Integrated Robotics System For Autonomous Mapping And 3D Scene Reconstruction In Structural Health Monitoring Applications

Mohsen Azimi[1*], Sina Tavasoli[1], Kasra Jabbari[2], T. Y. Yang[3],

[1]PhD Candidate, Department of Civil Engineering, The University of British Columbia, Vancouver, BC, Canada
[2]Research Assistant, Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran
[3]Professor, Department of Civil Engineering, The University of British Columbia, Vancouver, BC, Canada
[*]mohsen.azimi@ubc.ca (Corresponding Author)

ABSTRACT

This paper presents a novel integrated unmanned aerial vehicle (UAV) and unmanned ground vehicle (UGV) system for structural health monitoring (SHM). In the aftermath of a disaster, efficient and accurate data collection and processing is crucial for assessing the structural integrity of buildings. However, current limitations in SHM systems include a lack of autonomy, leading to manual remote operations that are labor-intensive and time-consuming. To address that, the proposed system utilizes a wireless robot operating system (ROS) network to coordinate navigation and data collection of the UAV and UGV. Additionally, improved simultaneous localization and mapping (SLAM) algorithm is employed to generate a 2D map of the environment. High-resolution data collected by the UAV is used for 3D reconstruction at the scene and component level, making it suitable for post-earthquake assessments. The effectiveness of the proposed system was validated through experiments using a custom-built ground robot and a micro aerial vehicle for indoor inspections. Results from both qualitative and quantitative analyses demonstrate the potential for this system to enhance the efficiency and effectiveness of post-earthquake SHM and remote sensing through the use of cost-effective and portable robots.

Keywords: Cyber-physical system, structural health monitoring, simultaneous localization and mapping (SLAM), Neural Radiance Field (NeRF), 3D reconstruction

INTRODUCTION

The advances in edge computing have catalyzed an advancement in the development and deployment of AI-based robotic applications, which substantially reduces the cost of data transmission and cloud computing [1]. The implementation of such an approach has proven to be highly effective in supporting applications that require real-time decision-making with low latency. One such application is autonomous structural health monitoring and damage detection using unmanned aerial vehicles (UAVs) and unmanned ground vehicles (UGVs) [2-5]. The use of traditional methods for path-planning and data collection may not always accurately capture the unstructured, dynamic, and complex nature of inspection sites after an event such as an earthquake, and human-based inspections may leave workers at risk of potential hazards [6]. In addition, manual approaches that largely rely on manual inspections, are time-consuming and subjective and thus not feasible for large sites [7].

The emerging technologies have paved the way for the development of new efficient cyber-physical systems that reduces the cost and time for a full inspection of buildings and infrastructures. Several researchers have investigated computer-aided and artificial intelligence (AI)-based approaches for intelligent monitoring, such as building information modeling (BIM)-based methods [1, 8-10]. Zhang et al. [11] developed a BIM-based proactive approach for hazard identification, visualization, and safety planning using construction schedules. However, without BIM support, such techniques cannot be scaled or generalized.

The integration of edge computing and AI-based computer vision into cyber-physical systems (e.g., robots) provides a promising solution to address the demands for post-event inspections. Recently, several studies have been conducted to explore automated approaches in post-earthquake inspections and damage detection [12-14]. One of the most popular methods is computer vision-aided frameworks for the identification of damaged structural components using stationary cameras or UAVs

[15, 16]. Such methods usually use convolutional neural networks (CNNs) for the recognition and localization of visual damages [17-19].

Due to the high payload capacity and mobility of Ground robots, such as UGVs, they offer more feasible and robust solutions for various applications, including structural health monitoring, surveillance, and construction site monitoring. Therefore, further research is required to develop robust and reliable UGVs that can efficiently navigate in a dynamic and complex environment and perform tasks autonomously. Traditionally, UGVs were designed to process sensor data using mounted high-performance computers. An efficient solution for indoor monitoring using UGVs is utilizing the Robot Operating System (ROS) [20], a flexible and scalable framework that facilitates the development of robotics applications. ROS can be used to develop various functionalities of UGVs, including collision-free navigation and mapping, as well as vision-based object detection and tracking.

Kim et al. [21] propose a SLAM-based UGV system for autonomous data collection using an object-detection algorithm. Such systems are not efficient for indoor sites due to their size as well as the cost of production, which hampered deployment for indoor inspections. Recently, Ojha et al. [6, 22] investigated the feasibility and performance of a small and low-cost UGV equipped with a camera and a light detection and ranging (LiDAR) sensor for the detection and localization of fall hazards in indoor construction sites. They utilized a low-resolution RGB camera for detecting fall hazards, alongside mechanical and cost-efficient LiDAR sensors for simultaneous localization and mapping (SLAM). However, despite their successful results, the system's performance is still limited due to hardware and computation constraints. Specifically, their use of a single RGB camera with limited resolution and field of view for detection has inherent limitations such as susceptibility to environmental factors like lighting and shadows, making accurate geo-tagging of objects of interest challenging.

To address the limitations of the previous research, this study proposes a novel integrated cyber-physical system that consists of ground and aerial robots for efficient data collection and monitoring. For this purpose, a high-performance single-board computer is used to deploy the lightweight navigation and detection algorithms for the UGV. A wireless ROS network was established to connect UAV and UGV to a host portable computer where the AI models were deployed for faster processing of high-resolution images. To generate the map of the environment, an integrated method including Hector SLAM mapping, rapidly exploring random-tree (RRT) and Voronoi Diagram (VD)-based path-planning algorithms were developed. To assess the performance and validate the feasibility and efficacy of the proposed cyber-physical system, an experimental study was conducted at an indoor structural laboratory at the University of British Columbia. The recorded videos by the UAV were used to reconstruct the scene in 3D using a NeRF algorithm [23].

METHODOLOGY

Overview

The proposed pipeline for autonomous 2D mapping and 3D scene reconstruction using integrated aerial-ground robotics is shown in Figure 1. Two fully programmable aerial and ground robots were used in this study. A wireless ROS network was established to efficiently communicate with each robot through standard IP-based data transmission protocols. First, the ground robot starts generating the 2D grid map of the surroundings. Based on the detected obstacles, the flight path for the UAV was generated and the recorded video frames were collected and processed for camera pose detection. This information was fed to the NeRF model to generate the 3D scene of the scanned environment. Details of the hardware and software of the proposed system are described in the following sections.
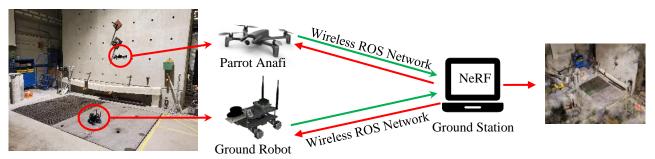

Figure 1. The proposed pipeline for integrated aerial-ground robotics for mapping and data collection.

Hardware components

The proposed cyber-physical system consists of an affordable micro-aerial vehicle (MAV) and a cost-effective custom-built robot that consists of two mechanical LiDAR sensors, an NVIDIA Jetson Nano (4GB), an Arduino Mega 2560, power supply units, and the omnidirectional UGV chassis with 12 V mecanum wheels.

To measure the distance data points between the robot and the objects (i.e., laser scanning), a cost-efficient mechanical LiDAR sensor (RPLiDAR A1M8) was utilized. This sensor is capable of scanning and generating 2D point sets to measure the distance from objects within 12 meters and 360° [24]. The sensor provides accurate estimation by emitting infrared pulses and receiving the reflected light from objects in its vicinity, which is essential for collision-free navigation and mapping in diverse environments.

The Jetson Nano (4GB) [25], a powerful and low-power consumption single-board computer with Ubuntu 18.04 LTS and ROS Melodic, is used to connect the UGV to the ROS network for efficient navigation and data collection. This compact edge computer features a quad-core ARM Cortex-A57 CPU, 4GB of RAM, and NVIDIA Maxwell GPU with 128 NVIDIA CUDA cores, which is ideal for running a wide range of light-weight computer vision, AI, and robotics applications including Robot Operating System (ROS) [20] and OpenCV [26]. To establish a ROS network, the Intel-AC8265 dual-mode wireless networking card module provides a much faster and more reliable Wi-Fi and Bluetooth connection compared to simpler single-board computers such as Raspberry Pi 4 [27]. Four encoder gear motors are used to move the robot in the desired direction. The 4WD UGV uses mecanum wheels to provide more moving options for navigation and mapping without changing the viewport. Four of major directions are depicted in Figure 2. Finally, A portable laptop, with 16 GB RAM, Intel Core i7-12700H CPU, and NVIDIA RTX 3070 GPU, is used as the Operation Control Unit (OCU).
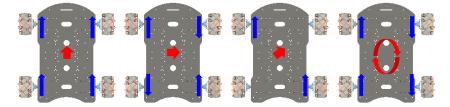

Figure 2. The working principle and motion of the 4WD mecanum wheel UGV.

Simultaneous localization and mapping (SLAM)

In the case of a mobile robot, simultaneous localization and mapping (SLAM) involve an iterative process that fuses diverse measurements from onboard sensors to enable both localization and mapping. This process is continuous and operates in real-time to update the robot's internal map while simultaneously determining its position within that map. Among the popular Simultaneous Localization and Mapping (SLAM) algorithms, Hector SLAM was used in this study because it does not require odometer information. The process of scan matching for the Hector SLAM algorithm involves optimizing the alignment of the 2D laser scan endpoints at the current time with the global map. A GausspNewton optimization approach is used to predict the robot pose. Hector SLAM utilizes a rigid body transformation:

$$\xi = \left(p_x, p_y, \psi\right)^T \tag{1}$$

where p_x, p_y, ψ are the coordinates and the yaw angle of the robot in the robot coordinate system, and it is minimized in order to achieve the best alignment between the laser scan and the map.

$$\xi^* = argmin_\xi \sum_{i=1}^{n}[1 - \boldsymbol{M}(\boldsymbol{S_i}(\xi)]^2 \tag{2}$$

where S_i represents the world coordinates of the scan endpoints as described by Wang [28] as follows:

$$\boldsymbol{S_i}(\xi) = \begin{Bmatrix} \boldsymbol{S_{i,x}}(\xi) \\ \boldsymbol{S_{i,y}}(\xi) \end{Bmatrix} = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \begin{Bmatrix} s_{i,x} \\ s_{i,y} \end{Bmatrix} + \begin{Bmatrix} p_x \\ p_y \end{Bmatrix} \tag{3}$$

where $\boldsymbol{S_{i,x}}(\xi)$ and $\boldsymbol{S_{i,y}}(\xi)$ are the ground coordinates of the ith endpoint scanned by the LiDAR sensor. Assuming small movement of the robot, $\nabla\xi$, we have:

$$\nabla\xi = H^{-1} \sum_{i=1}^{n} \left[\nabla M\left(S_i(\xi)\right) \frac{\partial S_i(\xi)}{\partial \xi}\right]^T [1 - MS_i(\xi)] \tag{4}$$

where, and H is the Hessian matrix,

$$H = \left[\nabla \mathrm{M}\big(\mathrm{S_i}(\xi)\big)\frac{\partial S_i(\xi)}{\partial \xi}\right]^T \left[\nabla \mathrm{M}\big(\mathrm{S_i}(\xi)\big)\frac{\partial S_i(\xi)}{\partial \xi}\right] \tag{5}$$

and $\nabla \mathrm{M}\big(S_i(\xi)\big)$ is the gradient of the current map.

Path planning for UAV

The Voronoi diagram is a mathematical construct that partitions a map into different regions based on the proximity of the points. In the context of path planning for an autonomous UAV-based data collection, Voronoi diagrams can be used to construct a roadmap of the environment, with each Voronoi region representing the set of points that are closest to a specific obstacle or feature in the environment. In this study, the Voronoi diagram is utilized to generate a roadmap of the environment based on the SLAM-generated occupancy grid map. The roadmap is then used to plan the UAV's flight path through the environment, to minimize the distance traveled while avoiding collisions with obstacles.
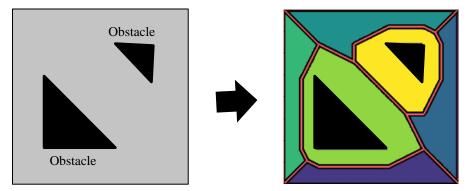

Figure 3. Voronoi diagram generation for collision-free path planning for UAV.

3D Reconstruction

In recent years, several innovative approaches have been proposed to execute 3D reconstruction utilizing Structural-from-Motion (SfM) techniques. The process of 3D reconstruction through SfM comprises several stages, including feature extraction, feature matching, geometric verification, and reconstruction [29]. These stages play a crucial role in producing an accurate and complete 3D representation of the target object or scene, making them essential components of the 3D reconstruction. Neural field (NF) is a term that was recently used to represent a neural network for parametrizing a signal [30]. The fully connected (FC) neural networks are utilized in most variations of NFs to encode the properties of scenes or objects, with the exception that in contrast to the standard ML models, NFs are intentionally overfitted to encode the scene's features as weights. In traditional approaches, voxel grids and polygon mesh were used in order to store 3D objects and scenes, which were not efficient. Due to the fact that NFs are differentiable and continuous, they have gained significant attention. Because they provide an efficient way to represent a 3D object. The most prominent models from the NF family are Neural Radiance Fields (NeRFs).

NeRF was introduced by Mildenhall et al. [23] for synthesizing photo-realistic 3D reconstruction of complex scenes using sets of 2D images from different angles. NeRFs are state-of-the-art NFs that learn to represent a 3D scene as a continuous function, and thus they are different from the traditional methods that require manual texture mapping and lighting. In contrast to the traditional 3D reconstruction methods, NeRFs do not rely on geometric primitives, such as triangles and voxels; instead, they train a neural network to directly model the scene's appearance. Another advantage of NFs lies in their ability to generate distinct representations of a point when observed from different angles, enabling the capture of various lighting effects such as reflections and transparencies. As a result, neural fields are suitable for rendering different views of the same scene, making them a superior representation when compared to voxel grids or meshes. As a result, NeRFs have emerged as a promising method for generating highly accurate and realistic 3D digital scenes and have been successfully applied to a wide range of applications [31, 32]. As shown in Figure 4, a NeRF model takes a single continuous 5D coordinate (i.e., location and viewing angle) and outputs the corresponding color intensities and the volume density.
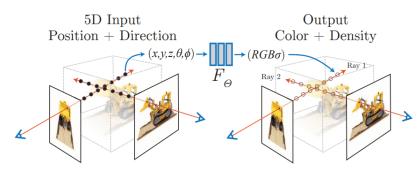
Figure 4. The Neural Radiance Fields [23].

EXPERIMENTAL RESULTS AND DISCUSSION

In this study, as the first step, the occupancy grid map of the environment was generated using the LiDAR sensor mounted on a custom-built 4WD ground robot and utilizing an autonomous exploration as well as the Hector SLAM algorithm. To efficiently distribute the computational resources, the navigation algorithm and laser scanning (RPLiDAR) ROS nodes were deployed on the robot, while the SLAM algorithm and visualization tools were developed and launched from the ground station computer (i.e., the ROS core). The ROS topic tree as well as the generated grid map of the environment is shown in Figure 6. This map provides an accurate representation of the surrounding obstacles and enables the UAV to navigate autonomously. To measure the accuracy of the map the Hausdorff distance (HD) metric was used [33-35]. The Voronoi diagram, Figure 6(a), is used to generate a collision-free path for the UAV (Figure 6(b)).
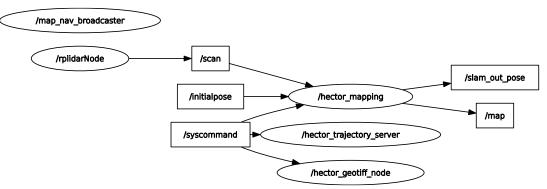

Figure 5. The ROS topic tree for Hector SLAM.

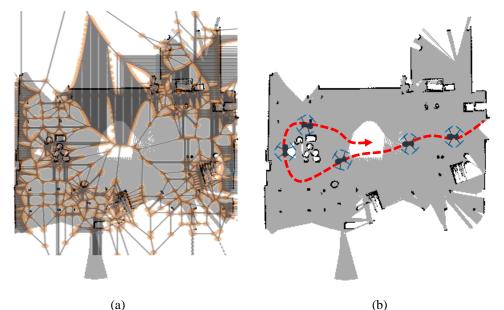(a)   (b)

Figure 6. UAV path planning: (a) the Voronoi diagram, (b) safe flight path based on the Voronoi diagram.

Once the recorded frames by the UAV were received by the host computer, the NeRF algorithm was used to generate the 3D representation of the scanned area. Due to the limitations of the graphical processing unit, only around 100 video frames were extracted for reconstruction purposes. The camera information was generated using COLMAP [29]. Using the Segment Anything (SAM) segmentation model [36], an original input video frame from the UAV was segmented and compared with the segmentation results for the synthetic new viewpoint generated using the 3D reconstruction with NeRF. As shown in Figure 8, the segmentation result demonstrates that the reconstructed scene is highly accurate and the fine details could be used for accurate damage assessment in structural health monitoring applications.
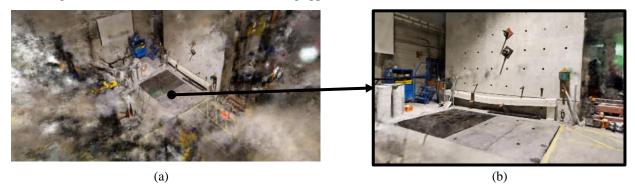

(a) (b)

Figure 7. 3D scene reconstruction of the structural laboratory at UBC (a), with the zoomed view (b).
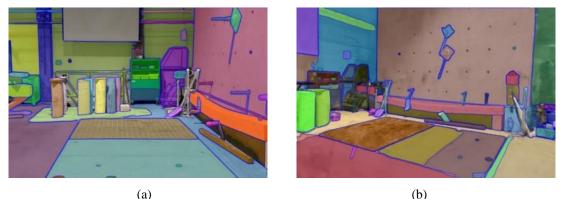

(a) (b)

Figure 8. Segmentation results for (a) the UAV-recorded video and (b) the synthetic novel view from NeRF.

CONCLUSIONS

This study presents a framework for the integrated robotic system. An affordable custom-built ground robot and a programmable aerial quadcopter were used to achieve the optimal solution for 2D mapping and 3D scene reconstruction. For the 2D mapping, the ground robot was used and the Hector SLAM algorithm was utilized to generate the grid map of obstacles. The optimal flight path was generated using the Voronoi diagram for the UAV. The 3D scene of the inspected area was generated using a state-of-the-art NeRF model. The obtained results for 2D mapping and 3D reconstruction demonstrate the effectiveness of the proposed system for real-world indoor inspections and structural health monitoring.

REFERENCES

[1] Y. Xiao *et al.*, "Autonomous Inspection and Construction of Civil Infrastructure Using Robots," in *Automation in Construction toward Resilience: Robotics, Smart Materials & Intelligent Systems*, E. Noroozinejad Farsangi *et al.* Eds.: CRC Press, 2023, ch. 1.

[2] S. Shayesteh and H. Jebelli, "Toward human-in-the-loop construction robotics: Understanding workers' response through trust measurement during human-robot collaboration," in *Construction Research Congress 2022*, 2022, pp. 631-639.

[3] S. Shayesteh, A. Ojha, and H. Jebelli, "Workers' Trust in Collaborative Construction Robots: EEG-Based Trust Recognition in an Immersive Environment," *Automation and robotics in the architecture, engineering, and construction industry,* pp. 201-215, 2022.

[4] M. Azimi, A. D. Eslamlou, and G. Pekcan, "Data-Driven Structural Health Monitoring and Damage Detection through Deep Learning: State-of-the-Art Review," *Sensors,* vol. 20, no. 10, p. 2778, 2020.

[5] S. Poorghasem and Y. Bao, "Review of robot-based automated measurement of vibration for civil engineering structures," *Measurement,* p. 112382, 2022.

[6] A. Ojha, Y. Liu, S. Shayesteh, H. Jebelli, and W. E. Sitzabee, "Affordable Multiagent Robotic System for Same-Level Fall Hazard Detection in Indoor Construction Environments," *Journal of Computing in Civil Engineering,* vol. 37, no. 1, p. 04022042, 2023.

[7] A. Albert, M. R. Hallowell, M. Skaggs, and B. Kleiner, "Empirical measurement and improvement of hazard recognition skill," *Safety science,* vol. 93, pp. 1-8, 2017.

[8] M. Kiviniemi, K. Sulankivi, K. Kähkönen, T. Mäkelä, and M.-L. Merivirta, "BIM-based safety management and communication for building construction," 2011.

[9] H. Pan, M. Azimi, F. Yan, and Z. Lin, "Time-Frequency-Based Data-Driven Structural Diagnosis and Damage Detection for Cable-Stayed Bridges," *Journal of Bridge Engineering,* vol. 23, no. 6, 2018.

[10] M. Azimi and G. Pekcan, "Structural health monitoring using extremely compressed data through deep learning," *Computer-Aided Civil and Infrastructure Engineering,* vol. 35, no. 6, pp. 597-614, 2020.

[11] S. Zhang, K. Sulankivi, M. Kiviniemi, I. Romo, C. M. Eastman, and J. Teizer, "BIM-based fall hazard identification and prevention in construction safety planning," *Safety science,* vol. 72, pp. 31-45, 2015.

[12] M. R. Al Zaabi, H. Alhendi, H. Alkhoori, M. Alkhoori, S. Alkhemriri, and Y. Abu-Kheil, "Automatic Site Inspection System in Construction Sites (ICI–Intelligent Camera Inspection)," in *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, 2022: IEEE, pp. 1-6.

[13] I. Jeelani, K. Asadi, H. Ramshankar, K. Han, and A. Albert, "Real-time vision-based worker localization & hazard detection for construction," *Automation in Construction,* vol. 121, p. 103448, 2021.

[14] L. Katebi, M. Tehranizadeh, and N. Mohammadgholibeyki, "A generalized flexibility matrix-based model updating method for damage detection of plane truss and frame structures," *Journal of Civil Structural Health Monitoring,* vol. 8, pp. 301-314, 2018.

[15] D. Kang and Y.-J. Cha, "Damage detection with an autonomous UAV using deep learning," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018*, 2018, vol. 10598: International Society for Optics and Photonics, p. 1059804.

[16] B. Lei, N. Wang, P. Xu, and G. Song, "New crack detection method for bridge inspection using UAV incorporating image processing," *Journal of Aerospace Engineering,* vol. 31, no. 5, p. 04018058, 2018.

[17] L. Deng, H.-H. Chu, P. Shi, W. Wang, and X. Kong, "Region-Based CNN Method with Deformable Modules for Visually Classifying Concrete Cracks," *Applied Sciences,* vol. 10, no. 7, p. 2528, 2020.

[18] H. Liu and Y. Zhang, "Image-driven structural steel damage condition assessment method using deep learning algorithm," *Measurement,* vol. 133, pp. 168-181, 2019/02/01/ 2019, doi: https://doi.org/10.1016/j.measurement.2018.09.081.

[19] Y. Xu, S. Wei, Y. Bao, and H. Li, "Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network," *Structural Control and Health Monitoring,* vol. 26, no. 3, p. e2313, 2019, doi: 10.1002/stc.2313.

[20] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," in *ICRA workshop on open source software*, 2009, vol. 3, no. 3.2: Kobe, Japan, p. 5.

[21] P. Kim, J. Chen, J. Kim, and Y. K. Cho, "SLAM-driven intelligent autonomous mobile robot navigation for construction applications," in *Advanced Computing Strategies for Engineering: 25th EG-ICE International Workshop 2018, Lausanne, Switzerland, June 10-13, 2018, Proceedings, Part I 25*, 2018: Springer, pp. 254-269.

[22] A. Ojha, Y. Liu, S. Shayesteh, and H. Jebelli, "Developing an affordable robotic system for automated fall hazard detection and localization in indoor construction environments," in *Computing in Civil Engineering 2021*, 2021, pp. 1041-1049.

[23] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM,* vol. 65, no. 1, pp. 99-106, 2021.

[24] M. A. Markom, A. H. Adom, E. S. M. M. Tan, S. A. A. Shukor, N. A. Rahim, and A. Y. M. Shakaff, "A mapping mobile robot using RP Lidar scanner," in *2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, 2015: IEEE, pp. 87-92.

[25] A. O. Prasad *et al.*, "Design and development of software stack of an autonomous vehicle using robot operating system," *Robotics and Autonomous Systems,* vol. 161, p. 104340, 2023.

[26] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[27] E. Bauer, B. G. Cangan, and R. K. Katzschmann, "Autonomous Vision-based Rapid Aerial Grasping," *arXiv preprint arXiv:2211.13093,* 2022.

[28] Z.-s. Wang, "Implementation of Odometry with EKF for Localization of Hector SLAM Method," NSYSU, 2016.

[29] J. L. Schonberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4104-4113.

[30] Y. Xie *et al.*, "Neural fields in visual computing and beyond," in *Computer Graphics Forum*, 2022, vol. 41, no. 2: Wiley Online Library, pp. 641-676.

[31] K. Gao, Y. Gao, H. He, D. Lu, L. Xu, and J. Li, "Nerf: Neural radiance field in 3d vision, a comprehensive review," *arXiv preprint arXiv:2210.00379,* 2022.

[32] F. Zhu, S. Guo, L. Song, K. Xu, and J. Hu, "Deep Review and Analysis of Recent NeRFs," *APSIPA Transactions on Signal and Information Processing,* vol. 12, no. 1, 2023.

[33] C. Knauer, M. Löffler, M. Scherfenberg, and T. Wolle, "The directed Hausdorff distance between imprecise point sets," *Theoretical Computer Science,* vol. 412, no. 32, pp. 4173-4186, 2011.

[34] D. M. Turnage, "Simulation results for localization and mapping algorithms," in *2016 Winter Simulation Conference (WSC)*, 2016: IEEE, pp. 3040-3051.

[35] M. Walter, F. Hover, and J. Leonard, "SLAM for ship hull inspection using exactly sparse extended information filters," in *2008 IEEE international conference on robotics and automation*, 2008: IEEE, pp. 1463-1470.

[36] A. Kirillov *et al.*, "Segment Anything," *arXiv preprint arXiv:2304.02643,* 2023.