Drone Image Processing for Efficient Obstacle Avoidance in Transmission Line Inspections

Yul Yunazwin Nazaruddin[1], A.S. Reinard Rahardian[1], Steven Bandong[1], Muhammad Z. Romdlony[2], Tua A. Tamba[3], Vebi Nadhira[1] and Narendra Kurnia Putra[1]

[1] Instrumentation, Control and Automation Research Group, Department of Engineering Physics, Institut Teknologi Bandung, Bandung, Indonesia
[2] School of Electrical Engineering Telkom University Bandung, Indonesia
[3] Dept. Electrical Engineering Parahyangan Catholic University Bandung, Indonesia
yul@itb.ac.id

Abstract. This paper presents a system designed to estimate object positions and detect the presence of objects using image data captured by a monocular camera mounted on a drone. The testing of the method was conducted using a controller algorithm, where a drone was flown laterally in front of two miniature plants, capturing images for object detection and position estimation via triangulation. The system was able to reduce triangulation errors after multiple detections of the same object, achieving improved accuracy. Further real-world testing demonstrated the system's ability to avoid obstacles and successfully navigate toward predefined target positions. The results indicate that the presented image processing algorithm can serve as a reliable input for target search and pathfinding applications, with potential future applications in monitoring tasks, such as transmission line inspections.

Keywords: Object detection, Position estimation, Triangulation, Drone navigation, Image processing, Transmission line.

1 Introduction

Geographical challenges and limited infrastructure make electricity provision in remote areas more complex, requiring effective technology to ensure power reaches every region. Technologies such as drones[1–3] and smart sensors [4–6] are now used to monitor transmission networks in real-time and detect faults. With the help of these technologies, disruptions to the transmission network can be minimized, ensuring a stable electricity supply to remote areas.

The use of smart sensors and drones offers distinct advantages for transmission line monitoring. Smart sensors provide continuous, real-time data on line conditions, ensuring accurate detection of faults at specific locations. However, their stationary nature limits coverage to the areas where they are installed. In contrast, drones offer greater flexibility, capable of inspecting large sections of transmission lines, including remote or hard-to-reach areas, by providing real-time visual data.

In [1], the drone is used for fast inspection of long power transmission lines and acts as a central communication hub or relay point, gathering data from sensors placed along the lines. In addition to enhancing communication, drones are also used to directly monitor the condition of transmission lines [3] or employ deep learning to assess potential failures caused by harsh environmental conditions [7]. Others are also working to expand drone applications to further improve the reliability of power transmission lines [8] [9].

However, during flight, the drone must have the capability to effectively avoid obstacles while remaining focused on its primary target, such as the need for monitoring power transmission lines. Obstacles the drone may encounter could include structures like buildings or trees. Therefore, the drone's navigation system must be supported by technologies such as sensors that can detect and determine the position of nearby obstacles in real time. In this paper, we use the camera already integrated into the drone and apply deep learning methods to detect the presence of obstacles and triangulate the position of both obstacles and targets. The obstacle's position is then fed into the drone controller, enabling the drone to avoid obstacles and pursue the target.

Since the operation needs to be real-time, the image-based detection algorithm YOLOX [10]was chosen for its superior computation time and good accuracy. For tracking, Unitrack [11]is employed. Additionally, the Iteratively Reweighted Midpoint Method (IRMP) [12]is used for triangulation, which iteratively utilizes object appearance information to infer the position of obstacles. The integration of these methods with the control system presented in[13] demonstrates good capabilities in avoiding obstacles while maintaining a fast response toward reaching the target.

2 Methods

Obstacle avoidance refers to the effort of avoiding obstacles in the path of the drone's movement, and it generally relies on input from an obstacle detection system to identify the presence and position of obstacles. The most standard avoidance methods utilize rangefinders, radar, or depth cameras, providing information about the distance between the drone and the obstacles. Saitoh [14]developed an obstacle avoidance technique based on a monocular camera, as this type of camera is relatively inexpensive and widely available in the market.

Considering[14], this paper presents a system designed to estimate object positions and detect the presence of objects using image data collected from a monocular camera mounted on a drone. The system is intended for monitoring the condition of power grid transmission lines. Additionally, it is capable of detecting and avoid obstacles during the drone's flight. However, this paper does not cover the development of the drone controller or the specifics of transmission line monitoring, as these will be discussed in a separate paper[13]. Instead, the focus is on image processing for obstacle avoidance while multiple drones attempt to reach the transmission line for monitoring.

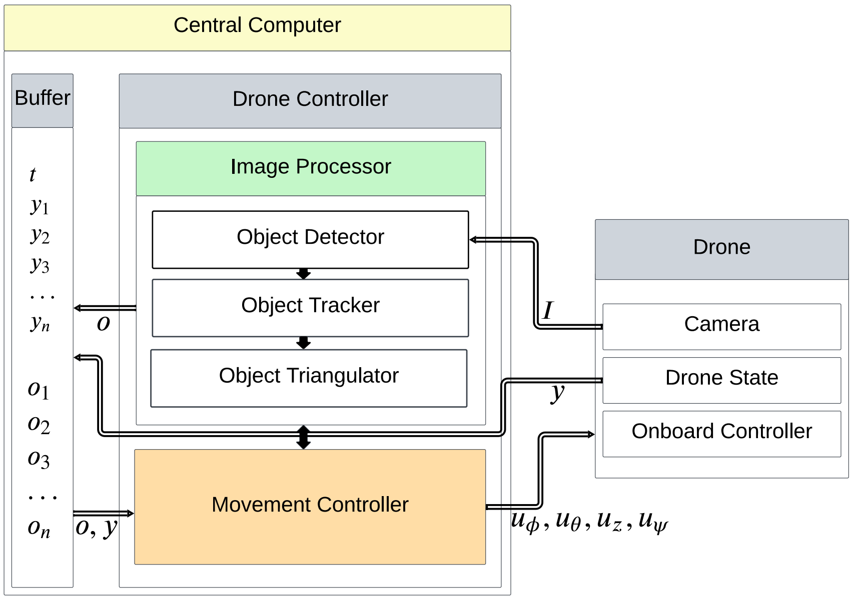

Fig. 1. Overview of the Designed System

The overall system is depicted in Fig 1. The system will run a control system on the central computer for each drone, as indicated by the Drone Controller block. Each control system includes an Image Processor block that reads the drone state information y along with the camera images I, and outputs the position of obstacles o and targets t to the Buffer block, which records the positions of objects and swarm member drones. The position records in the Buffer are then read by the Movement Controller block, which processes this information into control inputs $(u_\phi, u_\theta, u_z, u_\psi)$. For those interested, the controller methods used in this paper are Parallel Navigation and PID Control, as published in [13] . The system shown in Fig 1 will be implemented using the Robot Operating System (ROS) framework, which will be run on the central computer and connected to the drones via a Wireless Fidelity (Wi-Fi) network.

Fig 1 shows that the Image Processor block consists of three components: object detection, object tracking, and object triangulation. Object detection and tracking are used to detect objects along with their centroid direction from the drone in the form of a bearing vector (b), and to identify if the same object appears in different camera frames. Object triangulation is used to estimate the 3D position of the object based on the readings from object detection and tracking. The flowchart for obstacle detection is illustrated in Fig 2. b_i is the bearing vector, b_i' is the set of bearing vectors identified by the object tracker as originating from the same object, and $\hat{p}_i$ is the triangulated position of the object.

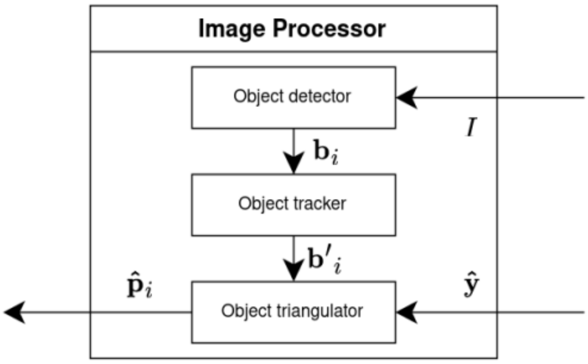

Fig. 2. Flowchart of the Image Processing System

2.1 Object Detection and Tracking System

YOLOX [10] is used as framework for object detection which a development of YOLO. It has advantage for real-time application as it has good detection performance while has fast computation speed. The detection-tracking process is carried out using the Unitrack [11]framework, which integrates various tracking processes to utilize the same image detection model. It includes three key components: the Appearance Model, which extracts feature maps from source images using the YOLOX object detector; Propagation and Association, where detected objects are linked across frames; and Multiple Object Tracking (MOT), which tracks multiple objects and identifies their recurrence using consistent bounding boxes across frames.

Before starting the detection stage, to ensure the bearing vector obtained from this stage is accurate, the camera images must first be corrected to produce images with minimal distortion. This stage is performed using Python's OpenCV library with 37 test images of a chessboard pattern (Fig 3). The calibration results include the transformation matrix $\boldsymbol{K}$, the tangential distortion coefficient $\boldsymbol{k}$, and the radial distortion coefficient $\boldsymbol{p}$ from the camera, along with the calibration error value as indicated in Equation 3.22. These parameters will be used to perform image undistortion using OpenCV.

$$\begin{gathered} \boldsymbol{K} = \begin{bmatrix} 504.87 & 0 & 329.40 \\ 0 & 500.67 & 158.59 \\ 0 & 0 & 1 \end{bmatrix} \\ \boldsymbol{k} = [-0.82824 \quad 0.71702 \quad 0.02304] \\ \boldsymbol{p} = [0.01456 \quad -0.01822] \\ RMSE = 0.2355\, pixel \end{gathered} \tag{1}$$

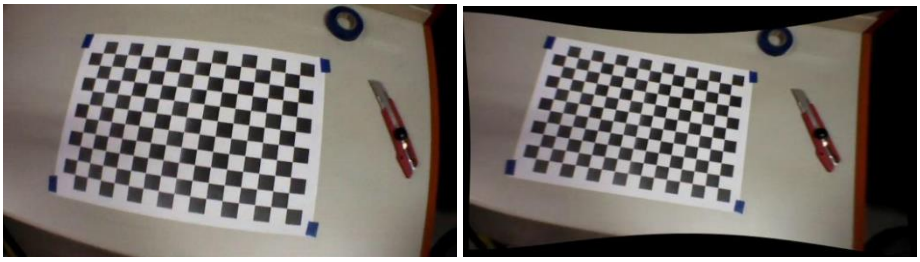

Fig. 3. Before and After Image of the Undistortion Operation (Left: Before, Right: After)

The image, after undergoing undistortion operations, is then fed into the object detection system. The output from this operation consists of a set of bounding boxes around the detected objects in the image. To minimize noise, the object detection system is configured to detect only a single object class, specifically chairs, which are used as test obstacles. For object detection, the YOLOX-L model size is used. YOLOX was chosen due to its proven performance when integrated with the Unitrack framework, and the YOLOX-L version was selected because it provides a sufficiently high inference frequency without sacrificing accuracy, as demonstrated in the benchmark results shown in Table 1. In the table, Inference Time refers to the time required for object detection (inference) on a single image, Inference Rate indicates the number of inferences that can be performed per second, and Miss Rate represents the percentage of test objects—specifically plastic plants—that were not detected when they appeared in the camera. The detection results are then fed into the object tracker system to associate detections with the same object across different images.

Table 1. Performance Benchmarking of YOLOX Variants on RTX 3050 GPU

Network Size	Inference Time (RTX 3050)	Inference Rate (RTX 3050)	Miss Rate
YOLOX-S	0.052 s	19.0 fps	50%
YOLOX-M	0.068 s	14.6 fps	38%
YOLOX-L	0.097 s	10.3 fps	5%
YOLOX-X	0.161 s	6.2 fps	5%

2.2 Object Triangulation

The Iteratively Reweighted Midpoint Method (IRMP)[12] is a triangulation technique used to estimate the position of an object by iteratively refining the midpoint of the detected object based on data from different viewpoints or sensor readings. In this method, initial estimates are made based on the detected object's position from multiple perspectives, and these estimates are iteratively adjusted to improve accuracy. The weighting of the estimates is updated at each iteration, focusing more on reliable data points, which helps minimize error accumulation over time. This method uses the appearance of the object in multiple frames or sensor readings to infer its 3D position.

For each detected object, the bearing vector (b_i) is calculated relative to the drone's position by performing 3D reconstruction of the centroid point for each bounding box. This reconstruction operation is carried out using the following equation

$$\begin{aligned} b_i &= R_i^T K^{-1} [u \quad v \quad 1]^T \\ T_i &= [x \quad y \quad z]^t \end{aligned} \tag{2}$$

Here, K^{-1} is the inverse of the camera matrix obtained from calibration, R_i^T is the inverse of the camera transformation matrix, equivalent to the rotation matrix representing the drone's orientation, and T_i is the reference point for the bearing vector, which corresponds to the drone's position. Each pair of R_i^T and T_i values constitute one data set from a single object detection. To minimize errors, triangulation is performed every time new data is available, but only after five or more data points are collected. The triangulation results are then stored in the object list, which is used for obstacle avoidance.

3 Results and Discussion

The testing of the presented method was conducted using the controller algorithm outlined in [13]. Two plastic plants were placed side by side, approximately 1 meter apart. A drone then took off and flew laterally at a speed of 0.5 m/s in front of the two plants. During the flight, the drone captured images, which were transmitted to the object detection and tracking system. The results were then fed into the triangulation system on the central computer, producing images similar to those shown in Fig 4. Once the same object is detected three times, triangulation is performed to determine the object's position. Triangulation is repeated each time the object is detected again to produce a more accurate position estimate. The results of the triangulation are shown in Fig 5 to 6

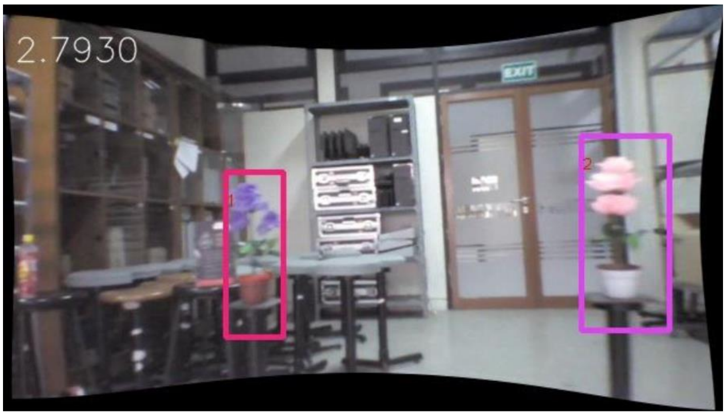

Fig. 4. Sample Image from Object Detection Test Results

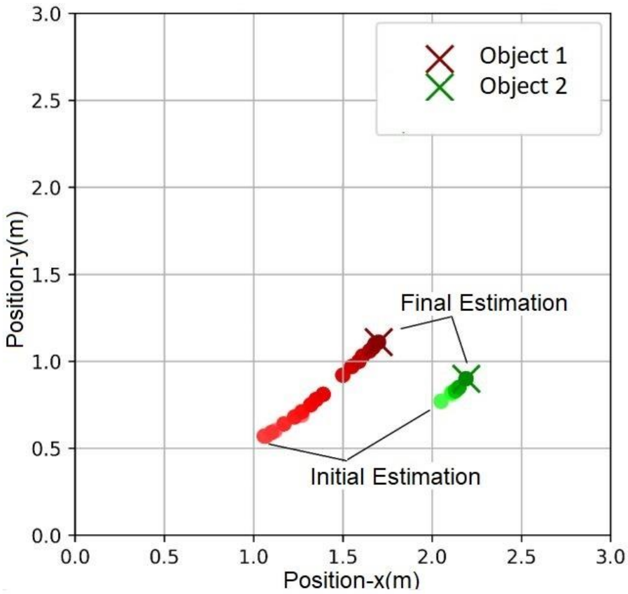

Fig. 5. Position estimation results from the image processing system.

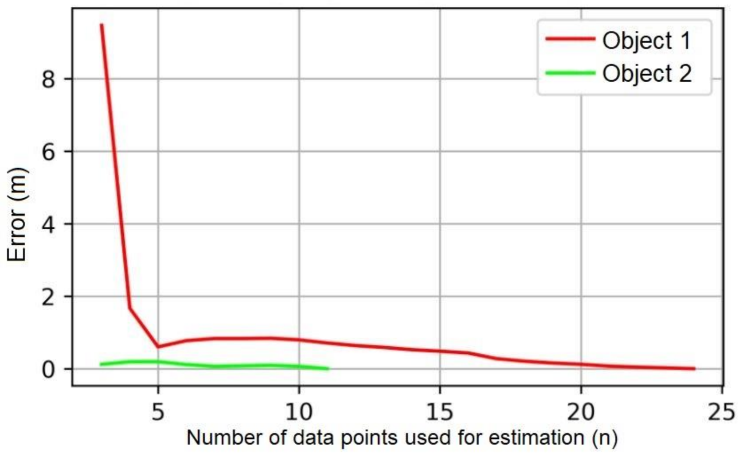

Fig. 6. Triangulation error of the object based on the last detected position.

In Fig 6, assuming the final triangulation is the closest to the actual position of the object, it is observed that the initial triangulation error for Object 1 reached 9 meters, only approaching the true position on the third triangulation, when there were 5 bearing vector data points. However, for Object 2, the initial triangulation error was less than 20 cm. Based on this data, in subsequent tests, triangulation will only be performed after there are more than 5 bearing vector data points for an object. Fig. 6 demonstrates that 5 bearing vectors yield a small error, while increasing the required initial number of bearing vectors could limit the availability of early-stage position information due to the need to wait for sufficient vector data. To further test the method in a real environment, a drone will fly to avoid stationary obstacles and reach a predefined target position. The control being used follow our research paper in [13]. The resulting path is shown in Fig 7 and the time needed to reach the target is presented on Table 2. The real experiment is depicted on Fig 8. From these tests, it was found that the system was able to avoid obstacles, indicating that the image processing algorithm can now be used as an input for target search. For instance, this method could be applied in the future for transmission line path searches for monitoring purposes.

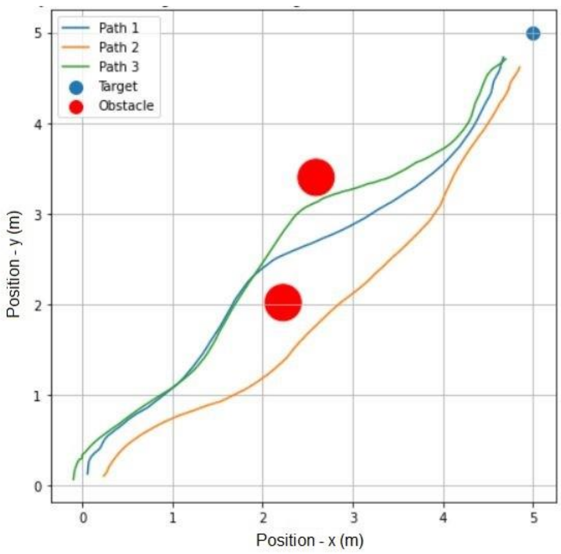

Fig. 7. Path map for predefined target search with real obstacles

Fig. 8. Example of predefined target search results with real obstacles.

Table 2. Target pursuit time with real obstacles

Path	1	2	3
T(s)	11.30	10.25	11.60

4 Conclusions

The method presented in this paper successfully demonstrated its capability to estimate object positions and detect the presence of objects at a laboratory scale, similar to transmission line inspection, using a drone-mounted monocular camera. The triangulation process improved position accuracy with multiple detections, as evidenced by the decreasing error for Object 1 and the initially lower error for Object 2. Performing triangulation only after accumulating more than five bearing vector data points effectively reduced error rates. In real-world testing, the system successfully avoided obstacles and reached its target. In the context of transmission line inspection, the obstacles used in the experiment represent plants and buildings, while the target represents transmission line components. In the future, the proposed method will be integrated into a multi-drone system to enhance transmission line monitoring, enabling greater coverage and faster data collection across large areas.

References

1. Weng, L., Zhang, Y., Yang, Y., Fang, M., Yu, Z.: A mobility compensation method for drones in SG-eIoT. Digital Communications and Networks. 7, (2021). https://doi.org/10.1016/j.dcan.2020.07.011.
2. Luo, Y., Chen, Y.: Energy-aware dynamic 3d placement of multi-drone sensing fleet. Sensors. 21, (2021). https://doi.org/10.3390/s21082622.
3. Menéndez, O., Pérez, M., Cheein, F.A.: Visual-based positioning of aerial maintenance platforms on overhead transmission lines. Applied Sciences (Switzerland). 9, (2019). https://doi.org/10.3390/app9010165.
4. Raja, P.D.A., Thapa, K.N.K., Harsha, K.S.S., Swami Naidu, J.H.M., Krishna, K.S., Sivakumar, A.: Implementation of Transmission Line Fault Detection System using Long Range Wireless Sensor Networks. International Journal on Recent and Innovation Trends in Computing and Communication. 11, (2023). https://doi.org/10.17762/ijritcc.v11i5.6580.
5. Li, L., Zhao, H.: Power Line Monitoring Data Transmission Using Wireless Sensor Network. Journal of Power and Energy Engineering. 03, (2015). https://doi.org/10.4236/jpee.2015.38008.
6. Lin, J., Zhu, B., Zeng, P., Liang, W., Yu, H., Xiao, Y.: Monitoring power transmission lines using a wireless sensor network. Wirel Commun Mob Comput. 15, (2015). https://doi.org/10.1002/wcm.2458.

7. Siddiqui, Z.A., Park, U.: A Drone Based Transmission Line Components Inspection System with Deep Learning Technique. Energies (Basel). 13, (2020). https://doi.org/10.3390/en13133348.
8. Ambatkar, H.P., Dhatrak, R.K.: Review: Drone applications in Transmission Line. In: 2022 International Mobile and Embedded Technology Conference, MECON 2022 (2022). https://doi.org/10.1109/MECON53876.2022.9752257.
9. Morita, M., Kinjo, H., Sato, S.: Autonomous flight drone for infrastructure (transmission line) inspection (3). In: ICIIBMS 2017 - 2nd International Conference on Intelligent Informatics and Biomedical Sciences (2017). https://doi.org/10.1109/ICIIBMS.2017.8279692.
10. He, Q., Xu, A., Ye, Z., Zhou, W., Cai, T.: Object Detection Based on Lightweight YOLOX for Autonomous Driving. Sensors. 23, (2023). https://doi.org/10.3390/s23177596.
11. Wang, Z., Zhao, H., Li, Y.L., Wang, S., Torr, P.H.S., Bertinetto, L.: Do Different Tracking Tasks Require Different Appearance Models? In: Advances in Neural Information Processing Systems (2021).
12. Yang, K., Fang, W., Zhao, Y., Deng, N.: Iteratively Reweighted Midpoint Method for Fast Multiple View Triangulation. IEEE Robot Autom Lett. 4, (2019). https://doi.org/10.1109/LRA.2019.2893022.
13. Reinard Rahardian, A.S., Nazaruddin, Y.Y., Nadhira, V., Bandong, S.: Implementation of Parallel Navigation and PID Controller for Drone Swarm Pursuit. In: IFAC-PapersOnLine (2023). https://doi.org/10.1016/j.ifacol.2023.10.1300.
14. Saitoh, T., Tada, N., Konishi, R.: Indoor mobile robot navigation by central following based on monocular vision. IEEJ Transactions on Electronics, Information and Systems. 129, (2009). https://doi.org/10.1541/ieejeiss.129.1576.