KHARITEH: Toward Simultaneous Localization and Mapping Through Deep Learning

Ashkan Aleali and Saman Biookaghazadeh

Abstract—We aim to present a method to boost robot localization and mapping by using deep learning algorithms. In robotics, the particle filter-based SLAM (Simultaneous Localization and Mapping) algorithm has many applications. SLAM consists of multiple parts; Landmark extraction, data association, state estimation, state update and landmark update. There are many ways to solve each of the smaller parts. Traditinally, laser beams (for indoor environment) and sonar (for marine environment) have been utilized to construct a real time map. A promising approach for solving the SLAM problem, is using vision techniques, which provides higher accuracy compared to laser and sonar and also can provide much more information compared to classical approaches. In this work, we are going to explore applicability of vision approaches using deep learning algorithms on SLAM problem.

Index Terms—

1 Introduction

In robotic mapping, simultaneous localization and mapping (SLAM) [2] is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it.

SLAM systems consists of important characteristics such as: (1) detecting features, which is able to find landmarks in an image, (2) matching landmarks in consecutive frames, (3) constructing a map of surrounding, given landmarks, and (4) memorizing previously visited areas in the map.

Deep learning We aim to present a method to boost robot localization and mapping by using deep learning algorithms. In robotics, the particle filter-based SLAM (Simultaneous Localization and Mapping) algorithm has many applications. SLAM consists of multiple parts; Landmark extraction, data association, state estimation, state update and landmark update. There are many ways to solve each of the smaller parts. Traditionally, laser beams (for indoor environment) and sonar (for marine environment) have been utilized to construct a real time map.

Another option is to use vision. We believe using computer vision techniques will provide more informative results and also is more intuitive as it resembles the way humans look at the world. Traditionally, using computer vision techniques were not feasible due to intensive computational overhead and inaccuracy in dark environments. Recent advances in hardware, computational power, and object detection techniques have shown promising results in this field . For instance, in [3], the authors used stereo and triclops cameras to come up with a vision based solution using SIFT feature detection [1].

References

[1] Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision 60*, 2 (2004), 91–110.

[2] Mountney, P., Stoyanov, D., Davison, A., and Yang, G.-Z. Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2006), Springer, pp. 347–354.

[3] Se, S., Lowe, D., and Little, J. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *The international Journal of robotics Research 21*, 8 (2002), 735–758.