Real-Time 3D UAV Pose Estimation by Visualization †

Benjamaporn Nenchoo and Suradet Tantrairatn *

Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; benz_2469@hotmail.com

* Correspondence: suradetj@sut.ac.th; Tel.: +66-955-623-555

† Presented at the Innovation Aviation & Aerospace Industry—International Conference 2020 (IAAI 2020), Chumphon, Thailand, 13–17 January 2020.

Abstract: This paper presents an estimation of 3D UAV position in real-time condition by using Intel RealSense Depth camera D435i with visual object detection technique as a local positioning system for indoor environment. Nowadays, global positioning system or GPS is able to specify UAV position for outdoor environment. However, for indoor environment GPS hasn't a capability to determine UAV position. Therefore, Depth stereo camera D435i is proposed to observe on ground to specify UAV position for indoor environment instead of GPS. Using deep learning for object detection to identify target object with depth camera to specifies 2D position of target object. In addition, depth position is estimated by stereo camera and target size. For experiment, Parrot Bebop2 as a target object is detected by using YOLOv3 as a real-time object detection system. However, trained Fully Convolutional Neural Networks (FCNNs) model is considerably significant for object detection, thus the model has been trained for bebop2 only. To conclude, this proposed system is able to specifies 3D position of bebop2 for indoor environment. For future work, this research will be developed and apply for visualized navigation control of drone swarm.

Keywords: deep learning; indoor positioning system; visualization

1. Introduction

The application to deploy drone in indoor environments has been needed, for instance using UAV to silos inspection, disaster relief, warehouse management etc., Nevertheless, positioning for indoor environment is the main challenge for researcher. Because GPS hasn't a capability to determine UAV position as discussed by Mainetti et al. (2014) [2] and Mautz et al. (2009) [4]. Thus, sensors with high precision, accuracy performance (centimeter level) and low latency are necessary for UAV indoor applications.

Motion tracking systems use reflected markers to detect position, speed and orientation of object. UAVs are able to flight at the same time by using markers with different shapes. Each camera has coverage area depending on the field of view. A ground station receives and processes camera data, allowing motion-tracking reconstruction. For conclude, Motion tracking systems work as an artificial GPS for indoor environment. Therefore, with high performance and accuracy they have to exchange with high cost for maintenance and calibration [7].

Stereo or IR depth perception cameras technologies in robotics are able to reconstruct 3D model and understand the environment. The high accuracy position and orientation accuracy make these sensors suitable to perform visual odometry algorithms [7]. Intel RealSense Depth camera D435i combines the robust depth sensing capabilities of the D435 with the addition of an inertial measurement unit (IMU). The maximum range Approx. 10 m. Accuracy depending on calibration, scene, and lighting condition. Depth field of view approx. 87° ± 3° × 58° ± 1° × 95° ± 3° with Intel

RealSense Vision Processor D435i as a vision processor board. Due to lower cost than motion tracking systems stereo depth camera used in this research as a sensor.

Therefore, this paper present 3D position estimation by using object detection with Intel RealSense Depth camera D435i as sensor for indoor environment application. For conclusion, this proposed system is able to specifies 3D position of bebop2 for indoor environment.

The paper is organized as follows: Section 2 Materials and Method describe about model training, detection, connection and working flow chart. Section 3 Results and Discussion describes testing result in static and flying test. Section 4 Conclusion describes conclusions of the work and future work.

2. Materials and Methods

This section provides a description of the Model Training and Detection and Positioning Method used in this research.

2.1. Model Training and Detection

Model training is one of procedure for deep learning. Accuracy of trained model depends on dataset, epochs, learning rates and other parameters. In this research, model is trained to detect bebop2 while flying, take-off and landing with 8000 images of dataset that based on YOLOv3(coco) model with Supervisely platform. Dataset's preparation is important procedure before training. After images was labeled, these have to do data transformation language (DTL) process viz vertical flip and horizontal flip. This process is defined with a JSON based config file. YOLO architecture is provided FCNN (fully convolutional neural network). The images (nxn) are passed through the FCNN, then output is (mxm) prediction. The convolutional neural network has 24 convolutional layers with 2 fully connected layers. Alternating 1 × 1 convolutional layer reduce the features space from preceding layers. The convolutional layers pretrained on the ImageNet classification task at half the resolution (224 × 224 input image) and then double the resolution for detection [1] as shown in Figure 1.

However, the connection of YOLOv3 with Intel RealSense Depth camera D435i aren't able to directly connect each other. By the way, this problem solved by using YOLOv3 base on Robot Operating Systems (ROS) or Darknet ROS package work as ROS bridge.
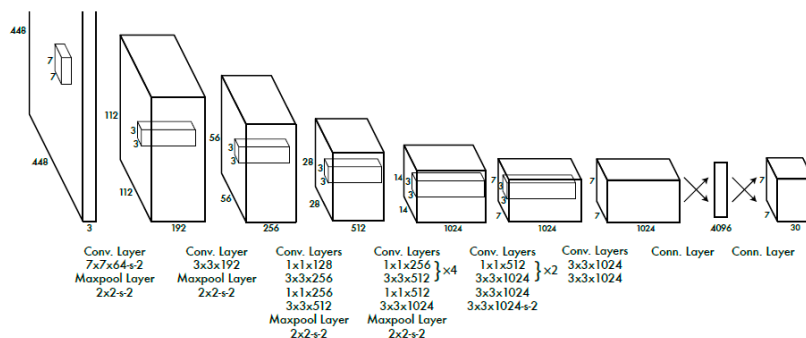

Figure 1. YOLO's architecture [1].

2.2. Positioning Method

3D positioning systems for indoor environment using Intel RealSense Depth camera D435i is utilized as a positioning sensor which set origin coordinates X = 0 at middle of camera's frame, Y = 0 at ground level and Z = 0 at position of depth camera and detect Bebop2 with YOLOv3 object detection. The bounding box pixels of target object on screen size 640 × 480 pixels send to points cloud topics of Intel RealSense Depth camera D435i that collect the coordinate x, y and z in every single pixel. For receives the coordinate x, y and z in meter unit, coordinates pixel center x and y from darknet ros [3] send to points cloud topic with point cloud library (PCL) for transformation of pixel to meter. Moreover, there is some noise from the position x, y and z in meter after transformation

that could filter. This research, using average method with 50 window size of data as a Noise filtration method. In additional the schematic diagram of working flow shown in Figure 2.
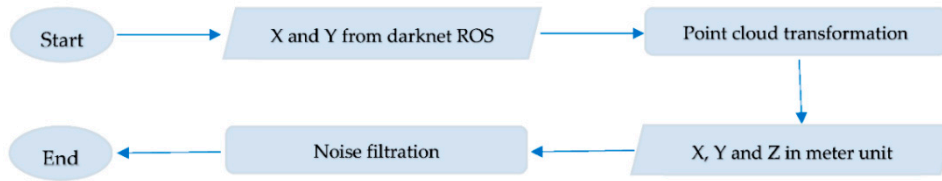

Figure 2. The schematic diagram of working flow.

3. Results and Discussion

In test implementation, Bebop2 was employed as a target object the camera by setting above the ground 45 cm. Testing area for indoor environment with area size width (X) 2.5 m., depth (Z) 2.5 m. and height (Y) 1.5 m. and minimum depth 0.5 m. The camera calibrated before testing in 6 positions randomly. The coordinate X and Z is the camera position and Y is ground. Setting up for testing is illustrated in Figure 3.
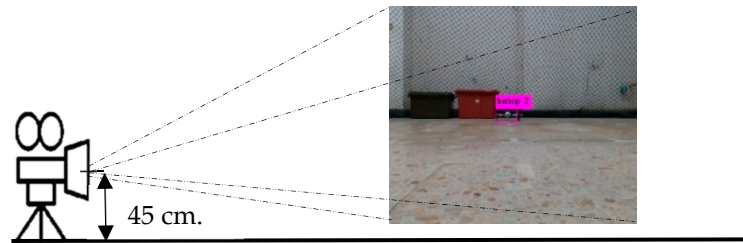

Figure 3. Setting up of camera for testing.

3.1. Static Testing

The testing performed to estimate the indoor position of UAVs. The estimated x, y and z position of the UAV at several positions are collected and the example results is shown in the Table 1.

Table 1. The indoor positioning system result. x, y and z between the measure and actual position of the UAV.

X (m.)	Y(m.)	Z(m.)	Mean X(m.)	Mean Y(m.)	Mean Z (m.)
1	1.5	2.5	0.99	1.5	2.55
0.5	1	2	0.58	1.09	1.97
−0.5	0.5	1.5	0.46	0.48	1.5
0	0	1	0.04	−0.06	0.97

The system able to measure accurate position in x, y and z where the mean of the estimated position values is close to the actual values. The system has standard deviation about 0 due to noise filtration from calculation process in coding. Moreover, comparison with GPS system for outdoor environment which has accuracy about ±3 m, Thus, this positioning system is more reliable. The additional results of object detection are demonstrated in Figure 4.
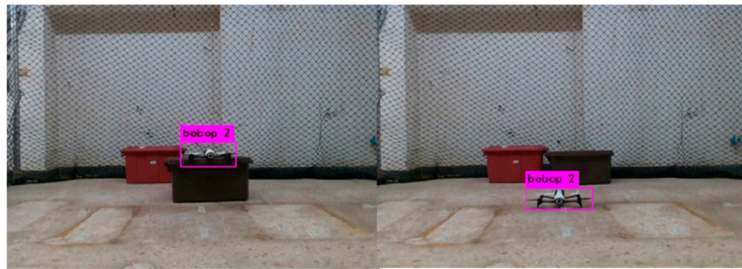
Figure 4. Result of object detection while static testing.

3.2. Flying Testing

Testing perform to estimate indoor position of UAV while flying and 3D trajectory plot with MATLAB program. Height (Y values from system) is altitude from ground in unit of meter, depth (Z from system) is distance from depth camera to UAV in unit of meter and X value is width from middle of camera. Moreover, the calculation process with fast process since position is update at a frequency less than 10 Hz and variance is acceptable. However, there is some mistake during testing due to model that is sometime unable to continuously detect. The result of flying test is shown in Figure 5.
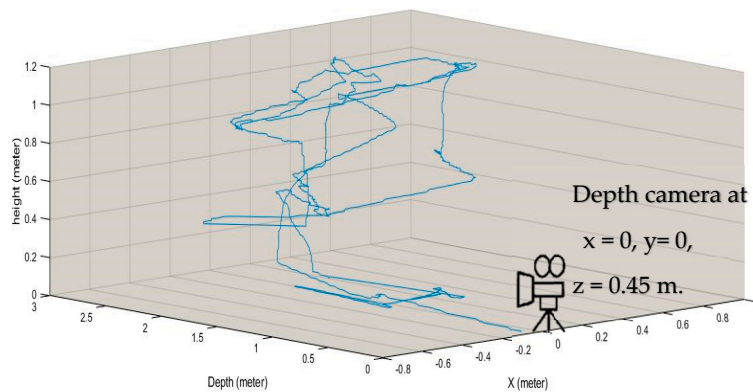

Figure 5. 3D trajectory plot of UAV by MATLAB with indoor positioning system.

4. Conclusions

UAV for indoor application has become increasing popular in application viz disaster release, mapping, indoor inspection etc. This research present one of choices to specify 3D position of UAV for indoor environment by using deep learning with Intel RealSense Depth camera D435i. The testing results show satisfy performance of system with low variance, high accuracy with maximum error about 10 cm. and fast by updating position at frequency less than 10 Hz. However, sometimes the camera system was unable to detect object due to trained model. Therefore, for future work, the trained model performance need to more improve by training more accuracy model. Moreover, filtration process will be applied with advanced techniques including Karman filter for more accuracy and this system will be applied to specify 3D position and navigation of swarm drone.

References

1　　Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv* **2016**, arXiv:1506.02640v5.

2　　Mainetti, L.; Luigi, P.; Ilaria, S. A survey on indoor positioning systems. In Proceedings of the 2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 17–19 September 2014.

3　　Marko B. YOLO ROS: Real-Time Object Detection for ROS. 2017. Available online: https://github.com/leggedrobotics/darknet_ros (accessed on 1 November 2019).

4　　Mautz, R. Overview of current indoor positioning systems. *Geodezija ir kartografija* **2009**, *35*, 18–22.

5　　Yasir, M.; Amelia, W.; Fajril, A. Indoor UAV Positioning Using Stereo Vision Sensor. *Procedia Eng.* **2012**, *41*, 575–579.

6　　Yohanes, K.; Izabela, N. A system of UAV application in indoor environment. *Prod. Manuf. Res. Open Access J.* **2016**, *4*, 2–22.

7　　Li, Y.; Scanavino, M.; Capello, E.; Dabbene, F.; Guglieri, G.; Vilardi, A. A novel distributed architecture for UAV indoor navigation. *Transp. Res. Procedia* **2018**, *35*, 13–22.