Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-01 | March-2023

Vision-Based Monocular SLAM in Micro Aerial Vehicle

Saira Khan[1], Abdul Basit[2]
[1,2]Department of Computer Science and Information Technology,
University of Balochistan, Quetta, Pakistan.
drabasit@um.uob.edu.pk

DOI: 10.5281/zenodo.8200317

ABSTRACT

Micro Aerial Vehicles (MAVs) are popular for their efficiency, agility, and lightweights. They can navigate in dynamic environments that cannot be accessed by humans or traditional aircraft. These MAVs rely on GPS and it will be difficult for GPS-denied areas where it is obstructed by buildings and other obstacles. Simultaneous Localization and Mapping (SLAM) in an unknown environment can solve the aforementioned problems faced by flying robots. A rotation and scale invariant visual-based solution, oriented fast and rotated brief (ORB-SLAM) is one of the best solutions for localization and mapping using monocular vision.

In this paper, an ORB-SLAM3 has been used to carry out the research on localizing micro-aerial vehicle Tello and mapping an unknown environment. The effectiveness of ORB-SLAM3 was tested in a variety of indoor environments. An integrated adaptive controller was used for an autonomous flight that used the 3D map, produced by ORB-SLAM3 and our proposed novel technique for robust initialization of the SLAM system during flight. The results show that ORB-SLAM3 can provide accurate localization and mapping for flying robots, even in challenging scenarios with fast motion, large camera movements, and dynamic environments. Furthermore, our results show that the proposed system is capable of navigating and mapping challenging indoor situations.

Keywords: Simultaneous localization and mapping (SLAM); Monocular SLAM; micro ariel vehicles; oriented fast and rotated brief (ORB-SLAM); flying robots; feature extraction; visual odometry; object recognition.

Cite as: Saira Khan, Abdul Basit. (2023). Vision-Based Monocular SLAM in Micro Aerial Vehicle. *LC International Journal of STEM (ISSN: 2708-7123), 4*(1), 41–51. **https://doi.org/10.5281/zenodo.8200317**

INTRODUCTION

Quadrocopters or micro helicopters are grasping the attention of people for their agility and efficiency. These are used widely in applications such as aerial photography, surveillance and reconnaissance, environmental monitoring, disaster response, and search and rescue operations. Though MAVs are getting popularity because of their advanced features, trilling performance, versatility, and mobility in all search and rescue operations in every environment.

Micro aerial vehicles (MAVs) are small unmanned aerial vehicles (UAVs) that are basically designed to fly in close proximity to objects and navigate in environments that are difficult to access by human beings and by traditional aircraft. MAVs weigh less than 1 kilogram and have a wingspan of less than 50 centimeters. They still struggle for autonomous flights in GPS-denied environments.

MAVs need built-in maps [1] in environments where GPS is denied, unstructured, and unknown. This problem can be solved by simultaneous localization and mapping (SLAM). Robots and drones can operate autonomously in uncharted and chaotic situations without human supervision, thanks to SLAM. Without SLAM, a robot or drone would need to already have a map of its surroundings and be aware of where it is in relation to that map in order to navigate successfully. A pre-existing map cannot be made or is impractical in many real-world situations, such as disaster areas or space exploration.

Oriented FAST and rotated BRIEF (ORB) SLAM is a visual simultaneous localization and mapping (SLAM) system, used in the field of Robotics and computer vision applications. It is a library with open source that provides a real-time 3D mapping of the environment with a single RGB camera. The ORB SLAM algorithm finds and matches locations in successive frames of a video stream using ORB characteristics. It then optimizes a 3D point cloud of the seen image utilizing bundle adjustment algorithms to determine the camera's position and orientation in that environment (see Fig. 1).

Figure 1: Micro aerial vehicle (MAV) Tello with visual sensor and odometry for the localization and mapping.

In the rest of the paper, the related work is discussed in Section II, in Section III the details of the proposed system are presented. The experimental setup and the results are drafted in Section IV. Finally, the conclusion of the manuscript is given in Section V.

LITERATURE REVIEW

In this section, we detail the existing technologies present in the visual monocular SLAM.

Mur et al. [2] made the initial presentation of ORB-SLAM. The authors suggested a monocular SLAM algorithm that employs real-time bundle adjustment techniques for optimization together with ORB (Oriented FAST and Rotated BRIEF) features for feature recognition and description. It has been demonstrated that the algorithm works well in both indoor and outdoor settings.

Wang et al. [3] used a monocular camera to introduce a different visual SLAM approach with dense planar reconstruction. This approach exploits planar template-based trackers for computing camera poses and reconstructing maps. The authors focused on three areas: depending on heterogeneous

information like key points etc, the second area is about deep learning segments and detection of planar regions, and the third area is exploiting planar maps for relocalization.
Mur et al. [4] introduced ORB-SLAM2, an improved version of ORB-SLAM that can handle multiple types of cameras, including monocular, stereo, and RGB-D cameras. This system operates in real time with the capabilities of loop closing and re-localization. It works efficiently in both indoor and outdoor environments.

Campos et al. [5] presented the first system presented, known as ORB-SLAM3 which is characterized as being able to use monocular, stereo, and RGB-D cameras with pin-hole and fisheye lens models to carry out visual, visual-inertial, and multi-map SLAM. One of the key innovations is a feature-based, tightly integrated visual-inertial SLAM system that solely depends on Maximum-a-Posteriori (MAP) estimation during the setup of the IMU (inertial measurement unit). The multiple map system uses a ground-breaking place recognition method with improved memory. ORB-SLAM3 generates a new map which is then combined with earlier maps when the mapped area is revisited. This allows ORB-SLAM3 to live for a prolonged duration of insufficient visual input.

Martine et al. [6] discussed various methods for estimating 6D camera posture meanwhile creating a 3D map of the observed scene. The method was based on visual signals and is known as a novel visual simultaneous localization and mapping (SLAM) system. A binary descriptor ORB was used in this method, often referred to as ORB-SLAM, for all visual tasks such as feature matching, re-localization, and loop closure. Moreover, ORB-SLAM combines graph-based global bundle updating with local updates to enable scaled map construction
without sacrificing real-time speed. The author described how to execute autonomous flight in a low-cost micro aerial vehicle using ORB-SLAM as a visual positioning system to supply
pitch, roll, and yaw commands to a PD controller.

Zhao et al. [7] chose the features that were crucial for posture VSLAM's estimation. In contrast to traditional feature selection works that focus primarily on efficiency, the author significantly improved posture tracking accuracy while adding little overhead. They offered the Max-logDet metric, which connected to the conditioning of least squares pose optimization problem, to do this by directing the feature selection. A cutting-edge visual SLAM system that incorporates Max-logDet feature selection saw accuracy gains with no additional overhead.

Chen et al. [8] suggested that visual simultaneous localization and mapping and multi-rotor unmanned aerial vehicle navigation in an unknown environment have achieved fame in research and training. Yet, because of its complex hardware setup, safety concerns, and battery constraints, rigorous physical testing can cost high and time-consuming. Before field trials, alternative techniques like simulation tools make testing and validating algorithms simpler. The authors provided a simulation solution for the UAV VSLAM and navigation inquiry that was fully configured for the ROS-Gazebo-PX4 simulator. A selection of localization, mapping, and path-planning software was also included in the simulation platform. Many factors, including intricate settings and onboard sensors, can simultaneously affect the navigation framework in simulation.

Engel et al. [9] used featureless monocular SLAM. The model created extensive maps that were persistent. They provided precise pose estimates for key-frames and generated a 3D environment for them in real-time, along with semi-dense depth maps. The two improvements they produced are as follows: the first was a direct tracking methodology that used a sim to detect scale drift, and the second incorporated the impact of noisy depth values into tracking. A CPU powered the resulting real-time direct monocular SLAM system.

Chappellet et al. [10] evaluated the accuracy and dependability of some camera types' localization in VSLAM. The current Open VSLAM framework is suggested for use with this methodology. Due to its adaptability, it may be used with various cameras, including fish-eye and 360-degree monocular or stereoscopic cameras using RGB or RGB-D approaches. Each camera's output varies a little bit from the others. The lowest localization rate and greatest precision were demonstrated by RGB-D vision. The stereo-fisheye camera supports localization rates and precision for both RGB-D and 360-degree vision.

Yang et al. [11] demonstrated a simultaneous localization and mapping (SLAM) system that made use of numerous cameras to provide reliable posture tracking of any micro aerial vehicle (MAV) in a challenging environment. Pose tracking and map improvement were examined in many cameras of VSLAM. In order to guarantee precise optimization, an analysis was performed. The final implementation of a known monocular VSLAM was claimed improved by using two cameras with non-overlapping fields of view (FoVs).

A VSLAM system makes MAVs autonomous navigation in an unknown environment. This idea can be successfully implemented in configuring multi-camera situations with onboard computational capability. In large environments' operations, visual SLAM can be modified to a constant-time robust visual odometry.

Lv et al. [12] designed the ORB-SLAM technique to track the robot's location and provided rich 3D reconstruction while the robot is exploring the area in real-time. The method successfully estimated the camera positions and built a sparse 3D map of the environment based on photographs. Yet, the scanty chart was useless for either navigating or avoiding obstacles. By using octrees and probabilistic occupancy estimates to improve ORB-mapping SLAM, the author was able to produce 3D reconstructions that might be used in the robot industry. Benchmark datasets are used to illustrate this method's effectiveness. Finally, studies show that when the enlarged SLAM system is used with a portable Kinect 2.0, the camera position is accurately monitored and an automap is produced in real-time.

Steder et al. [13] employed flying objects to concentrate on learning a visual map of the ground. The researcher assumed that the vehicles had one or two inexpensive downward-facing cameras working in tandem with an attitude sensor. This created a visual map to be used for navigation. One of the benefits of this approach was being relatively simple to implement, effectively handling noisy camera images, and working with either a monocular camera or a stereo camera system. It made use of visual features and a PROSAC algorithm variation to estimate the correspondences between features. Graph approaches are to solve the SLAM problem, which extracts spatial constraints between camera postures. It also deals with the issue of effectively locating loop closures. The author conducted a number of flying vehicle studies that show our technology can create maps of both huge outdoor and inside settings.

Celik et al. [14] utilized the architectural orthogonality of the indoor environment, The authors developed a novel monocular camera-based indoor navigation and range technique to estimate range and vehicle states from a monocular camera for vision-based SLAM. The navigation approach assumed a previously unidentified interior or indoor-like man-made environment, one that is not GPS-enabled and may be represented by straight architectural lines and energy-based feature points. The research experimentally evaluated the proposed methods using a completely self-contained micro aerial vehicle (MAV) with robust onboard image processing and SLAM capabilities. It built and outfitted a tiny aerial vehicle to fly in limited locations, especially when GPS signals weren't there and there weren't many sensor choices available.

Chen et al. [8] suggested that the surrounding texture of scenes is low or repeated when mobile robots are operating in indoor unknown locations. Due to the tendency of robots to wander back and forth within a small space, it is difficult to assess their poses, and picture details are readily lost when tracking them. To address such tracking challenges, authors proposed a one-circle feature-matching approach, also known as a series of circle matching for the time after space (STCM), and an STCM-based visual-inertial simultaneous localization and mapping (STCM-SLAM) method. This approach improved the indoor pose prediction of the mobile robot by tightly linking the stereo camera and the inertial measurement unit (IMU). Visual characteristics are tracked using optical flow in both directions. The absolute accuracy and relative accuracy of STCM are both 129.167 percent and 37.869 percent higher than that of correlation flow, respectively. In terms of scale error, operating frequency, and CPU load, their investigations show that STCM-SLAM performs significantly better than the OKVIS technique.

Ozbek et al. [15] concentrated on maintaining airplane navigation in situations where GPS is not available. The two best-known algorithms in the literature for visual-inertial navigation systems, VINS-Mono and ORB-SLAM3, were evaluated and their performances were compared.

Dissanayake et al. [16] presented the structure of SLAM. According to the author estimating map convergence to any relative map is developed with zero certainty, where the exact accuracy of the map and location of the vehicle is shown and the Vehicle location reaching to lower bound is explained with vehicle initial uncertainty. According to these results, autonomous vehicles can start moving in an unknown location. By the use of relative observations, it builds a world map and simultaneously computes bounded estimation of the location of the vehicle.

They also discussed the significance of the algorithm implementation on vehicles moving in an environment that is outdoor by using Millimeter Wave (MMW) radar to provide map's relative observations. By implementing this, main issues for example map management and data association are tackled in an outdoor environment. Outputs are then compared with the locations of the map which are gained by surveying.

SYSTEM OVERVIEW

For flying robots, hardware and software setup depend on the specific application and platform. Some general guidelines are here:

- A flight controller is needed for stabilizing the drone, controlling its movements, and reading sensor data.
- ORB SLAM requires a camera to capture images of the environment, which it uses to estimate the drone's position and orientation. A popular drone used in our research is the DJI Tello with an onboard nose-mounted camera.
- For communicating with the ground station or other devices during flight, the drone uses wireless communication modules such as WiFi or Bluetooth to achieve this.
- Flying robots require a reliable and powerful battery and power distribution system to operate efficiently.
- We need an onboard computer with enough processing power to run the algorithm in real time. The system's specifications for the computer are AMD A10 quad-core processor with 4GB in RAM. Ubuntu and ROS (Robot Operating System). See Fig. 2 for the details of the system.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-01 | March-2023

Initialization

For initializing ORB SLAM, the robot should be moved around in the environment to capture images or extract features from different viewpoints. The algorithm will then use these images to estimate the initial position and orientation of the robot, as well as the features in the environment.

Feature Detection and Matching

ORB SLAM works by detecting and matching features in the images captured by the camera(s). This process involves identifying key points and descriptors in images, which are used to track the robot's position and build the environment's map. The algorithm uses the ORB (Oriented FAST and Rotated BRIEF) feature detector and descriptor, designed to be fast and robust.

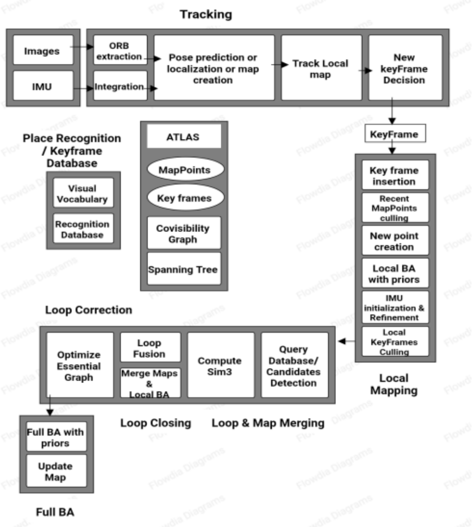


Fig. 2. An overview of the proposed ORB SLAM 3 algorithm for the micro aerial vehicle Tello.

Camera Tracking

Once the algorithm has been initialized, it will track the robot's position and orientation in real time as it moves around the environment. This involves continuously detecting and matching features in the images captured by the camera(s), and updating the position of the robot and the environment's map accordingly.

Map Building

For loop detection and re-localization, our system uses bags of words place recognition module (DBoW2). Visual words are a sampling of descriptor space called Visual vocabulary or codebook which is a key component of the bags-of-words (BoW) approach used for image classification, object recognition, and place recognition. Visual vocabulary is the set of visual words or visual features that are extracted from a large collection of images with ORB descriptors and used to represent images as histograms of visual words.

The process of building a visual vocabulary involves several steps. First, a large collection of images is acquired and pre-processed to extract visual features, such as ORB descriptors. Second, a database will be built with an inverted index to store visual words in the vocabulary and the already-seen frames. It will make database access very quick. The culling procedure will be used to update the database.

Loop Closing

For a flying robot, loop closing in ORB SLAM 3 is essential for simultaneous localization and mapping (SLAM). Recognizing previously visited areas and adjusting the robot's predicted trajectory are required for loop closing. The bag-of-words method, in which images are described as histograms of visual words, is used in ORB SLAM 3 to close loops. The current image's and earlier photos' histograms are compared to determine the order to discover the best match, the histograms of the current image and earlier images are compared. If the match is strong enough, the algorithm takes the likelihood of a loop closure into consideration.

Due to the flying robot's high-speed movement and the environment's quick changes, loop closing can be difficult. The ORB SLAM 3 algorithm attempts to correct the motion blur brought on by the camera's movement by taking into account the robot's motion model. The Loop Closure consists of two steps: it first starts the Loop Detection step and then it does the Loop Closure. Despite of sensors' accurate information and reliability of the odometry and mapping algorithm, noise and approximation introduce errors into the pose and map estimation process.The Loop Closure corrects the errors. The main large classes of ORB SLAM are Bundle Adjustment (BA) methods and filter methods.

Map Optimization

In order to reduce the reprojection inaccuracy of the observed features, ORB SLAM uses bundle adjustment techniques to optimize the camera positions and the 3D map.

Localization

When tracking fails, a reliable SLAM system has to localize the camera again. By using a proper motion model, ORB-SLAM3 resolves the relocalization problem. The algorithm must be accurately modeled because the motion of a flying robot is more complex than that of ground-based robots since it occurs in three dimensions. The sensor setup is another crucial element. Several sensors, including cameras, lidars, and IMUs, are frequently found on flying robots and can be used for localization. The ORB SLAM algorithm's accuracy and robustness can be increased by integrating these sensors and fusing the data they produce. In addition, the flying robot's size and speed need to be considered. The algorithm must be tuned to manage the robot's rapid movements while preserving precision and consistency. Using the well-known SLAM dataset EuRoC, the localization accuracy of visual SLAM and visual-inertial SLAM approaches has been evaluated. Even though it was gathered using a micro-aerial vehicle, we came to the conclusion that this dataset is the best option for comparing STCM-SLAM, ORB-SLAM2, and OKVIS. The stereo camera photos, IMU data, and actual robot motions are all included in the package.

EXPERIMENTS AND RESULTS

We used a Tello drone for our research where we did experiments by letting our drone fly autonomously in an indoor environment. We also set coordinates (x, y, z) for the purpose of teaching and repeating the approach. We use the IMU sensor in conjunction with ORB-SLAM 3 algorithms to estimate the

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-01 | March-2023

pose (position and orientation) of the robot in real time. It gives us information on the linear acceleration and angular velocity of the robot, which helps in predicting the robot's motion and updates the ORB-SLAM 3 algorithm.

Fig.4 indicates that we first selected an indoor environment, then took tello with a front-facing camera to move in that particular environment and extract features of that environment. The green squares that can be seen on the map are basically the features of that environment. That is used for feature-based extraction.

We have chosen different indoor environments to test our algorithm. Our flying robot works wonders in every indoor environment and extracts features of that environment in a great manner. Indoor environments consist of classrooms, libraries, and departments and we have also used the EuroC dataset.

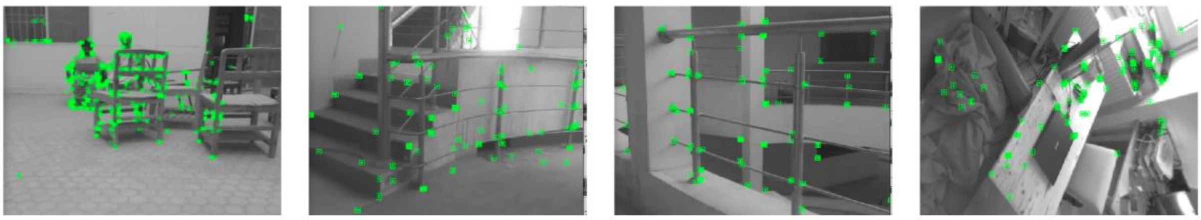
Experiment I: Feature Extraction

In experiment I, We have used Tello with the capability of the front-facing camera which is to create a 3D map of the environment in real-time through feature extraction, Keyframes are allocated via the server, and a local map is created. Tello is used for extracting features. When Tello takes off or starts its flight, it moves in the environment and when moves forward, objects that it finds, it saves as key frames and it continues doing this till it reaches its destination. Green squares which are visible in our map are basically the keyframes.

In the first map, the drone moved in an indoor scenario with coordinates alright set, and it moved in a specified path and extracted features (see Fig. 3-(a)).

Fig. 3-(b) shows a straight autonomous flight where the MAV moved in a straight trajectory, where it went forward and extracted features of that area, and then the controller stored them which identified the pose of the MAV and features of the environment in real-time. The flight time was 45 seconds. Fig. 3-(c) shows that MAV moved in a corridor or in a small indoor environment and extracted its features and keyframe matching. It also created a map of the environment.

In Fig. 3-(d) MAV moved into a small and feature-rich area, it moved in that area and extracted its features, and stored them in the map. Additionally, We also used the standard dataset EuroC, TUM, and Bags of Words to examine the feature extraction of the proposed method.



(a) (b) (c) (d)

Fig. 3. An overview of the proposed ORB SLAM 3 algorithm for the micro

Experiment II: Flight with Various Trajectories

We have generated simple and challenging trajectories using the micro flying robot Tello. For the custom map, we used the department's indoor environment and generated simple and complex trajectories to test the robustness of the proposed ORB SLAM 3 method. In the first

Published by Logical Creations Education Research Institute. www.lceri.net

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-01 | March-2023

path, the robot moved in a straight path generating a feature map and odometry readings, see Fig. 4-(a).

In Fig. 4-(a) the micro-ariel vehicle moved on the straight path. Making trajectory by extracting features of a straight path. Whereas in Fig. 4-(b) it is a loop closure scenario, Tello moved in an area and collected keyframes when it revisited the same area, it found the same keyframes. Fig. 4-(c), MAV moved in a corridor where the flying robot had random or autonomous trajectories. These key frames are stored in its database. In Fig. 4-(d) a dataset EuroC Visual-Inertial and Bags of Words were used to generate the map and drone's trajectory in its autonomous flight in real times in a traversed trajectory. When the EuroC dataset is applied to our proposed algorithm, it will detect all the keyframes correctly generate a map, and read all points successfully.

Here, Multiple colors are visible in our 3D map with different representations. The green color in our map shows the autonomous flight of the Tello device, which means Tello's flight is represented by green color. The blue color is to show the target's or our flying robot's trajectory. The red color shows waypoints and the Black color is used to show MAV's start position.

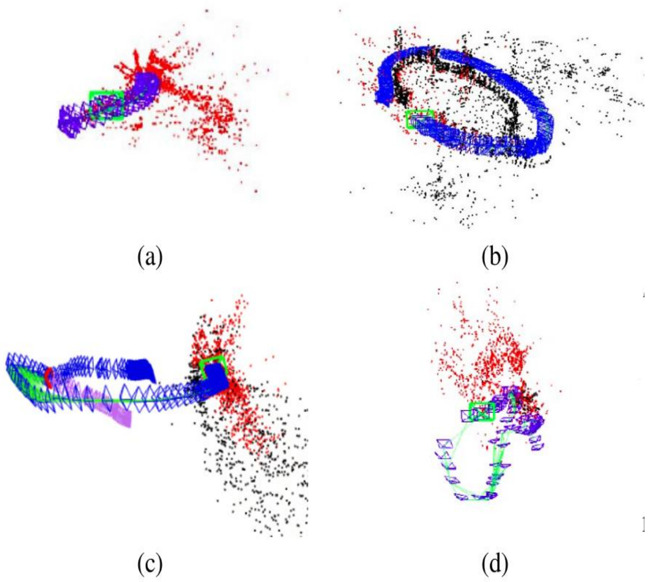


Fig. 3. An overview of the proposed ORB SLAM 3 algorithm for the micro

sExperiment III: Various Dataset Used for Experiments

We have used the EUROC dataset, Bundle Adjustment (BA), and Bags of words(DBoW) in ORB-SLAM3 algorithm with IMU to estimate camera pose and localization, loop closure, and tracking and mapping. A ground station was assigned to PC. We have shown the results of the creation of a 3D map and recording of the waypoints from this traversed trajectory at the mapping stage.

Table 1: Accuracy test: Experiment for testing the accuracy of the proposed method

Chapters	# of Vehicles	True Positive (TP)	False Positive (FP)
Video set I	40	38	4

Video set II	35	32	5
Video set III	30	25	3

CONCLUSION

We have used ORB SLAM 3 for our DJI Tello drone with the onboard frontal camera used as a sensor to give us the images of any unknown indoor or outdoor environment. Our quadrocopter was allowed to navigate any dynamic environment either by giving it instruction through setting its coordinates' values where it could move around those coordinates and it could fly high by controlling it manually. Or we left it for autonomous flight in any setting. ORB SLAM 3 gave its flight a robust initialization which resulted in accurate localization and mapping. Our proposed method has the capability to estimate the 6D positions of a monocular camera.

We also needed Bluetooth or wifi to connect our UAV with the controller. For getting imagery from MAV and communicating with our robot, the Robot Operating System (ROS) has been used with its complete packages. For map building and tracking we used Bags of Words (BoW), which created visual vocabulary consisting of visual features of surroundings. We used the Loop closure technique to enable our system to recognize already visited areas and adjust them with the trajectory.

ACKNOWLEDGMENT

We are thankful to the government of Pakistan for supporting and funding Saira Khan's research work.

REFERENCES

[1] Weiss, S., Scaramuzza, D., & Siegwart, R. (2011) Monocular-SLAM–based navigation for autonomous micro helicopters in GPS-denied environments, Journal of Field Robotics, vol. 28, no. 6, pp. 854–874.
[2] Mur-Artal, R., Montiel, J. M. M., & Tardos, J. D. (2015) Orb-slam: a versatile and accurate monocular slam system, IEEE transactions on robotics, vol. 31, no. 5, pp. 1147–1163.
[3] Wang, X., Christie, M., & Marchand, E. (2021) Tt-slam: Dense monocular slam for planar environments, in 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 11 690–11 696.
[4] Mur-Artal, R., & Tardós, J. D. (2017) Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras, IEEE transactions on robotics, vol. 33, no. 5, pp. 1255–1262.
[5] Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M. & Tardós, J. D. (2021) Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam, IEEE Transactions on Robotics, vol. 37, no. 6, pp. 1874–1890, 2021.
[6] Martínez-Carranza, J., Loewen, N., Márquez, F., García, E. O., & Mayol-Cuevas, W. Towards autonomous flight of micro aerial vehicles using orb-slam, in 2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS). IEEE, pp. 241–248.
[7] Zhao, Y., & Vela, P. A. (2020) Good feature matching: toward accurate, robust vo/vslam with low latency, IEEE Transactions on Robotics, vol. 36, o. 3, pp. 657–675.
[8] Chen, S., Chen, H., Zhou, W., Wen, C.-Y., & Li, B. (2020) End-to-end uav simulation for visual slam and navigation, arXiv preprint arXiv:2012. 00298.
[9] Engel, J., Schöps, T., & Cremers, D. (2014) Lsd-slam: Large-scale direct monocular slam, in Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part II 13. Springer, pp. 834–849.

[10] Chappellet, K., Caron, G., Kanehiro, F., Sakurada, K., & Kheddar, A. (2021) Benchmarking cameras for open vslam indoors, 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021, pp. 4857–4864.
[11] Yang, S., Scherer, S. A., Yi, X., & Zell, A. (2017) Multi-camera visual slam for autonomous navigation of micro aerial vehicles, Robotics and Autonomous Systems, vol. 93, pp. 116–134.
[12] Lv, Q., Lin, H., Wang, G., Wei, H., & Wang, Y. (2017) Orb-slam-based tracing and 3d reconstruction for robot using kinect 2.0, 29th Chinese control and decision conference (CCDC). IEEE, 2017, pp. 3319–3324.
[13] Steder, B., Grisetti, G., Stachniss, C., & Burgard. W. (2008). Visual slam for flying vehicles, IEEE Transactions on Robotics, vol. 24, no. 5, pp. 1088–1093.
[14] K. Celik and A. K. Somani. (2013) Monocular vision slam for indoor aerial vehicles. Journal of electrical and computer engineering, vol. 2013, pp. 4–4, 2013.
[15] Özbek, B. K., & Turan, M. Research on the availability of vins-mono and orb-slam3 algorithms for aviation, WSEAS Transactions on Computers, ISSN/E-ISSN, pp. 1109–2750.
[16] Dissanayake, M. G., Newman, P., Clark, S., Durrant-Whyte, H. F., & Csorba, M. (2001) A solution to the simultaneous localization and map building (slam) problem, IEEE Transactions on robotics and automation, vol. 17, no. 3, pp. 229–241.