Accurate Pose Estimation Method using ToF-Stereo Camera Data Fusion

Sukwoo Jung
Contents Convergence Research Center
Korea Electronics Technology Institute
Seoul, South Korea
swjung@keti.re.kr

Harrison Yun
Innovation Laboratories
Deutsche TelekomAG
Berlin, Germany
Harrison-Hongseo.Yun@telekom.de

KyungTaek Lee
Contents Convergence Research Center
Korea Electronics Technology Institute
Seoul, South Korea
ktechlee@keti.re.kr

***Abstract*— Recent studies on sensor pose estimation have expanded due to their applications across a range of fields, including augmented reality (AR), virtual reality (VR), and robotics. Within this context, we present an innovative approach that diverges from traditional techniques involving images or IMU sensors. Instead, we used pre-reconstructed 3D maps to enhance the precision of pose estimation. By training on these 3D maps, we can track objects using data from ToF-Stereo fusion and utilize this information to calculate precise positional coordinates within indoor environments. Through experimentation with a ToF-Stereo prototype, we collected data and tested the algorithm to validate the proposed method.**

Keywords—3D reconstruction, pose estimation, navigation, feature matching, object detection

I. Introduction

sensor-based position estimation involves determining the location of objects, devices, or individuals in the physical world using various sensors and technologies. This process has applications in fields such as navigation, robotics, metaverse, augmented reality, virtual reality, and Internet of Things (IoT) [1]. we introduce the importance of accurate position estimation in various applications and outline the motivation behind sensor-based methods. We discuss the significance of multi-sensor fusion and the need for real-time and high-precision positioning solutions.

Within the confines of this discourse, we investigate the repertoire of sensors commonly employed for position estimation. Notable sensors in this context encompass Global Positioning System (GPS) units, accelerometers, gyroscopes, magnetometers, Light Detection and Ranging (LiDAR) devices, cameras, and Ultra-Wideband (UWB) technology. Leveraging these instrumentalities, a spectrum of techniques has been investigated to estimate the location of mobile sensors [2].

Of particular prominence are visual-based methodologies, such as visual odometry and Simultaneous Localization and Mapping (SLAM), which have found extensive utility in domains like robotics, autonomous vehicles, and drones. Certain scholars [3-4] have harnessed cameras and computer vision techniques to deduce position and generate maps of environments.

Using smart phone[5], emerging trends are the signal processing method such as integration of 5G and cellular signals for position estimation, Bluetooth and BLE beacons for indoor positioning, and the role of edge computing in reducing latency.

Nonetheless, the inherent limitations in accuracy and reliability exhibited by individual sensors have propelled us towards a strategic integration of multiple sensors, thereby improving the precision and robustness of pose estimation [6-7]. Within this paradigm, we expound upon sensor fusion methods, such as Kalman filtering, particle filtering, and Bayesian inference. Our paper proposes a method using 3D reconstructed data to augment the accuracy of pose estimation.

II. Proposed Method

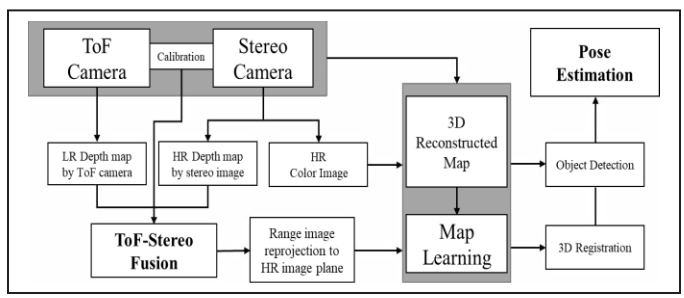


Fig. 1. Overall procedure of the proposed method

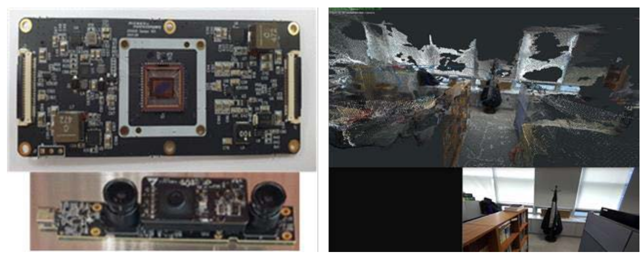
The overall procedure of the proposed approach is depicted in Figure 1. We affixed Time-of-Flight (ToF) and Stereo cameras onto a rigid frame and established their spatial relationship through calibration. The ToF camera captures a depth map at a lower resolution, whereas the stereo camera captures a higher-resolution depth map in conjunction with color images. By amalgamating these two datasets and applying color mapping via reprojection[7], we derive comprehensive 3D information of the scene. In order to accumulate the 3D data and construct a 3D map, we execute 3D registration[8] between successive data points. This process enables object detection from fusion data by utilizing the data from the acquired 3D map.

The resultant 3D map data serves as the foundation for object recognition. Through this process, specific positions within the map become discernible, thereby enabling the determination of

979-8-3503-1327-7/23/$31.00 ©2023 IEEE

camera-to-object relationships via the registration of objects within the 3D map and their real-world counterparts. Importantly, the pose information obtained through this method surpasses the precision achieved by conventional techniques reliant on images or IMU data.

As illustrated in Figure 2, we have established the testing environment outlined in (a), followed by the data acquisition procedure as shown in (b). Subsequently, we conducted a comprehensive testing of the proposed algorithm, comparing its accuracy with that of conventional methods.



(a) (b)

Fig. 2. Pose estimation using ToF-Stereo fusion test configuration. (a) ToF-Stereo test device; (b) pose estimation with 3D data test result.

III. CONCLUSION

We have introduced a novel approach to pose estimation that utilizes 3D map data and ToF-Stereo fusion data. Our method sets a new trajectory by fusing these diverse data streams, and the veracity of our approach has been underscored through rigorous testing. The results of these tests not only affirm the feasibility of our approach but also highlight its penchant for yielding remarkably accurate outcomes compared to the conventional algorithms. By virtue of this approach, we unlock a gateway to extracting high-performance pose estimation outcomes, resonating proficiently across a spectrum of diverse environments and scenarios.

The limitations are complexities entailing sensor calibration, noise reduction, sensor drift, and the seamless transition between indoor and outdoor scenarios. Each layer of complexity has been meticulously peeled away, revealing the heart of these challenges and driving our exploration of potential solutions. Within this framework, we delve into a panorama of prospective remedies and fertile avenues for future research. The possibilities encompasses the integration of machine learning strategies, the deployment of advanced error correction techniques, and the formulation of sensor calibration methods that are not just effective but also remarkably efficient.

In these concluding moments, we want to highlight how important ongoing research and constant innovation are in the field of sensor-based position estimation. We emphasize the importance of continued research and innovation in sensor-based position estimation methods to meet the demands of evolving technologies and applications.

ACKNOWLEDGMENT

This work was supported by the Technology Innovation Program (20023743, Development of Digital Twin-based AMR Operation Service Technology for Industrial Use Cases) funded by the Ministry of Trade, Industry & Energy (MOTIE, South Korea)

REFERENCES

[1] Rusnida Romli, et al. "Mobile augmented reality (AR) marker-based for indoor library navigation." IOP Conference Series: Materials Science and Engineering. Vol. 767. No. 1. IOP Publishing, 2020.

[2] Qabas A. Hameed, et al. "Development of Augmented Reality-based object recognition mobile application with Vuforia." Journal of Algebraic Statistics 13.2 (2022): 2039-2046.

[3] Talha Khan, et al. "A systematic comparison of the accuracy of monocular RGB tracking and LiDAR for neuronavigation." Healthcare Technology Letters 9.6 (2022): 91-101.

[4] Rieder, Jonas SI, Danielle H. Van Tol, and Doris Aschenbrenner. "Effective close-range accuracy comparison of Microsoft HoloLens Generation one and two using Vuforia ImageTargets." 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE, 2021.

[5] Moechammad Sarosa, et al. "Developing augmented reality based application for character education using unity with Vuforia SDK." Journal of Physics: Conference Series. Vol. 1375. No. 1. IOP Publishing, 2019.

[6] Sukwoo Jung, et al. "3D reconstruction using 3D registration-based ToF-stereo fusion." Sensors 22.21 (2022): 8369.

[7] Sukwoo Jung, Seho Park, and KyungTaek Lee. "Pose tracking of moving sensor using monocular camera and IMU sensor." KSII Transactions on Internet and Information Systems (TIIS) 15.8 (2021): 3011-3024.

[8] Sukwoo Jung, et al. "Range image registration based on 2D synthetic images." Computer-Aided Design 94 (2018): 16-27.