Towards Robust Multi-Object Pose Estimation under Sparse and Noisy Data in Robotic Systems

Heoncheol Lee
Department of IT Convergence Engineering, School of Electronic Engineering
Kumoh National Institute of Technology
Gumi, South Korea
hclee@kumoh.ac.kr

2025 8th Conference on Cloud and Internet of Things (CIoT) | 979-8-3315-6002-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/CIOT67574.2025.11410147

Abstract— **This paper addresses the problem of estimating the poses of multiple objects under sparse and noisy data in robotics systems in unknown environments. Due to the data sparsity caused by limited data acquisition policies for saving resources and the data noises caused by environmental disturbance, the robustness of the conventional methods for the estimation degenerates. This paper proposes a robust multi-object pose estimation method to solve the problem, which can be applicable to not only ground environments but also space environments. First, a robot builds a map as a model its surrounding environments by simultaneous localization and mapping (SLAM). Then, multiple objects are detected by a visual sensor with a deep learning-based algorithm. Finally, the poses of multiple objects are robustly estimated by probabilistic filters with range data and localize them in the built map. The performance of the proposed method was tested with a real robotic system with visual and range sensors. Experimental results showed that the proposed method can robustly estimate the poses of multiple objects under sparse and noisy data.**

Keywords—Multi-object pose estimation, Sparse and noisy data, Robotic systems

I. INTRODUCTION

Accurately estimating the location and pose of multiple objects in unknown environments is a fundamental capability in autonomous robotic systems [1,2]. In particular, searching for target objects in unknown areas is a resource-intensive task. As the number of target objects and the exploration area increases, the demand for computational power, processing time, and labor also increases significantly. To address this challenge, we propose a robust multi-object pose estimation method based on mobile robotics and sensor fusion under sparse and noisy data conditions. Deep learning (DL)-based object detection algorithms [3] have been widely adopted for object localization tasks due to their ability to generalize across various object appearances and conditions. When trained with sufficient data, these algorithms can achieve high accuracy in identifying and localizing objects in diverse scenes. However, their performance significantly degenerates when sparse and noisy data are acquired.

This work was supported in part by the Institute of Information Communications Technology Planning Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2025-RS-2020-II201612, 50%), and in part by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2025 RS-2024-00438430, 50%) supervised by the IITP(Institute for Information Communications Technology Planning Evaluation).

To overcome these limitations, we integrate simultaneous localization and mapping (SLAM) techniques [4] into the object localization framework. The proposed method leverages SLAM to incrementally construct a map of the unknown environment while simultaneously estimating the robot's pose. Among various SLAM approaches, we adopt feature-based SLAM because of its robustness and real-time performance, which reduces computational complexity while maintaining reliable mapping accuracy. Although it occasionally generates noisy point clouds in vertically structured environments, the SLAM algorithm can enhance stability in both structured and unstructured environments with reduced computational overhead.

To conduct multi-object recognition within this SLAM-generated map, we incorporate a depth camera and apply a deep learning-based object detection algorithm for real-time object detection. The RGB-D information is fused with the SLAM output to associate detected objects with accurate 3D poses. The object detection algorithm was selected for its high inference speed and adaptability to newly trained object classes, outperforming earlier versions in both accuracy and efficiency. All components of the proposed system are implemented on the Robot Operating System (ROS), and the overall framework supports real-time operation even with the simultaneous execution of 3D mapping and object detection. Finally, the poses of multiple objects are robustly estimated by probabilistic filters with range data and localize them in the built map.

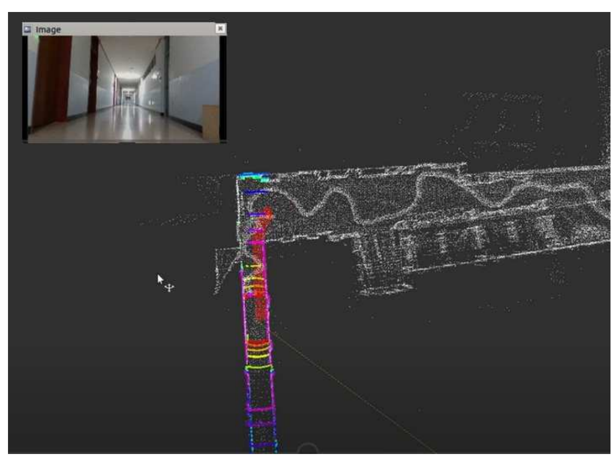

Fig. 1. An example of acquiring sparse and noisy data by visual and range sensors of robotic systmes in unknown environments.

II. MULTI-OBJEST POSE ESTIMATION

When estimating the poses of multiple objects using visual and range sensors, robustness may degenerate due to the data sparsity caused by limited data acquisition policies for saving resources and the data noises caused by environmental disturbance such as variations in ambient lighting, reflective surfaces, and the presence of transparent or semi-transparent objects. Besides, in complicated scenes, background clutter may cause the target object to be incorrectly identified or entirely missed. Due to limitations in the resolution and sensing range of depth sensors, it becomes difficult to accurately detect distant or extremely small objects. Additional sources of error include a lack of robustness under diverse environmental conditions, sensor noise, distortion, and restricted fields of view. To address these challenges, improvements in algorithmic robustness, sensor fusion techniques, and fine-grained sensor data analysis are required. A more detailed understanding of environmental variables and object physical characteristics can help adapt the data processing pipeline to improve estimation accuracy.

Particle filters (PF) and particle swarm optimization (PSO) can be applied to improve the robustness of estimating multi-object poses under nonlinear and uncertain motion conditions of target objects. However, their performance is highly sensitive to the number of particles used. In cases where the target exhibits nonlinear or unpredictable motion patterns, particle convergence tends to degrade, requiring a significantly larger number of particles to maintain estimation accuracy. This increase in particle count leads to a proportional rise in computational load, resulting in substantial processing time. Such computational demands become a critical limitation in small robotic systems that require real-time performance. Therefore, reducing the execution time of particle filter computations is essential for achieving efficient and real-time object tracking in compact robotic platforms.

This paper proposes a probabilistic method based on the PSO-PF [5] to improve the robustness of estimation while a parallelization method based on a GPU (Graphic Processing Unit) [6] is also applied to relax the increased computational complexity due to the large number of particles. In the GPU-based parallelization method with CUDA (Compute Unified Device Architecture), each thread processes data, and all threads share the same kernel to perform identical operations. Threads are organized into two hierarchical levels: a block, which is a group of threads, and a grid, which is a group of blocks. To access data within the kernel, indexing is used based on block and grid variables. To allocate memory for weight computation, the thread size is defined, and the number of blocks is dynamically determined. The relevant data required for the weight computation kernel is copied to the GPU. Using indexing, each thread executes its designated computation in parallel. The kernel computes the difference between each particle and the measurement, and assigns weights based on a Gaussian probability density function, consistent with CPU-based implementations. Finally, the resulting weights are copied back to the CPU to be used in the subsequent resampling step.

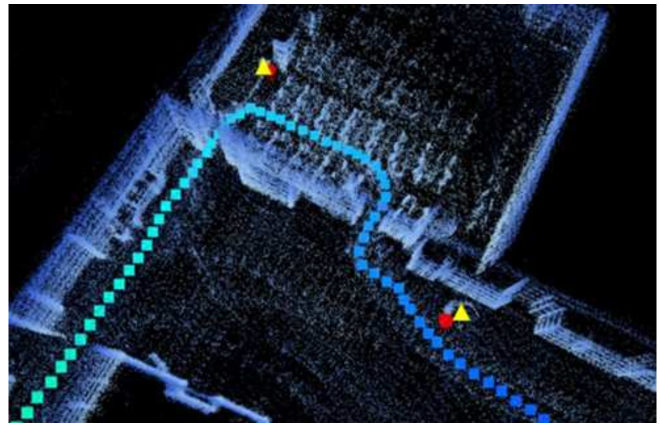

Fig. 2. Experimental results of multi-object pose estimation under sparse and noisy data. The estimated object poses are localized in the map built by SLAM. Red circles and yellow triangles represent the true poses and estimated poses of objects, respectively, which indicates the estimation errors were not significant.

III. CONCLUSIONS

This paper proposed a probabilistic filter-based method to robustly estimate the poses of multiple objects under sparse and noisy sensor data. The proposed method integrates the particle filter and the particle swarm optimization algorithm, which enhances convergence and maintains particle diversity through cooperative interactions among particles, thereby achieving more stable estimations. To relax the increased computational cost caused by using a large number of particles, the GPU-based parallelization method was also employed. Also, the poses of multiple objects are robustly estimated by probabilistic filters with range data and localize them in the built map. The performance of the proposed method was tested with a real robotic system with visual and range sensors. Experimental results showed that the proposed method can robustly estimate the poses of multiple objects under sparse and noisy data.

REFERENCES

[1] J. Kim, S. Nam, G. Oh, S. Kim, S. Lee and H. Lee, "Implementation of a Mobile Multi-Target Search System with 3D SLAM and Object Localization in Indoor Environments," in *Proc. Int. Conf. Control, Automation and Systems*, Jeju Island, South Korea, Oct. 2021.

[2] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam and S. S. Aher, "Moving Object Tracking Using Kalman Filter," in *Proc. Int. Conf. Advances in Communication and Computing Technology*, pp. 544-547, Nov. 2018.

[3] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 779-788, June, 2016.

[4] T. Shan and B. Englot, "LeGo-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," In *Proc. IEEE/RSJ Int. Conf. Inteligent Robots and Systems*, 2018.

[5] H. C. Lee, S. K. Park, J. S. Choi and B. H. Lee, "PSO-FastSLAM: An Improved FastSLAM Framework Using Particle Swarm Optimization," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics*, pp. 2763-2768, Oct. 2009.

[6] D. Kim, Y. Han, H. Lee, Y. Kim, H. Kwon, C. Kim and W. Choi, "Accelerated Particle Filter with GPU for Real-Time Ballistic Target Tracking," *IEEE Access*, vol. 11, pp. 12139-12149, Jan. 2023.