





Visual-SLAM based 3-dimensional modelling of indoor environments

Simla Özbayrak ¹, Veli İlçi ^{*1}

¹ Ondokuz Mayıs University, Department of Geomatics Engineering, Samsun, Türkiye, simlaozbayrakk@gmail.com, veli.ilci@omu.edu.tr

Cite this study: Özbayrak, S., & İlçi, V. (2024). Visual-SLAM based 3-dimensional modelling of indoor environments. International Journal of Engineering and Geosciences, 9 (3), 368-376

<https://doi.org/10.26833/ijeg.1459216>

Keywords

Indoor Modelling
Visual-SLAM
Unmanned Ground Vehicle
Stereo Camera

Research Article

Received: 26.03.2024
Revised: 06.05.2024
Accepted: 10.05.2024
Online Published: 17.11.2024



Abstract

Simultaneous localization and mapping (SLAM) is used in many fields to enable robots to map their surroundings and locate themselves in new circumstances. Visual-SLAM (VSLAM), which uses a camera sensor, and LiDAR-SLAM, which uses a light detection and ranging (LiDAR) sensor, are the most prevalent SLAM methods. Thanks to its benefits, including low-cost compared to LiDAR, low energy consumption, durability, and extensive environmental data, VSLAM is currently attracting much attention. This study aims to produce a three-dimensional (3D) model of an indoor environment using image data captured by the stereo camera located on the Unmanned Ground Vehicle (UGV). Easily measured objects from the field of operation were chosen to assess the generated model's accuracy. The actual dimensions of the objects were measured, and these values were compared to those derived from the VSLAM-based 3D model. When the data were evaluated, it was found that the size of the object produced from the model could be varied by ± 2 cm. The surface accuracy of the 3D model produced has also been analysed. For this investigation, areas where the walls and floor surfaces were flat in the field were selected, and the plane accuracy of these areas was analysed. The plain accuracy values of the specified surfaces were determined to be below ± 1 cm.

1. Introduction

The emergence of sensor, robotic, and artificial intelligence technologies has attracted corporate and academic interest in autonomous vehicle (AV) research. These vehicles must be able to locate themselves and map their environment precisely to operate safely [1]. In open-sky areas, Global Navigation Satellite Systems (GNSS) systems can satisfy the requirement for highly accurate positioning of AV's [2]. Multiple satellite system data integration enables more received signals to improve position accuracy [3-4]. The inability to receive sufficient and healthy satellite signals in locations like urban areas, where direct access to the receiver is prohibited, might cause the position accuracy to drop to an undesirable level even with numerous satellite signals [5]. Differential GNSS positioning techniques like Real-Time Kinematic (RTK) offer accuracy at the centimetre level in open areas but not in places where signals are predominantly obstructed [6-7]. In these places, GNSS is frequently integrated with the Inertial Measurement Unit (IMU) to provide continuous and precise coordinates [8]. For AVs to operate securely in any environment, they must be able to map their surroundings, precisely

position themselves, and distinguish between moving and stationary objects. Light detection and ranging (LiDAR) or camera sensors included with GNSS/IMU integration are used for mapping in AVs. Using LiDAR and camera sensors, Simultaneous Localization and Mapping (SLAM), which has recently gained much attention, is also utilized for positioning and mapping.

Autonomous robots are another kind of vehicle that technology is bringing into our lives and using more and more of each day. These robots are frequently used in huge indoor environments, such as shopping malls, airports, and warehouses, and moreover, they are used in our houses as robotic vacuum cleaners. As with AVs, these robots must be able to accurately locate themselves and simultaneously map their surroundings inside the indoor spaces to move around safely. Since the signals cannot reach the receiver, location accuracy in GNSS-based positioning in interior environments is either non-existent or erroneous. While there are studies on the use of Inertial Navigation System (INS) [9], Wi-Fi [10-11], UWB (Ultra-Wide Band) [12], and Radio Frequency Identification (RFID) [13] sensors for indoor position determination where GNSS systems cannot be used efficiently, camera [14-16], Radio Detection and Ranging

(RADAR) [17], and LiDAR [18-20] based sensor solutions are generally studied for mapping.

SLAM is a method of determining the vehicle's position and generating a map of its surroundings using one or more sensor data. The SLAM method collects two-dimensional (2D) or three-dimensional (3D) geometric information from an unknown environment by sensor systems, then uses this data to predict the system's location and create a map [21]. In addition to its use in indoor space robots, SLAM can be used in a variety of areas, including outdoor [22], underwater [23] and aviation system [24] applications. The most widely used SLAM methods are LiDAR-SLAM and Visual-SLAM (VSLAM). LiDAR uses its laser signals to detect and visualize objects up to 300 meters away [25]. The LiDAR-based SLAM method has the ability to create precise maps of indoor and outdoor spaces. However, long-range LiDAR sensors are quite expensive, and low-cost LiDAR sensors are unsuitable for giant areas and outdoor applications [26].

Compared to LiDAR sensors, the camera-based SLAM method attracts great attention due to the advantages of cameras, such as being cheaper, requiring low energy consumption, and providing rich environmental information [27-28]. According to their functions, the cameras can be classified as monocular, stereo, and depth [29]. Monocular cameras use only one lens to capture an image of a target. Reflecting a scene onto the camera plane results in a two-dimensional representation of a three-dimensional world, and therefore, it is impossible to calculate the distance between the target and the camera using a single photograph. Stereo cameras are merged with multiple cameras used to calculate depth information. Unlike monocular and binocular cameras, RGB-D cameras can obtain direct pixel depth using light or time-of-flight (TOF) without the need for complicated calculations. SLAM applications typically use cameras that can provide depth information [30].

Many sensors that can be used for SLAM are being introduced to the market daily. However, analysing the accuracy of the models obtained from processing sensor data with SLAM algorithms is vital for the method's usability. This study investigated the object dimensions and surface accuracy in the 3D model produced by the VSLAM method. For this aim, we produce an accurate 3D model of an indoor area by using stereo camera data mounted on an Unmanned Ground Vehicle (UGV). The VSLAM method was used to produce the model from camera images. The produced 3D model was visualized, and the accuracy of the model was tested by comparing the dimensions of the objects on the model with the actual dimensions. Furthermore, a plane-fitting analysis was carried out to test the accuracy of the surfaces chosen from the walls and the ground.

2. Method

2.1. Simultaneous localization and mapping

In recent years, some state-of-the-art sensors have been adopted to generate new fusion technologies to exceed the inaccurate localization problem of AVs. One

well-known alternative sensor fusion methodology for localization and mapping is SLAM, which combines a set of perception sensors. The SLAM method simultaneously reconstructs a map of the vehicle's environment during movement according to the collected sensor data and determines the AV's current position within the built map [31-35]. In the SLAM method, the error accumulates through time and distance, and the vehicle's position is estimated in the local coordinate system [36-37]. The accuracy, reliability, and robustness of SLAM-based methods must also be enhanced for use in complicated scenarios [38]. The growing amount of data increases the computational difficulty and strains available storage [39-40]. Camera and LiDAR are commonly used sensors to perceive surroundings using VSLAM and LiDAR-SLAM methods [41].

2.1.1. Visual-SLAM

GNSS, IMU, RADAR, and LiDAR have been used in the majority of SLAM investigations; however, in the last few years, VSLAM has developed with better information, thanks to components like colour and texture received from cameras or RGB-D sensors. Thus, VSLAM has become an alternative that performs positioning and mapping tasks in indoor and outdoor environments with only visual data [42]. In recent years, VSLAM has been frequently used in different applications, such as robotics, AVs, augmented reality (AR), and service robots, where a 3D environment model needs to be created [43]. Using the visual data, VSLAM estimates the relative orientation and translation while tracking the features of successive images [44]. The classic structure of the VSLAM system is divided into five sections: camera sensors, front-end, back-end, loop closure, and mapping modules [45].

- The camera sensor module collects image data,
- The front-end module involves processes in order to obtain the prediction of the first motion of the camera and local matching, such as monitoring and extracting characteristics from two sequential images and matching or following the extracted features on different video frames,
- The back-end module is responsible for the geometric calculation, which converts these traces into camera positions and 3D coordinates based on the theory of multiple-image geometry,
- The loop-closure module calculates image similarity in large-scale environments, eliminates accumulated errors, and recreates the mapping module.

According to the adopted error model, VSLAM can be classified as a feature-based and direct approaches [46]. The feature-based approach works by referencing distinctive points. This approach saves key points to local sub-maps or keyframes and improves with a graphically based optimisation approach [47]. The feature-based approach can quickly fail in dynamic situations with poor texture and pure rotation, but it is accurate and reliable in static areas with rich texture [46]. Some shortcomings of the feature-based approach are remedied by the direct method, which optimises the photogrammetric error of the sensor measurement by directly using the intensity

of the image [48]. Direct approaches record all pixels of images directly to local maps and use photometric losses for measurement error predictions [49].

The commonly used feature-based approaches are MonoSLAM, Parallel Tracking and Mapping (PTAM), and Oriented FAST and Rotated BRIEF-SLAM (ORB-SLAM). The first feature-based monocular approach, MonoSLAM, was developed by [50]. In MonoSLAM, a typical filter-based approach, monitoring and mapping are carried out sequentially and simultaneously using the Extended Kalman Filter (EKF). The monitoring and mapping of the PTAM method, introduced in [51], is carried out in parallel on the CPU. This solves the calculated burden problem in MonoSLAM. ORB-SLAM is an open-source monocular SLAM method recommended for large-scale environments, consisting of three stages: monitoring, mapping, and loop detection. ORB-SLAM2, an enhanced version of ORB-SLAM, has extended the use of the method from monocular cameras to stereo and RGB-D cameras [52]. ORB-SLAM3 is the first system to use fish-eye lens models to process VSLAM with monocular, stereo, and RGB-D cameras, using only visual, visual-inertial, or multiple sensors.

Unlike feature-based approaches, direct approaches pose an optimization problem by predicting camera movements directly from pixel information. In the Direct Tracking and Mapping (DTAM) method proposed by [53], the camera's position is determined by the whole image directly aligned with the depth map. Another direct SLAM method, the Large-Scale Direct SLAM (LSD-SLAM), was developed by [54]. LSD-SLAM directly predicts the camera's position and recreates large-scale, consistent, and 3D maps of the environment in real-time using image alignment and depth prediction. Another direct method, the Direct Sparse Odometry SLAM (DSO-SLAM), is considered the best solution for estimating current accuracy and work efficiency [55].

Many studies have been carried out to solve the VSLAM problem. [56] indicates that it is more challenging to map and estimate the position in outdoors due to the continuous changes in the sun's angle at different times of the day. The study suggested an instant visual trajectory estimation method for AVs, considering that the mobile platform moves on a constant route. The proposed model was tested on the SeqSlam data set for classifying the images of the track followed, and a success rate of 78.5% was achieved. [57] proposed a VSLAM-based low-cost indoor positioning system and examined the threshold of acceptable positioning accuracy of basic features per image. The number of image entries that users need to collect for reliable positioning is also examined. A comprehensive VSLAM method has been developed to study matching and accuracy problems in the event of changes in lighting conditions. The proposed approach has achieved a positioning success of over 94% in a real building. The Semantic Optical Flow SLAM (SOF-SLAM) was improved based on the RGB-D mode of ORB-SLAM2 and was introduced as a visually meaningful SLAM system for dynamic environments [58]. This method proposes a dynamic attribute detection approach called semantic optical flow. This approach is designed with an integration that allows the efficient and

reasonable extraction of dynamic properties based on semantic and geometric knowledge. The pixel-based semantic segmentation results are used as a mask in flow to obtain a reliable base matrix to filter real dynamic properties. The monitoring and optimization module focus only on fixed features for accurate camera position estimates in dynamic environments. This method has been evaluated in general TUM RGB-D datasets and experiments in real-world environments. Compared to ORB-SLAM2, SOF-SLAM has improved 96.73% in highly dynamic scenarios.

2.2. Site survey

2.2.1. Mobile platform

In this study, a remote-controllable UGV was developed to produce a 3D model of indoor areas. This vehicle consists of the motion and imaging systems. The six-wheel drive (6WD) was used as a carrier platform for the motion system. The vehicle's motion is controlled remotely through telemetry. The motion of the vehicle is accomplished by the motor driver transmitting commands to the engines from the remote control. Lithium Polymer (LiPo) batteries power the vehicle's motion system. The vehicle imaging system consists of a stereo camera (ZED 2i), a processor (Jetson AGX Orin), a monitor, and a power supply (Figure 1). A screen was also used for instant data visualization. A strong power bank feeds the imaging system. Figure 2 shows the UGV.

2.2.2. VSLAM-based 3D modelling

In order to test the accuracy of the produced model, a corridor with doors, fire cabinets, etc., which can be easily measured, was chosen. The corridor is 34 meters long, 2.96 meters wide, and 2.85 meters high (Figure 3a). The visual data was collected with the Zed2i stereo camera. The ZED 2i stereo camera has a 120° field-of-view (FOV) and advanced IMU sensor that improve VSLAM performance by providing positional and spatial awareness. These sensors provide visual-inertial stereo SLAM with advanced sensor fusion and thermal compensation. ZED 2i can create applications in the fields of motion, artificial intelligence, and depth sensing thanks to this equipment. The images collected with the Zed2i stereo camera were evaluated using the ZEDfu application, which is part of ZED-SDK software (Figure 3b). The sequential image data was processed using the VSLAM method at ZEDfu and obtained in accordance with the actual colours of the 3D model consisting of 1,820,363 points of the study area (Figure 4). CloudCompare software was used to visualize the resulting model.

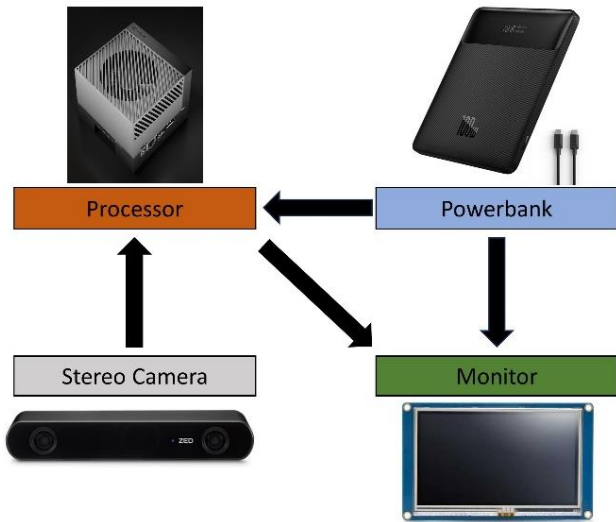


Figure 1. Monitoring system of the UGV

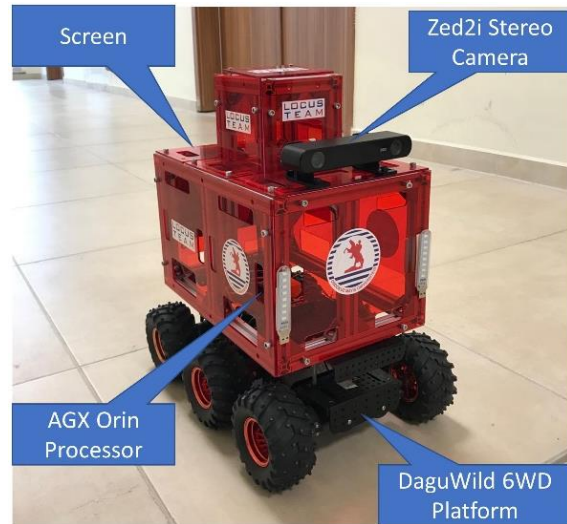


Figure 2. UGV and sensor placement



(a)



(b)

Figure 3. a. The study area, b. real-time 3D visualization



Figure 4. 3D model of the study area

3. Results

Using a measuring tape, the easily measured width and length of the items in the field of operation were

determined in order to assess the accuracy of the 3D model that was created. In addition, extra objects were placed in the corridor, such as a cabinet, during the modelling to increase the number of objects to be measured. Figure 5 shows the reference dimensions of the selected objects, and the dimensions measured using the CloudCompare software using the 3D model.

The accuracy values in Table 1 are attained when the measurement values derived from the model are compared with the reference values. Table 1 shows that the accuracy of five objects is in the range of $\pm 2\text{cm}$.

The surface flatness test is another analysis carried out for the accuracy of the model. As seen in Figure 6, a total of 16 flat surfaces from the two walls (8 surfaces from the Right Wall (R), and 8 surfaces from the Left Wall (L)) and five flat surfaces from the ground (G) were chosen for this investigation. Using CloudCompare, an accuracy examination of the surfaces was performed.

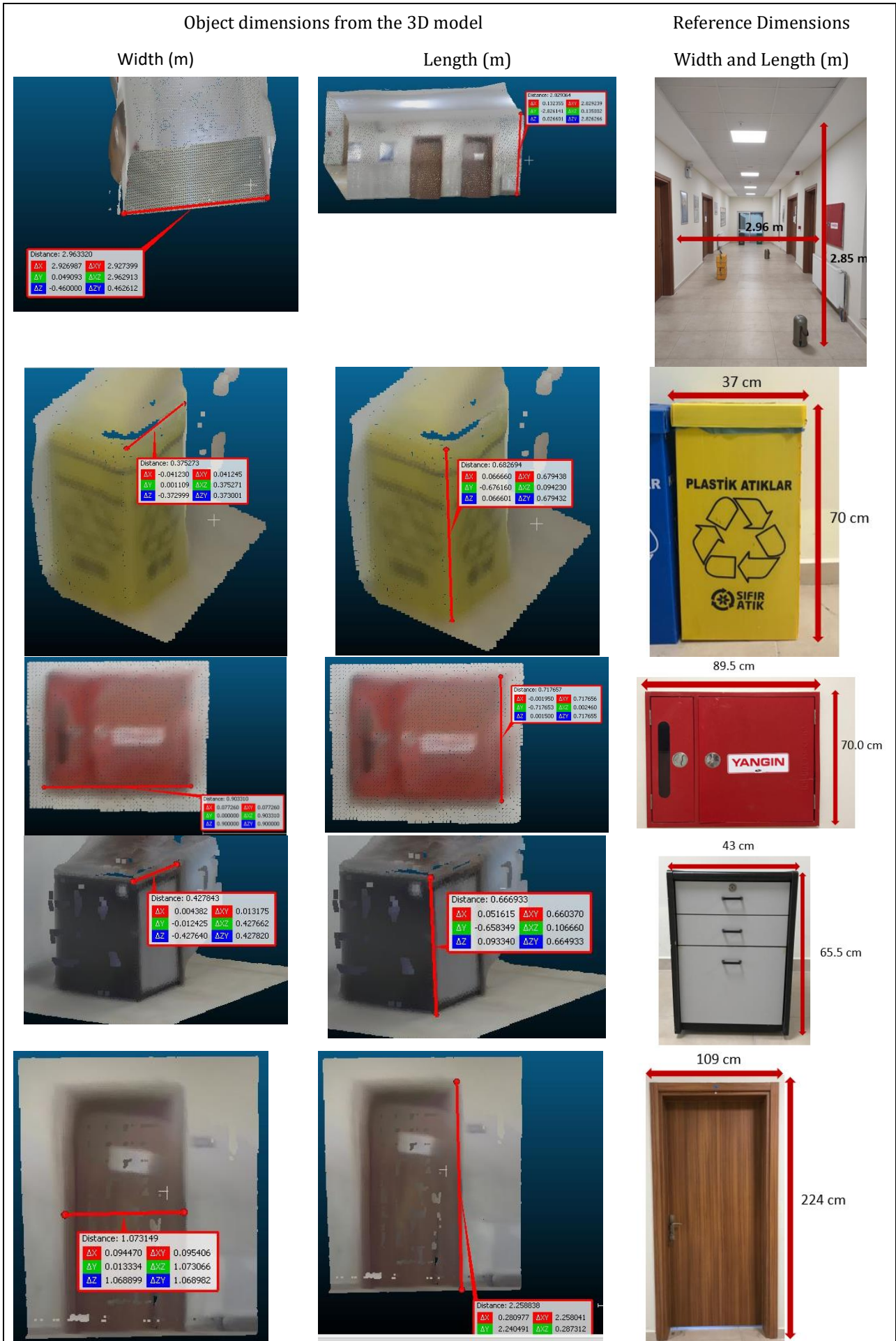
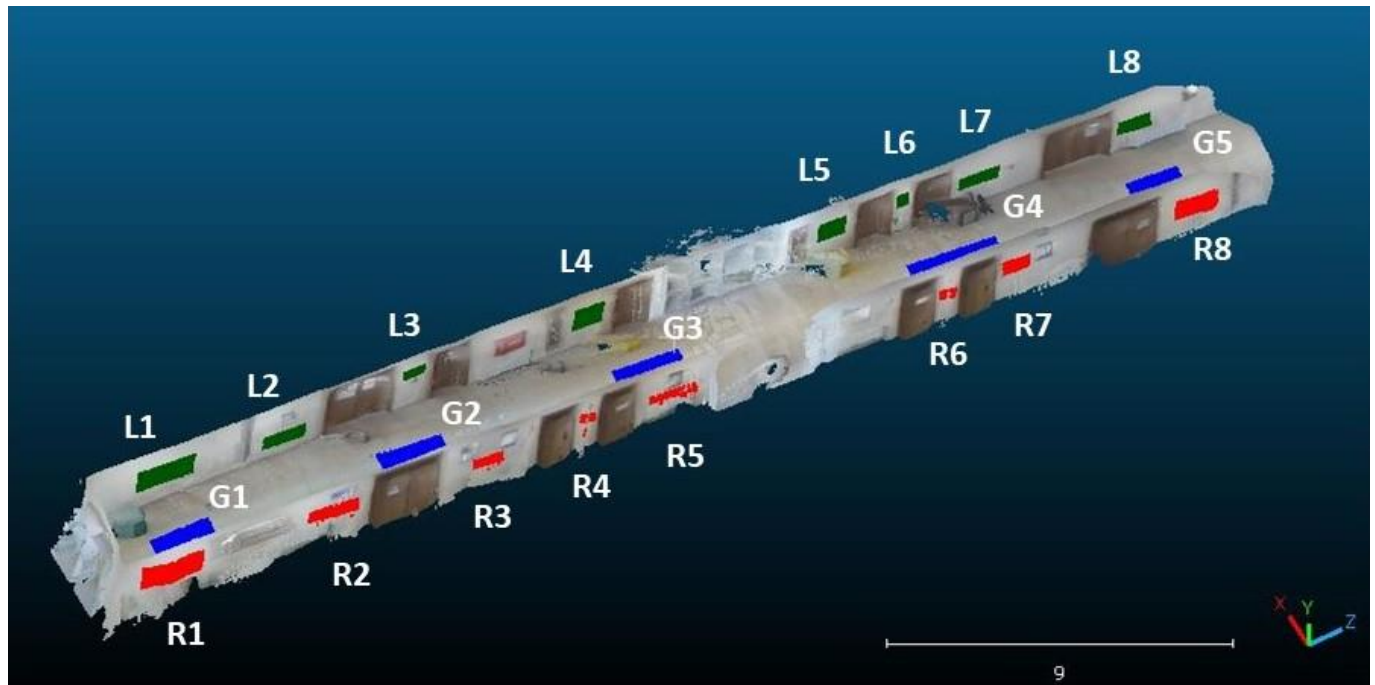


Figure 5. Reference and model dimensions of the selected objects in the 3D model

Table 1. Model accuracies

Object	Reference Dimensions		VSLAM Model Dimensions		Error (m)	
	Width (cm)	Length (cm)	Width (cm)	Length (cm)	Width (cm)	Length (cm)
Corridor	296	285	296	283	0	2
Recycling box	37	70	38	68	-1	2
Fire cabinet	90	70	90	72	0	-2
Cabinet	43	66	43	67	0	-1
Door	109	224	107	226	2	-2

**Figure 6.** The flat surfaces are chosen from the 3B model (green colour - left wall surfaces; red colour – right wall surfaces; blue colour – ground surfaces)

The Root Mean Square Error (RMSE) values for the plane of the surfaces are shown in Table 2. The RMSE values are calculated by using Equation 1.

$$RMSE = \sqrt{\frac{\sum e^2}{n}} \quad (1)$$

Here, e is the distance between the predicted and actual points, n is the number of points. The mean RMSE values were determined as 0.9 cm, 0.5 cm and 0.1 cm for the right, left and ground surfaces, respectively.

Table 2. RMSE values of the surfaces

Surface No	RMSE Values (cm)		
	Right	Left	Ground
1	2.4	0.1	0.1
2	0.8	0.4	0.1
3	0.5	0.2	0.1
4	0.4	0.4	0.2
5	1.2	0.5	0.1
6	0.4	0.3	-
7	0.6	0.7	-
8	0.8	1.5	-
Min	0.4	0.1	0.1
Max	2.4	1.5	0.2
Mean	0.9	0.5	0.1

4. Conclusion

In this study, the images collected with the stereo camera on the UGV were processed using the VSLAM method and produced a 3D model of the selected indoor area. Two analyses were carried out to test the accuracy of the produced 3D model. The dimensions of the objects on the 3D model were compared to their reference dimensions, and the model accuracy was determined as ± 2 cm or better. Secondly, surfaces with a homogenous distribution were chosen to test the surface's accuracy. The investigation yielded RMSE averages of 0.9, 0.5, and 0.1 cm for a total of 21 surfaces chosen from the floor, left wall, and right wall, respectively. These values indicate that the accuracy of the produced model is 1-2 cm and can be safely used in many indoor modelling applications. The colour of the model also makes the VSLAM method superior to other sensor models. However, VSLAM has many drawbacks as well as its benefits. The VSLAM method may have problems due to the low FOV of camera sensors. Because the height of the camera sensor in the UGV is very close to the floor and far away from the ceiling, the model's ceiling accuracy has been determined to be lower than that of the ground and side walls. It has also been determined that the light density in the environment may cause problems when the model is produced. In this study, the stereo camera was positioned in front of the UGV, so the details of the vehicle's direction of movement could be reflected well in the model, while the details in the opposite direction

of the UGV could not be obtained at the desired accuracy. This problem is believed to be solved with at least two cameras, which are integrated into the front and rear of the vehicle.

Acknowledgement

This study was funded by Ondokuz Mayıs University Scientific Research Projects (Projects No: PYO.MUH.1906.22.002, and PYO.MUH.1908.22.079). We also appreciate the LOCUS-TEAM members for their support during this study.

Author contributions

Simla Özbayrak: Field study, Data curation, Visualization, Writing-Original draft preparation
Veli İlçi: Conceptualization, Methodology, Validation, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

- Uçarlı, A. C., İlçi, V., Par, K., & Peker, A. U. (2022). Otonom araçlarda çoklu GNSS uydu sistemleri kullanımının konum doğruluğuna etkisinin araştırılması. *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 11(3), 672–680. <https://doi.org/10.28948/ngumuh.1082124>
- Li, N., Guan, L., Gao, Y., Du, S., Wu, M., Guang, X., & Cong, X. (2020). Indoor and outdoor low-cost seamless integrated navigation system based on the integration of INS/GNSS/LIDAR system. *Remote Sensing*, 12(19), 1–21. <https://doi.org/10.3390/rs12193271>
- Yurdakul, Ö., & Kalaycı, İ. (2022). The effect of GLONASS on position accuracy in CORS-TR measurements at different baseline distances. *International Journal of Engineering and Geosciences*, 7(3), 229–246. <https://doi.org/10.26833/ijeg.975204>
- Uçarlı, A. C., Demir, F., Erol, S., & Alkan, R. M. (2020). Farklı GNSS Uydu Sistemlerinin Hassas Nokta Konumlama (PPP) Tekniğinin Performansına Etkisinin İncelenmesi. *Geomatik*, 6(3), 247–258. <https://doi.org/10.29128/geomatik.779420>
- Ilci, V., & Toth, C. (2020). High definition 3D map creation using GNSS/IMU/LiDAR sensor integration to support autonomous vehicle navigation. *Sensors*, 20(3), 899. <https://doi.org/10.3390/s20030899>
- Atiz, O. F., Konukseven, C., Ogutcu, S., & Alcay, S. (2022). Comparative analysis of the performance of Multi-GNSS RTK: A case study in Turkey. *International Journal of Engineering and Geosciences*, 7(1), 67–80. <https://doi.org/10.26833/ijeg.878236>
- İlçi, V. (2020). CenterPoint RTX Teknolojisinin Doğruluk ve Tekrarlılığına Bilirliğinin Araştırılması. *Geomatik*, 5(1), 10–18. <https://doi.org/10.29128/geomatik.560026>
- Gurturk, M., & Ilci, V. (2022). The performance evaluation of PPK and PPP-based Loosely Coupled integration in wooded and urban areas. *Earth Sciences Research Journal*, 26(3), 211–220. <https://doi.org/10.15446/esrj.v26n3.100518>
- Kuang, J., Niu, X., & Chen, X. (2018). Robust pedestrian dead reckoning based on MEMS-IMU for smartphones. *Sensors*, 18(5), 1391. <https://doi.org/10.3390/s18051391>
- İlçi, V., Gülal, E., & Alkan, R. M. (2018). An investigation of different Wi-Fi signal behaviours and their effects on indoor positioning accuracy. *Survey Review*, 50(362), 404–411. <https://doi.org/10.1080/00396265.2017.1292672>
- İlçi, V., Gülal, E., & Alkan, R. M. (2020). Performance Comparison of 2.4 and 5 GHz WiFi Signals and Proposing a New Method for Mobile Indoor Positioning. *Wireless Personal Communications*, 110(3), 1493–1511. <https://doi.org/10.1007/s11277-019-06797-x>
- Wang, J., Wang, M., Yang, D., Liu, F., & Wen, Z. (2021). UWB positioning algorithm and accuracy evaluation for different indoor scenes. *International Journal of Image and Data Fusion*, 12(3), 203–225. <https://doi.org/10.1080/19479832.2020.1864788>
- Motroni, A., Buffi, A., & Nepa, P. (2021). A Survey on Indoor Vehicle Localization through RFID Technology. *IEEE Access*, 9, 17921–17942. <https://doi.org/10.1109/ACCESS.2021.3052316>
- Zhong, J., Li, M., Liao, X., & Qin, J. (2020). A real-time infrared stereo matching algorithm for RGB-D cameras' indoor 3D perception. *ISPRS International Journal of Geo-Information*, 9(8), 1–16. <https://doi.org/10.3390/ijgi9080472>
- Tarik, T., & ÖCALAN, T. (2020). PPK GNSS Sistemine Sahip İnsansız Hava Araçları İle Elde Edilen Fotogrametrik Ürünlerin Doğruluğunun Farklı Yaklaşımlarla İrdelenmesi. *Türkiye Fotogrametri Dergisi*, 2(53476), 22–28.
- Ceylan, M. C., & Uysal, M. (2021). İnsansız Hava Aracı ile Elde Edilen Veriler Yardımıyla Ağaç Çıkarımı. *Türkiye Fotogrametri Dergisi*, 3(1), 15–21. <https://doi.org/10.53030/tufod.912501>
- Guidi, F., Guerra, A., & Dardari, D. (2016). Personal mobile radars with millimeter-wave massive arrays for indoor mapping. *IEEE Transactions on Mobile Computing*, 15(6), 1471–1484. <https://doi.org/10.1109/TMC.2015.2467373>
- Petrlik, M., Krajník, T., & Saska, M. (2021). LIDAR-based Stabilization, Navigation and Localization for UAVs Operating in Dark Indoor Environments. 2021 International Conference on Unmanned Aircraft Systems, ICUAS 2021, (871479), 243–251. <https://doi.org/10.1109/ICUAS51884.2021.9476837>
- Nazari, S. W., Akarsu, V., & Yakar, M. (2023). Analysis of 3D Laser Scanning Data of Farabi Mosque Using Various Softwares. *Advanced LiDAR*, 3(1), 22–34.
- Zeybek, M., & Ediz, D. (2022). Detection of Road Distress with Mobile Phone LiDAR Sensors. *Advanced Geomatics*, 2(2), 48–53.

21. Bala, J. A., Adeshina, S. A., & Aibinu, A. M. (2022). Advances in Visual Simultaneous Localisation and Mapping Techniques for Autonomous Vehicles: A Review. *Sensors*, 22(22), 8943. <https://doi.org/10.3390/s22228943>
22. Ren, R., Fu, H., & Wu, M. (2019). Large-scale outdoor slam based on 2d lidar. *Electronics*, 8(6), 613. <https://doi.org/10.3390/electronics8060613>
23. Palomer, A., Ridao, P., & Ribas, D. (2019). Inspection of an underwater structure using point-cloud.pdf. *Journal of Field Robotics*, 36, 1333–1344. <https://doi.org/10.1002/rob.21907>
24. Hong, Z., Petillot, Y., Wallace, A., & Wang, S. (2022). RadarSLAM: A robust simultaneous localization and mapping system for all weather conditions. *International Journal of Robotics Research*, 41(5), 519–542. <https://doi.org/10.1177/02783649221080483>
25. Kazerouni, I. A., Fitzgerald, L., Dooly, G., & Toal, D. (2022). A survey of state-of-the-art on visual SLAM. *Expert Systems with Applications*, 205, 117734. <https://doi.org/10.1016/j.eswa.2022.117734>
26. Dai, Y., Wu, J., Wang, D., & Watanabe, K. (2023). A Review of Common Techniques for Visual Simultaneous Localization and Mapping. *Journal of Robotics*, 2023, 872822. <https://doi.org/10.1155/2023/8872822>
27. Mur-Artal, R., & Tardos, J. D. (2017). ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics*, 33(5), 1255–1262. <https://doi.org/10.1109/TRO.2017.2705103>
28. Taketomi, T., Uchiyama, H., & Ikeda, S. (2017). Visual SLAM algorithms: A survey from 2010 to 2016. *IPSP Transactions on Computer Vision and Applications*, 9(16). <https://doi.org/10.1186/s41074-017-0027-2>
29. Zhang, S., Zhao, S., An, D., Liu, J., Wang, H., Feng, Y., ... Zhao, R. (2022). Visual SLAM for underwater vehicles: A survey. *Computer Science Review*, 46, 100510. <https://doi.org/10.1016/j.cosrev.2022.100510>
30. Chen, W., Shang, G., Ji, A., Zhou, C., Wang, X., Xu, C., ... Hu, K. (2022). An Overview on Visual SLAM: From Tradition to Semantic. *Remote Sensing*, 14(13), 1–47. <https://doi.org/10.3390/rs14133010>
31. Fayyad, J., Jaradat, M. A., Gruyer, D., & Najjaran, H. (2020). Deep learning sensor fusion for autonomous vehicle perception and localization: A review. *Sensors*, 20(15), 4220. <https://doi.org/10.3390/s20154220>
32. Debeunne, C., & Vivet, D. (2020). A review of visual-lidar fusion based simultaneous localization and mapping. *Sensors*, 20(7), 2068. <https://doi.org/10.3390/s20072068>
33. Wang, H., Wang, C., & Xie, L. (2021). Intensity-SLAM: Intensity Assisted Localization and Mapping for Large Scale Environment. *IEEE Robotics and Automation Letters*, 6(2), 1715–1721. <https://doi.org/10.1109/LRA.2021.3059567>
34. Gostar, A. K., Fu, C., Chuah, W., Hossain, M. I., Tennakoon, R., Bab-Hadiashar, A., & Hoseinnezhad, R. (2019). State transition for statistical SLAM using planar features in 3D point clouds. *Sensors*, 19(7), 1614. <https://doi.org/10.3390/s19071614>
35. Takleh, T. T. O., Bakar, N. A., Rahman, S. A., Hamzah, R., & Aziz, Z. A. (2018). A brief survey on SLAM methods in autonomous vehicle. *International Journal of Engineering and Technology (UAE)*, 7(4), 38–43. <https://doi.org/10.14419/ijet.v7i4.27.22477>
36. Wen, W., Zhang, G., & Hsu, L. T. (2020). Object-Detection-Aided GNSS and Its Integration with Lidar in Highly Urbanized Areas. *IEEE Intelligent Transportation Systems Magazine*, 12(3), 53–69. <https://doi.org/10.1109/MITS.2020.2994131>
37. Min, H., Wu, X., Cheng, C., & Zhao, X. (2019). Kinematic and Dynamic Vehicle Model-Assisted Global Positioning Method for Autonomous Vehicles with Low-Cost GPS/Camera/In-Vehicle Sensors. *Sensors*, 19(24), 5430. <https://doi.org/10.3390/s19245430>
38. Van Brummelen, J., O'Brien, M., Gruyer, D., & Najjaran, H. (2018). Autonomous vehicle perception: The technology of today and tomorrow. *Transportation Research Part C: Emerging Technologies*, 89(July 2017), 384–406. <https://doi.org/10.1016/j.trc.2018.02.012>
39. Zhang, Y., Chen, L., XuanYuan, Z., & Tian, W. (2020). Three-Dimensional Cooperative Mapping for Connected and Automated Vehicles. *IEEE Transactions on Industrial Electronics*, 67(8), 6649–6658. <https://doi.org/10.1109/TIE.2019.2931521>
40. Dwijotomo, A., Rahman, M. A. A., Ariff, M. H. M., Zamzuri, H., & Azree, W. M. H. W. (2020). Cartographer SLAM method for optimization with an adaptive multi-distance scan scheduler. *Applied Sciences*, 10(1), 347. <https://doi.org/10.3390/app10010347>
41. Wu, Y., Li, Y., Li, W., Li, H., & Lu, R. (2021). Robust Lidar-Based Localization Scheme for Unmanned Ground Vehicle via Multisensor Fusion. *IEEE Transactions on Neural Networks and Learning Systems*, 32(12), 5633–5643. <https://doi.org/10.1109/TNNLS.2020.3027983>
42. Zou, D., Tan, P., & Yu, W. (2019). Collaborative visual SLAM for multiple agents: A brief survey. *Virtual Reality and Intelligent Hardware*, 1(5), 461–482. <https://doi.org/10.1016/j.vrih.2019.09.002>
43. Tourani, A., Bavle, H., Sanchez-Lopez, J. L., & Voos, H. (2022). Visual SLAM: What Are the Current Trends and What to Expect? *Sensors*, 22(23), 9297. <https://doi.org/10.3390/s22239297>
44. Chiang, K. W., Tsai, G. J., Li, Y. H., Li, Y., & El-Sheimy, N. (2020). Navigation engine design for automated driving using INS/GNSS/3D LiDAR-SLAM and integrity assessment. *Remote Sensing*, 12(10), 1564. <https://doi.org/10.3390/rs12101564>
45. Cheng, J., Zhang, L., Chen, Q., Hu, X., & Cai, J. (2022). A review of visual SLAM methods for autonomous driving vehicles. *Engineering Applications of Artificial Intelligence*, 114(October 2021), 104992. <https://doi.org/10.1016/j.engappai.2022.104992>

46. Lin, X., Wang, F., Yang, B., & Zhang, W. (2021). Autonomous vehicle localization with prior visual point cloud map constraints in GNSS-challenged environments. *Remote Sensing*, 13(3), 506. <https://doi.org/10.3390/rs13030506>
47. Razali, M. R., Athif, A., Faudzi, M., & Shamsudin, A. U. (2022). Visual Simultaneous Localization and Mapping: A review. *PERINTIS eJournal*, 12(1), 23–34.
48. Lin, X., Yang, B., Wang, F., Li, J., & Wang, X. (2021). Dense 3D surface reconstruction of large-scale streetscape from vehicle-borne imagery and LiDAR. *International Journal of Digital Earth*, 14(5), 619–639. <https://doi.org/10.1080/17538947.2020.1862318>
49. Sharafutdinov, D., Griguletskii, M., Kopanev, P., Kurenkov, M., Ferrer, G., Burkov, A., ... Tsetsserkou, D. (2023). Comparison of modern open-source Visual SLAM approaches. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 107(43), 1-22. <https://doi.org/10.1007/s10846-023-01812-7>
50. Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 1052–1067. <https://doi.org/10.1109/TPAMI.2007.1049>
51. Klein, G., & Murray, D. (2007). Parallel tracking and mapping for small AR workspaces. 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR, 225–234. <https://doi.org/10.1109/ISMAR.2007.4538852>
52. Beghdadi, A., & Malle, M. (2022). A comprehensive overview of dynamic visual SLAM and deep learning: concepts, methods and challenges. *Machine Vision and Applications*, 33, 54. <https://doi.org/10.1007/s00138-022-01306-w>
53. Newcombe, R. A., Lovegrove, S. J., & Davison, A. J. (2011). DTAM: Dense tracking and mapping in real-time. *Proceedings of the IEEE International Conference on Computer Vision*, 2320–2327. <https://doi.org/10.1109/ICCV.2011.6126513>
54. Engel, J., Schöps, T., Cremers, D. (2014). LSD-SLAM: Large-Scale Direct Monocular SLAM. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) *Computer Vision – ECCV 2014*. *ECCV 2014. Lecture Notes in Computer Science*, vol 8690. Springer, Cham. https://doi.org/10.1007/978-3-319-10605-2_54
55. Duan, C., Junginger, S., Huang, J., Jin, K., & Thuro, K. (2019). Deep Learning for Visual SLAM in Transportation Robotics: A review. *Transportation Safety and Environment*, 1(3), 177–184. <https://doi.org/10.1093/tse/tdz019>
56. Yalcin, H., & Cilasun, M. H. (2016). Derin Öğrenme Tabanlı Otonom Yön Belirleme. 2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings, 1645–1648. <https://doi.org/10.1109/SIU.2016.7496072>
57. Tseng, P. Y., Lin, J. J., Chan, Y. C., & Chen, A. Y. (2022). Real-time indoor localization with visual SLAM for in-building emergency response. *Automation in Construction*, 140, 104319. <https://doi.org/10.1016/j.autcon.2022.104319>
58. Cui, L., & Ma, C. (2019). SOF-SLAM: A Semantic Visual SLAM for Dynamic Environments. *IEEE Access*, 7, 166528–166539. <https://doi.org/10.1109/ACCESS.2019.2952161>



© Author(s) 2024. This work is distributed under <https://creativecommons.org/licenses/by-sa/4.0/>