/IITP. [2019-0-01309, Development of AI Technology for Guidance of a Mobile Robot to its Goal with Uncertain Maps in Indoor/Outdoor Environments].

References

[1] Brubaker, Marcus A., Andreas Geiger, and Raquel Urtasun. "Lost! leveraging the crowd for probabilistic visual self-localization." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013.

[2] Mirowski, P., Grimes, M., Malinowski, M., Hermann, K. M., Anderson, K., Teplyashin, D., & Hadsell, R. (2018). Learning to navigate in cities without a map. In Advances in Neural Information Processing Systems (pp. 2419-2430).

[3] Milford, M., & Wyeth, G. (2010). Persistent navigation and mapping using a biologically inspired SLAM system. The International Journal of Robotics Research, 29(9), 1131-1153.

[4] https://ai.googleblog.com/2019/02/using-global-localization-to-improve.html

[5] Badino, Hernán, Daniel Huber, and Takeo Kanade. "Real-time topometric localization." 2012 IEEE International Conference on Robotics and Automation. IEEE, 2012.

[6] Bruce, J., Sünderhauf, N., Mirowski, P., Hadsell, R., & Milford, M. (2018). Learning deployable navigation policies at kilometer scale from a single traversal. arXiv preprint arXiv:1807.05211.

[7] Kumar, Abhijeet, et al. "Towards View-Invariant Intersection Recognition from Videos using Deep Network Ensembles." 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018.

[8] Amini, A., Rosman, G., Karaman, S., & Rus, D. (2018). Variational End-to-End Navigation and Localization. arXiv preprint arXiv:1811.10119.

[9] Ventura, C., Pont-Tuset, J., Caelles, S., Maninis, K. K., & Van Gool, L. (2018). Iterative deep learning for road topology extraction. arXiv preprint arXiv:1808.09814.