International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Feature-Based Stereo Vision Relative Positioning Strategy for Formation Control of Unmanned Aerial Vehicles

Yousef Yaghoobi, Muhammad Rijaluddin Bahiki, Syaril Azrad

***Abstract:** As inspired by birds flying in flocks, their vision is one of the most critical components to enable them to respond to their neighbor's motion. In this paper, a novel approach in developing a Vision System as the primary sensor for relative positioning in flight formation of a Leader-Follower scenario is introduced. To use the system in real-time and on-board of the unmanned aerial vehicles (UAVs) with up to 1.5 kilograms of payload capacity, few computing platforms are reviewed and evaluated. The study shows that the NVIDIA Jetson TX1 is the most suited platform for this project. In addition, several different techniques and approaches for developing the algorithm is discussed as well. As per system requirements and conducted study, the algorithm that is developed for this Vision System is based on Tracking and On-Line Machine Learning approach. Flight test has been performed to check the accuracy and reliability of the system, and the results indicate the minimum accuracy of 83% of the vision system against ground truth data.*

***Keywords:** Flight formation, unmanned aerial vehicle, vision system, on-line machine learning, leader-follower.*

I. INTRODUCTION

As the flying machines decussating with autonomous and also semi-autonomous machines, the result is the birth of the unmanned aerial vehicles (UAVs) that are predicted to change the way of human travel, cargo transport, security surveillance and emergency response, among others. In general, there are four categories of UAV: multi-rotor, fixed-wing, single-rotor and fixed-wing hybrid. Each of these types of UAV have their own advantages and disadvantages. Among them, multi-rotor UAVs have obtained the most attention, primarily due to their advantages including ease of use, vertical take-off and landing (VTOL), hover flight and capability to operate in the confined areas. However, their small payload capacity has been one of the concerned disadvantages [1].

Such operational limitation of the single multi-rotor UAVs has led to the interest in flight formation and formation control of UAVs, which has notably grown over the past few years [2]. One of the key aspects of flight formation is spatial coordination or relative positioning between UAVs that are flying in close proximity to each other in order to avoid collision and achieve collective operation. Vision has been a critical component for birds to be able to respond to their neighbor's motion when flying in flocks [3]. In similar notion, the focus of this study is on the vision aspect of the UAVs as a primary sensor for relative positioning in the flight formation of a Leader-Follower scenario. There are two main approaches in the application of vision-based technique in an on-board computation for spatial coordination of UAVs flying in swarm: color-based (artificial marker detection) that is discussed in [4]–[6], and motion-based (optical flow) that is discussed in [7]–[8]. The color-based approach performance is highly affected by misdetection for indoor application and light intensity variation for outdoor application. On the other hand, the motion-based or optical flow approach suffers from lack of precision and high sensitivity to noise for both indoor and outdoor applications. Meanwhile, there seems to be a lack of research on the use of a feature-detection approach for the collision-free UAV flight formation that can help to overcome those operational limitations. One of the main objectives in any vision-based navigation system is the ability to detect and/or to track the target(s) from other objects in a scene. In computer vision, at least one of the following terms is used: object detection, object recognition and object tracking. For object detection, it corresponds to the capability to notice or discover something within an image or video, which can be anything that is of interest such as a car, a person, an UAV and others. Meanwhile, object recognition is a step further than object detection by identifying the nature of the detected object. For instance, in addition to detecting the presence of a person in the image or video, the algorithm for object recognition must be able to recognize who the person is. In order to enable this capability, the algorithm should be furnished with some previous knowledge regarding what to be detected and recognized in the input image or video. There are various methods that can be applied for object recognition algorithm such as appearance-based, feature-based, template (or pattern) matching, and also machine learning methods [9].

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Yousef Yaghoobi, Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, Malaysia.

Muhammad Rijaluddin Bahiki, Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, Malaysia.

Syaril Azrad*, Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, Malaysia. Email: syaril@upm.edu.my

In the meantime, object tracking refers to corresponds to the ability to follow a movement or progress of interested object in a video. The primary goal of an object tracking algorithm is to keep an eye on the interested object in a scene and predict where it is heading to. The tracking algorithm must be able to handle the video frames in relation to one another, therefore it can create a location history for position change of an interested object over the certain period of time. This information is used to construct the object's motion model, which becomes an input to techniques such as Kalman Filter that help in the estimation of the past, present and future states of the tracked object. These estimations are essential in the tracking process as the tracked object might get occluded with other objects. By having an idea of the tracked object's probable future state or location, this facilitates in keeping the identity of the object across the video frames in spite of the occlusion. This paper presents the development of a vision system that is capable of acquiring, processing and reacting to the visual sensor's data for relative positioning between the members of flocking UAVs. The tracking algorithm has been selected to be used for the system and the developed vision system is then tested in a real-time data processing on-board of UAVs flying in swarms.

II. METHODOLOGY

The schematic process flow of the developed vision system for the formation flight is illustrated in Fig. 1. Once the vision system (VS) has been initialized, the image data representing the current scene observed by the follower UAV will be made available to the user. It should be noted that the UAV has been equipped with installed onboard camera that acts like its eye. At this stage, the VS will request the user to select the leader UAV. After the selection has been made, the tracking module (TM) will take over the entire process of tracking the selected leader UAV without requiring any further assistance from the user. Nonetheless, the user is still allowed to change the leader UAV when required. TM will locate the selected leader UAV in consecutive frame and then report its location to the relative positioning module (RPM). Using this information, RPM will calculate the location of the leader UAV relative to follower UAV. Once the coordinates of leader UAV is known, the data extraction and transfer module (DETM) will extract location data and transfer it to the flight controller. For this test case study, there is only one leader UAV and one follower UAV. However, the system can also work in cases where there are more than two members in the flock or in situations where the leader UAV has to leave the group due to some reasons. In the latter situation, the flexibility in selecting new leader UAV by the user allows the flight formation to continue seamlessly and without any notable disruption. Fig. 2 depicts the main components of TM that is adopted from the tracking algorithm that is developed in [10]. In short, multiscale filter bank is an array of filters which separates the input samples into their different categories according to their scale. Multiscale image features are the Haar-like features that have been extracted from the input samples while the sparse measurement matrix is the matrix of random numbers. In the meantime, the compressed vectors are the projected Haar-like features from the image features' space (i.e. original Haar-like features from the sample image) into the low-dimensional or compressed space. Last but not least, classifier is an algorithm that handles the task of classification that identifies whether a new set of images is an object or a background. On the other hand, Fig. 3 shows the schematic process flow for the RPM. It starts with the reception of the coordinates for location of the leader UAV from TM. RPM will then use this information to calculate the relative location of leader UAV to the follower UAV using the triangulation method. The output from RPM is the three-dimensional relative distance between the leader UAV and the follower UAV, which is recorded as a .csv file for comparison with the ground truth, and the depth information will be sent to DTEM.

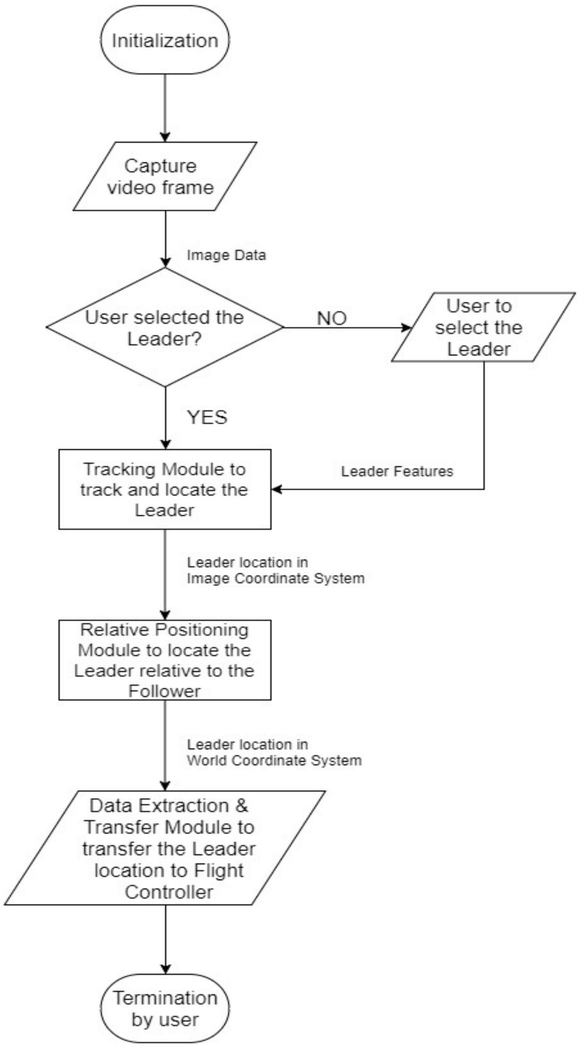
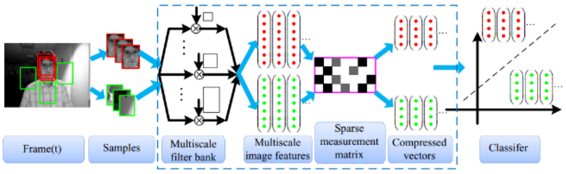
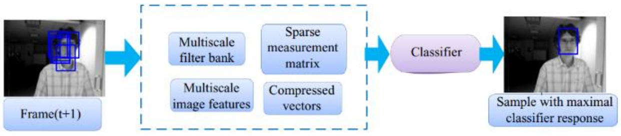


Fig. 1. Process flow of the developed vision system



(a) Updating classifier at t-th frame



(b) Tracking at (t+1)-th frame

Fig. 2. Main components of the tracking module [10]

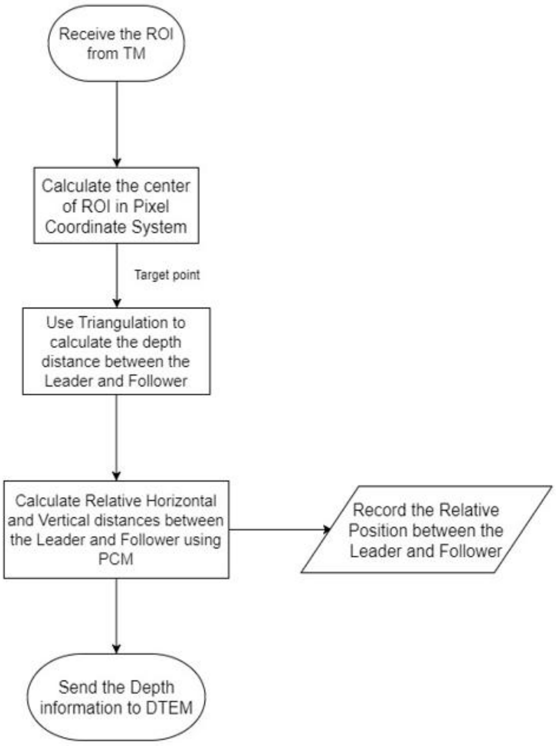


Fig. 3. Process flow of relative positioning module

In order to test this VS operation, a well-established flight controller unit is used. A reliable communication module that is known as Socket Programming is applied to bridge between the VS module and the flight controller unit because they are on different platform and programming language. Moreover, to carry out the real-time application of the developed VS, the programming platform is constructed upon Jetson TX1 as the on-board processing unit (OBPU) with a ZED stereo camera from Stereolabs as the primary vision interpretation medium. The technical specifications for Jetson TX1 and ZED stereo camera are tabulated in Table I and Table II, respectively.

Table-I: NVIDIA-Jetson TX1 technical specifications

Technical Unit	**Specification**
Central Processing Unit (CPU)	Quad ARM® A57/2 MB L2
Graphics Processing Unit (GPU)	NVIDIA Maxwell ™, 256 CUDA cores
Processing Power	1024 GFlops
Operating System	Ubuntu 14.04
Compatibility with *ZED*	YES
Compatibility with *ZED* SDK	YES

Table-II: ZED technical specifications

Technical Parameter	**Specification**
Frame Rate	60 fps
Image Resolution	720 p
Field of View (FoV)	110˚
SDK availability	YES
Weight	159 g

Two platforms are used for the flight formation: one is the leader UAV while another is the follower UAV. It should be noted that there is no direct communication between leader and follower UAVs, and the developed VS is the only sensing tool that is utilized by the follower UAV to navigate through the environment to follow the leader UAV. As shown in Fig. 4, the leader UAV platform used in this study is the DJI F450 quadcopter with cube flight controller. Its specifications are tabulated in Table III. Meanwhile, for the follower UAV platform, which carries the VS hardware, the DJI MATRICE 100 quadcopter is utilized and it is depicted in Fig. 5 while its specifications are tabulated in Table IV.

Fig. 4. Leader UAV platform

Table-III: Leader UAV platform specifications

Technical Parameter	**Specification**
Model	Flame Wheel 450 (F450)
Frame Weight	282g
Diagonal Wheelbase	450mm
Takeoff Weight	800g ~ 1600g
Recommended Propeller	10 × 3.8in; 8 × 4.5in
Recommended Battery	3S~4S LiPo
Recommended Motor	22 × 15mm or 22 × 12mm
Recommended ESC	15A OPTO

Fig. 5. Follower UAV platform

Table-IV: Follower UAV platform specifications

Technical Parameter	Specification
STRUCTURE	Diagonal Wheelbase: 650 mm Weight (with TB47D battery): 2355 g Max. Takeoff Weight: 3600 g Expansion Bay Weight: 45 g Battery Compartment Weight: 160 g
PROPULSION SYSTEM	Motor Model: DJI 3510 Propeller Model: DJI 1345s ESC Model: DJI E SERIES 620D
BATTERY	Model: TB47D Capacity: 4500 mAh Voltage: 22.2 V Type: LiPo 6S Energy: 99.9 Wh Net Weight: 600 g
PERFORMANCE	Hovering Accuracy: Vertical: 0.5 m, Horizontal: 2.5 m Max. Angular Velocity: Pitch: 300°/s, Yaw: 150°/s Max. Tilt Angle: 35° Max. Speed of Ascent: 5 m/s Max. Speed of Descent: 4 m/s Max. Wind Resistance: 10 m/s Max. Speed: 22 m/s

Last but not the least, the testing platform used to verify and validate the acquired results from the VS can be divided into two main subsystems: OptiTrack and Ground Control Station (GCS), located in the Satellite and Space System Laboratory of the Department of Aerospace Engineering, Universiti Putra Malaysia. OptiTrack is a system that is consisted of mesh of infrared (IR) cameras, which can be considered as an indoor GPS but with a much higher accuracy. The OptiTrack system that is available in the lab has 16 cameras in total, which are synced and formed a loop via three routers. The obtained data gets transferred to the GCS, which runs the Motive tracking software. In order to get the relative position (i.e. relative to an arbitrary reference point) of an object, which in this case is the UAV platform, special reflective markers that can reflect emitted IR beams from cameras are attached to the platforms. The camera arrangements and the GCS user interface are as shown in Fig. 6 and Fig. 7, respectively.

Fig. 6. OptiTrack camera arrangement

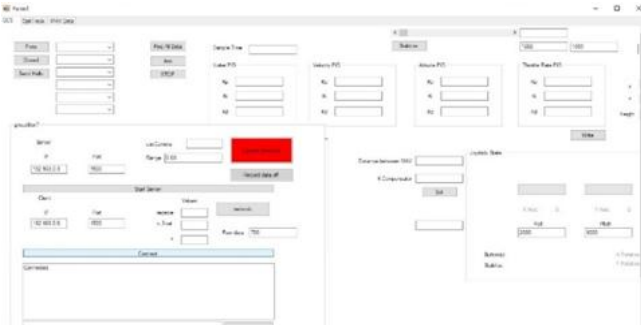


Fig. 7. GCS user interface

A leader-follower scenario is defined to put the developed VS into real-time testing. For this test case study, the follower UAV is set to be hovering at a height of about 1 m above the ground and the leader UAV is flying a random trajectory path. The accuracy of the VS is measured from the obtained results in this case study.

III. RESULTS AND DISCUSSION

The recorded data from the testing consists of two sets: one is three-dimensional position of the leader UAV that has been acquired from OptiTrack system (serving as the ground truth) and the other is three-dimensional position of the leader UAV relative to the follower UAV as obtained from the VS. Each set of data contains three values that represent the horizontal (X), vertical (Y) and depth (Z) information, and their plots are presented in Fig. 8, Fig. 9 and Fig. 10, respectively.

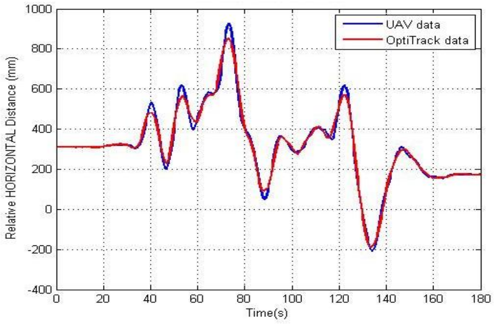


Fig. 8. Horizontal position (X) data

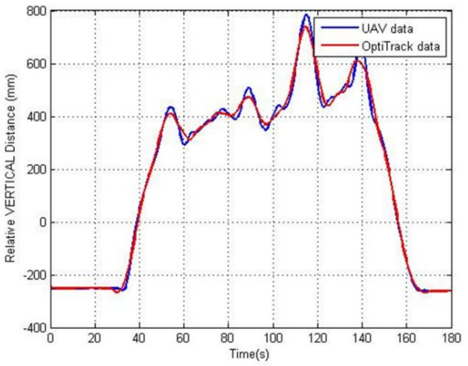


Fig. 9. Vertical position (Y) data

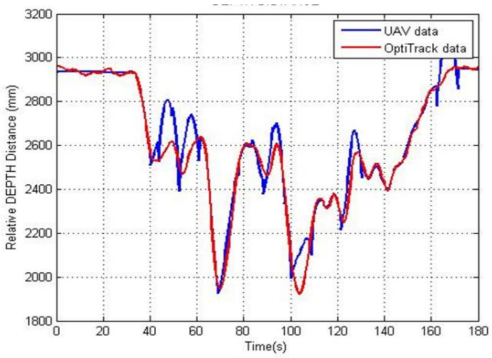


Fig. 10. Depth position (Z) data

The data from OptiTrack system is used as the benchmark in measuring the accuracy of the VS. From the obtained data, the mean variation of RPM in the horizontal direction is about 2.13% and the maximum variation is 10.74%. Meanwhile, the mean variation of RPM in the vertical direction is 1.95% and the maximum variation is 9.67%. Last but not least, the mean variation of RPM in the depth direction is roughly 2.33% and the maximum variation is 16.62%. The low accuracy in depth direction can be contributed to the fact that the three reflective markers that are pasted on leader UAV platform form a rigid body representing it with a single pivot point and the Motive uses that point to calculate its three-dimensional position. In constrast, the RPM uses the center of rectangle that is reported by TM as the representative of the leader UAV platform and thus, when the center points fall on the arm of platform, there is a difference of about 15 – 20 cm in measurement. Overall, the VS accuracy and the variation between VS and OptiTrack are summarized in Table V.

Table-V: Summary of VS performance

Direction	**Variation (%)**		**Accuracy (%)**	
	Mean	**Maximum**	**Mean**	**Maximum**
Horizontal (X)	2.13	10.74	97.87	89.26
Vertical (Y)	1.95	9.67	98.05	90.33
Depth (Z)	2.33	16.62	97.67	83.38

IV. CONCLUSION

The development of VS based on the feature-detection and stereo vision has been presented in this paper. The developed VS is capable to be utilized on-board of UAVs for real-time data processing. One of the targeted applications for such a system is as a visionary aid for the flight formation of UAVs, specifically in leader-follower flight formation. Based on the testing results, the minimum accuracy of the developed VS is shown to be 83.38%, which is obtained for the depth direction data. Overall, this can be taken to indicate very encouraging performance result for the application of the developed VS in with UAVs, especially for flying in swarms.

ACKNOWLEDGMENT

This research is supported by Fundamental Research Grant Scheme, FRGS/1/2015/TK09/UPM/02/04. The authors fully acknowledge the Ministry of Higher Education (MOHE) and Universiti Putra Malaysia for the approved fund, which makes this research possible.

REFERENCES

1. A. Chapman, “Types of drones: multi-rotor vs fixed-wing vs single rotor vs hybrid VTOL,” *Australian DRONE Magazine*, vol. 1, 2016, pp. 27-35
2. F. Borrelli, T. Keviczky, and G. J. Balas. “Collision-free UAV formation flight using decentralized optimization and invariant sets,” *43rd IEEE Conference on Decision and Control*, Nassau, Bahamas, 2004
3. R. Vidal, O. Shakernia, and S. S. Sastry. “Distributed formation control with omnidirectional vision-based motion segmentation and visual servoing,” *IEEE Robotics & Automation Magazine*, vol. 11, 2004, pp. 1-13
4. R. Jin, and J. Wang. “A vision tracking system via color detection,” *12th IEEE International Conference on Control and Automation (ICCA)*, Kathmandu, Nepal, 2016
5. D. Dias, P. Lima, and A. Martinoli. “On-board vision-based 3D relative localization system for multiple quadrotors,” *IEEE International Conference on Robotics and Automation*, Stockholm, Sweden, 2016
6. K. Sapkota, “Vision-based unmanned aerial vehicle detection and tracking for sense and avoid systems,” *International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, South Korea, 2016
7. N. Gageik, M. Strohmeier, and S. Motenegro, “An autonomous UAV with an optical flow sensor for positioning and navigation,” *International Journal of Advanced Robotic Systems*, vol. 10, 2013
8. Q. Juntong, Y. Naixin, and L. Xiang, “A UAV positioning strategy based on optical flow sensor and inertial navigation,” *IEEE Conference on Unmanned Systems (ICUS)*, Beijing, China, 2017
9. S. Achatz, *State of the Art of Object Recognition Techniques*, Technical University of Munich, Germany
10. K. Zhang, L. Zhang, and M. Yang, “Real-time compressive tracking,” *12th European Conference on Computer Vision*, Florence, Italy, 2012

AUTHORS PROFILE

Yousef Yaghoobi has obtained his Bachelor degree in mechanical engineering from Universiti Tenaga Nasional (UNITEN), Malaysia and MSc. degrees in the aerospace engineering from Universiti Putra Malaysia, Malaysia in 2013 and 2019, respectively. His research interests also include unmanned aerial vehicles and computer vision.

Muhammad Rijaluddin Bahiki received his Bachelor and MSc. degrees in the aerospace engineering field from Universiti Putra Malaysia in 2014 and 2018, respectively. Among others, his research interests include vision-based control, formation control and collision avoidance system of unmanned aerial vehicles.

Syaril Azrad received his B. Eng. and also M. Eng. from the Tokyo University of Science in 2000 and 2002, respectively. He received his Ph.D. degree in Mechanical Engineering in 2012 from Chiba University based on his research work on vision-based control of unmanned air vehicle (UAV). Currently, he is a senior lecturer in the Department of Aerospace Engineering, Universiti Putra Malaysia (UPM). His research interests are in unmanned aerial systems that include the areas of high-altitude balloon, cube satellite, and airborne wind energy system. Moreover, he is a committee member representing Malaysia and also UPM in the South East Asia Network on Aerospace Engineering (SANAE). Additionally, he is also UPM representative for the Belt and Road Aerospace Innovation Alliance (BRAIA).