Proceedings of the IEEE
International Conference on Information and Automation
Ningbo, China, August 2016

Real-Time Tracking a Ground Moving Target in Complex Indoor and Outdoor Environments with UAV*

Shiyu Chen, Shanshan Guo and Yanjie Li

School of Mechanical Engineering and Automation
Harbin Institute of Technology Shenzhen Graduate School
Shenzhen, Guangdong Province, China

Email: lyj@hitsz.edu.cn

***Abstract* – In this paper, we present a complete strategy of tracking a ground moving target in complex indoor and outdoor environments with an unmanned aerial vehicle (UAV) based on computer vision. The main goal of this system is to track a ground moving target stably and get the target back when it is lost. An embedded camera on the UAV platform is used to provide real-time video stream to the onboard computer where the target recognition and tracking algorithms are running. A vision-based position estimator is applied to localize the position of UAV. According to the position of the ground target obtained from the 2-dimensional images, corresponding control strategy of the UAV is implemented. Simulation of UAV and ground moving target with the target-tracking system is performed to verify the feasibility of the mission. After the simulation verification, a series of real-time experiments are implemented to demonstrate the performance of the target-tracking strategy.**

Index Terms – Tracking, Control strategy, Unmanned aerial vehicle, Computer vision.

I. Introduction

Target-tracking systems have been applied in various fields, such as traffic monitoring [1], military surveillance [2] and disaster relief [3] in recent years. Especially, with the development of UAV and computer vision, a series of intelligent target-tracking systems based on UAV have been proposed [4], [5]. Generally, three important parts are essential in a target-tracking system, including target-tracking algorithms, accurate position estimator of UAV and the control strategy of the system.

In order to obtain the relative location of the moving target from the real-time video stream, a target detection method is indispensable. Many kinds of image features can be used to detect the object. Edge detection method [6], [8] can lock the targets which have regular shapes, such as rectangle or ellipse. However, the objects in real environment have various appearances and even change their shapes along with time. Thus, other characters are required to determine the position of the target. In [7], local binary feature of the image has been used to make pixel comparisons to obtain a robust character of the target which does not change even if the scale of the target has transformed. Also, the target-tracking algorithm is important to estimate the motion of target. The major method is pyramid optical flow [9], [10] which is integrated in OpenCV library. Because of the lack-of-memory property of the target-tracking algorithm, the combination of the tracking algorithm and detection method is necessary for good performance [7].

All the video streams are got from the embedded camera on the UAV platform. So, a precise estimator for the position of UAV can ensure the stability of the target-tracking system. GPS is very popular to provide the global position [11] when UAV works in outdoor environment. However, the UAV usually performs a task indoor. With the development of computer vision, various vision-based position estimators have been proposed recently [12], [13], [14]. Multi-sensor fusion algorithm [15] based on extended Kalman filter (EKF) can fuse the information from the inertial measurement unit (IMU), GPS and vision-based position estimator to make the localization of UAV more accurate. Visual odometry [16] and simultaneous localization and mapping (SLAM) [17] methods are also used to get the pose of UAV.

In this paper, a complete control strategy for the target-tracking system is proposed. Firstly, we use the dynamics model of the quadrotor in [18] to set up the model of UAV. Then, the target-tracking problem is divided into three kinds of control modes, including target-searching mode, target-tracking mode and target-lost mode. For each mode, a corresponding control strategy is designed and the switch between different modes is operated well. Finally, simulations and experiments are implemented to verify the feasibility of the strategy.

The paper is organized as follows. The system overview is presented in Section II. In Section III, the process of recognizing and tracking the target is given. Section IV describes the UAV localization in complex environment. Section V presents the strategy of real-time tracking. Then, simulations and experiments are described in Section VI. Finally, Section VII concludes this paper and proposes some future research works.

II. System Overview

As shown in Fig. 1, the whole system framework considered in this paper consists of three parallel parts, including target recognition and tracking, position estimator of UAV and control strategy of the system. The target

* This work is partially supported by Shenzhen basic research programs, No. JCYJ20140417172417158, No. JCYJ20150731105106111 and No. JCYJ20140901003938996.

978-1-5090-4102-2/16/$31.00 ©2016 IEEE

recognition and tracking part provides relative position of the target in 2-dimensional images for the system control strategy part and combine with the position of UAV from the position estimator to implement a path planning of the UAV. Then, the UAV will follow the path to track the target immediately. All the parts are running synchronously on an onboard computer attached on UAV.

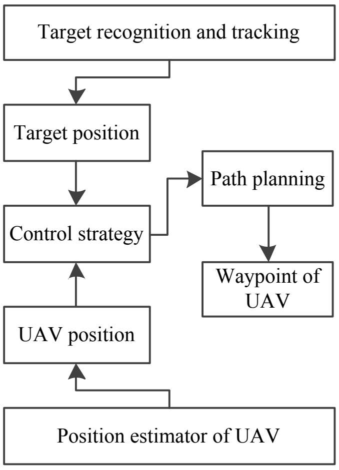

Fig. 1 System framework.

A. *Target Recognition and Tracking*

In this part, the robust and real-time target-tracking algorithm [7] is used to estimate the motion and relative position of the ground moving target in video stream. Both target detection and tracking methods are integrated in the algorithm, and a precise position can be got through synthesizing the two methods for different situations. An online learning algorithm makes the result more accurate.

B. *Position Estimator of UAV*

An accurate pose of UAV can make the target-tracking system performs more stable. Multi-sensors, such as GPS, optical flow camera and IMU, have been used to estimate the position of the quadrotor. A vision-based SLAM method [17] is adopted to obtain the position of the quadrotor in this paper. In order to enhance the robustness of the position estimator in complex indoor and outdoor environments, we use EKF algorithm to fuse the information of GPS, IMU and optical flow method together to calculate the position.

C. *Control Strategy*

The function of control strategy is to plan a path for the quadrotor to follow the ground moving target real-timely. Three modes are proposed to adapt different conditions of the target-tracking system. Control methods have been designed for each mode and the target can be tracked stably for a long time.

III. TARGET RECOGNITION AND TRACKING

The quadrotor needs to get the motion and relative position of the ground moving target before tracking it. We adopt a robust target recognition and tracking method based on an online learning algorithm [7]. This method consists of four components, including detection, tracking, learning and integration. The relationship between these components is shown in Fig. 2.

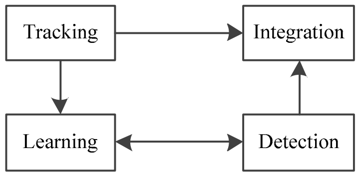

Fig. 2 Target recognition and tracking algorithm.

In the initialization stage of the algorithm, we need to select a target through the mouse. Then, the target model database is established by some basic image transformation of the target patch. At the same time, the tracking component starts to estimate the motion of the target between consecutive frames immediately. But the tracker may be fail and never recovery by itself once the target disappeared from the video stream. Therefore, we need the detection component to determine the position of the target in every independent frame.

In the detection component, we divide the image frame from the real-time video stream into many small patches. Then, the variances of the patches are calculated to roughly exclude the background patches. The patches with small variances, which may be the background patches will be abandoned and the rest of image patches will be collected to compare with the target model database by local binary feature. The patch which carries the feature most similar with the target will be selected as the target patch finally.

The information from detection and tracking components may be different sometimes, and the integration component can choose the closer result by compare them with the target model database respectively. Also, the integration component is used to recover the tracker when it fails to track the target.

Both detector and tracker are corrected instantly based on the online learning algorithm which is used to adjust the parameters of the both components. The positive sample also increases constantly through online learning, as is shown in Fig. 3. The target may change its direction along with time, but the detector can always extract the position of it by the updated parameters. Then, the target patch will be added to the positive sample space. Therefore, the sample space can always storage the latest positive sample of the target and the target recognition and tracking algorithm can be applied more generally and accurately to track a ground moving target for a long time.

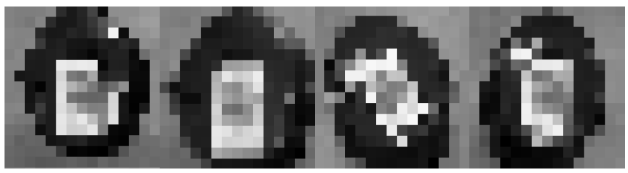

Fig. 3 The positive sample of a ground target.

IV. UAV Localization

Because of the complete autonomy of the target-tracking system, the quadrotor needs to know its accurate position in the world frame. The error of the information from GPS limits its application in the localization of UAV. In this paper, we consider the system works in complex indoor and outdoor environments. Therefore, vision-based position estimator is more suitable to provide the position data for UAV. In our system, we use a monocular SLAM method based on the ORB feature [17] to obtain the localization of the quadrotor. As shown in Fig. 4, a large number of features can be obtained in different scenes to estimate the motion of the camera and the localization of UAV can be computed based on the state of the camera. Since the information from the SLAM method may be invalid sometimes when the images in the video stream lack feature points. Thus, we may use a multi-sensor fusion algorithm [15] to compensate the fault. The architecture of UAV localization is shown in Fig. 5.

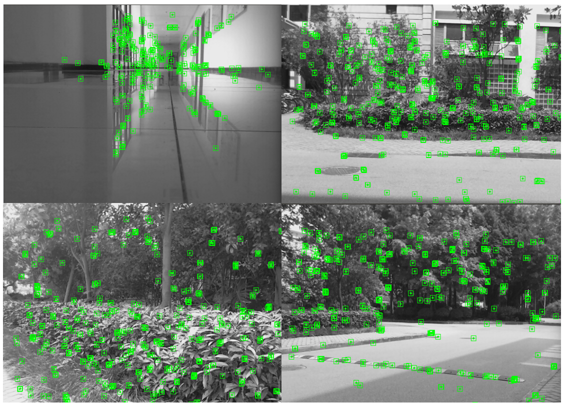

Fig. 4 ORB features in different scenes.

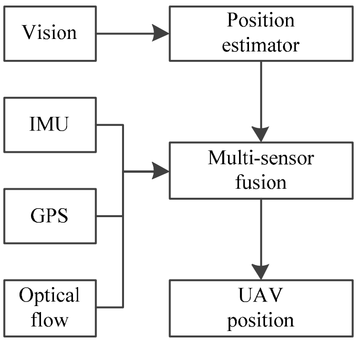

Fig. 5 UAV localization architecture.

V. Real-time Tracking Strategy

With the purpose of real-time tracking and getting the target back in the video stream when it is lost, the target-tracking strategy is divided into three different modes, as is shown in Fig. 6.

1) Target-searching: The quadrotor needs to take off and fly to the desired altitude firstly. Then, the quadrotor begins to search the target through the embedded camera attached on the quadrotor platform.

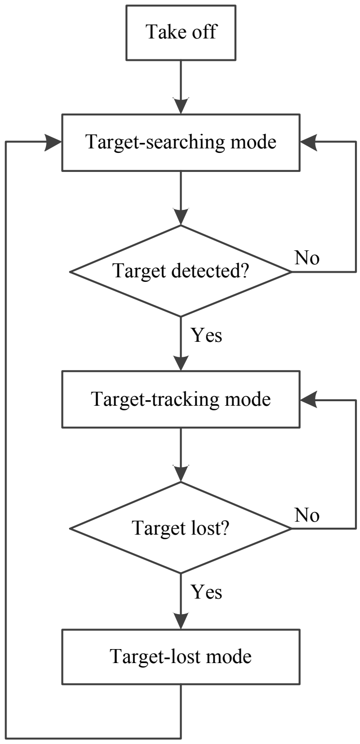

Fig. 6 The target-tracking strategy.

2) Target-tracking: In this mode, the target-tracking algorithm provides the relative-position of the ground target in the 2-dimensional images. Then, the planning path of the quadrotor can be generated real-timely according to the target position. The method of creating the path is shown in Fig. 7.

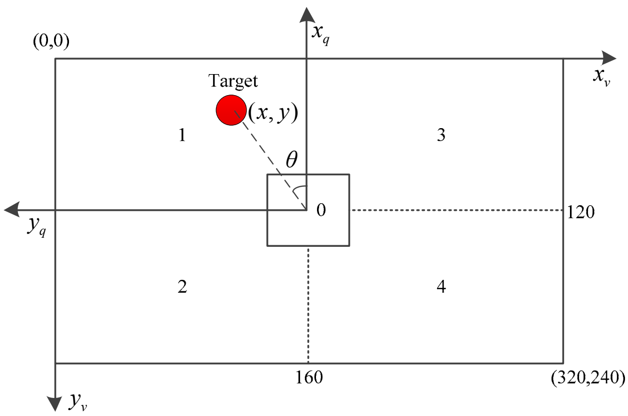

Fig. 7 Path planning of UAV based on 2-dimensional image.

The coordinate system (x_v, y_v) is established for 2-dimensional image and (x_q, y_q) is the reference frame of UAV. The image is divided into five regions, so that different strategies can be created to calculate the increase value of the position $(\Delta x, \Delta y)$ for the quadrotor. Suppose that the target is at the region 1 and the position of the target is (x, y). Then, the waypoint of the planning path (x_t, y_t) can be computed as follows:

$$x_t = x + \Delta x \tag{1}$$

$$y_t = y + \Delta y \tag{2}$$

$$\tan\theta = \frac{160 - x}{120 - y} \tag{3}$$

$$\Delta y = \Delta x \cdot \tan\theta \tag{4}$$

Following the waypoint, the quadrotor can track the ground moving target at a fixed altitude.

3) Target-lost: When the ground moving target is lost from the video stream, the quadrotor would increase its altitude to extend the vision of the camera until the target comes back again. As shown in Fig. 8, the highest altitude can ensure the safety of the target-tracking system.

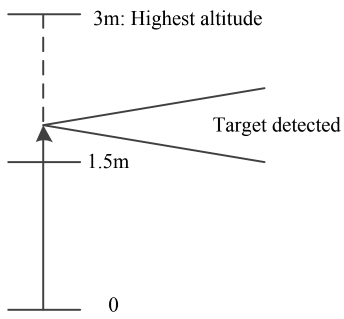

Fig. 8 The control strategy of target-lost mode.

VI. Simulation and Experiments

To verify the proposed target-tracking strategy, we develop a simulation platform and all the algorithms can be operated on it. After good performance is obtained in simulations, a series of real-time experiments have been implemented in complex indoor and outdoor environments to demonstrate the robustness and accuracy of the strategy.

A. *Simulation*

The simulation platform [19] is established based on Gazebo. A camera has been attached to the quadrotor model and the video sequence is published at 30 fps with a resolution 320*240. The robot named Innok-Heros is added to the world frame as the ground moving target. All the algorithms use robot operating system (ROS) as the interfacing robotics middleware.

In order to verify the controllability and stability of the quadrotor model, a test of flying a short line is put into effect. The quadrotor takes off and flies along with a line which is 3 meters long. The fly trajectory of the model is shown in Fig. 9. The position (x, y, z) has been described respectively in Fig. 10.

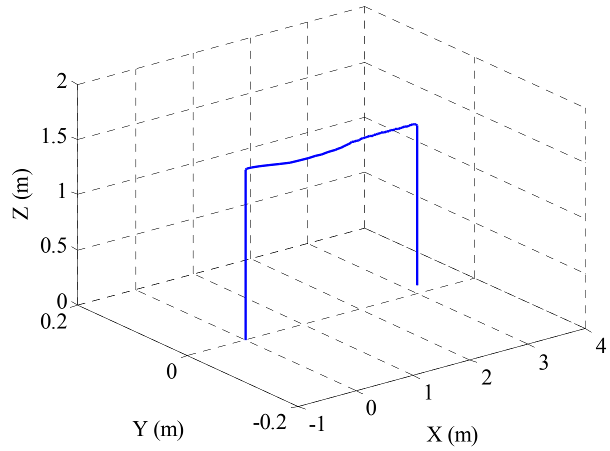

Fig. 9 The fly trajectory of the quadrotor model.

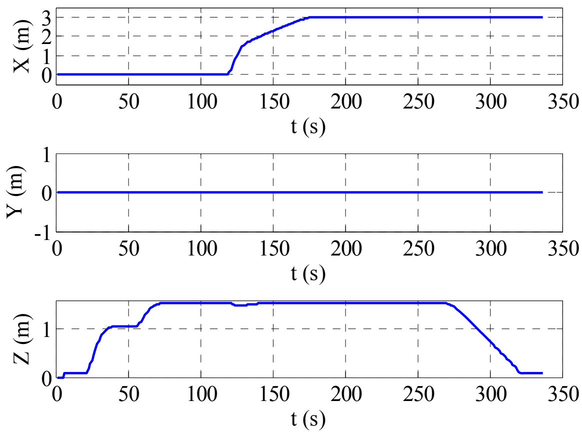

Fig. 10 The position of the quadrotor model.

Then, the target-tracking system is operated based on the simulation platform. Fig. 11 shows the components of the system which include a UAV model and a ground moving robot. The target-tracking trajectory is shown in Fig. 12 and the position of the model is shown in Fig. 13. The result of the simulation demonstrates that the strategy proposed for the target-tracking system is feasible.

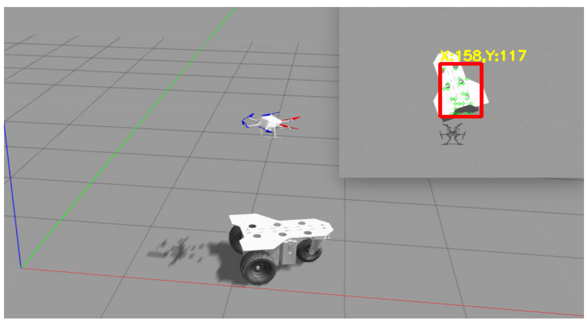

Fig. 11 Components of the target-tracking system.

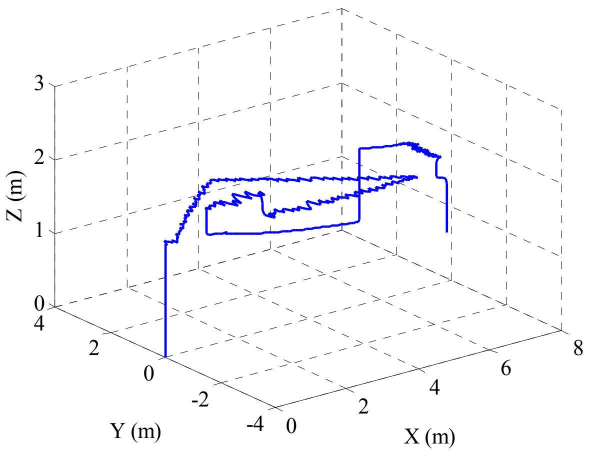

Fig. 12 The target-tracking trajectory of UAV model.

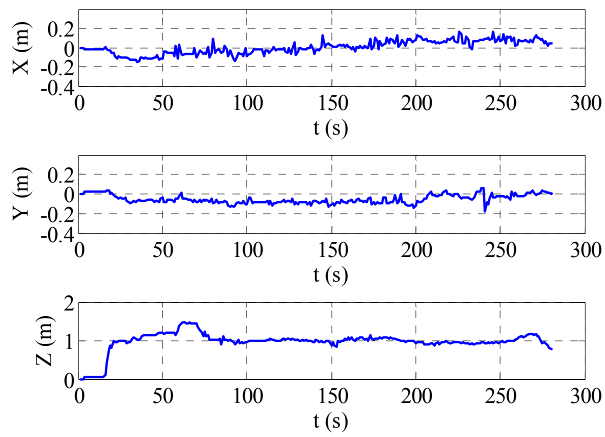

Fig. 14 The values of x, y, z axis in hovering experiment.

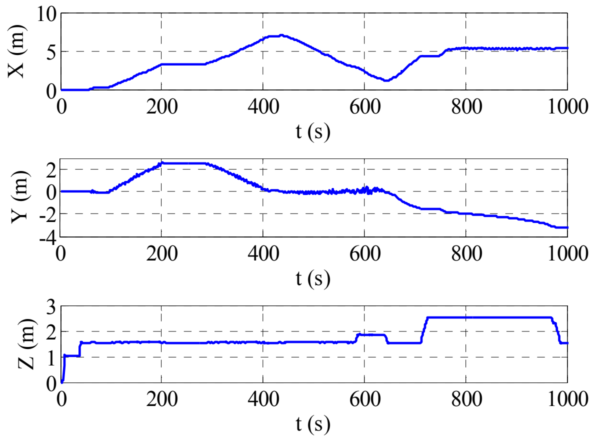

Fig. 13 The position of UAV model during target-tracking.

B. Experiments

The experimental platform is based on a quadrotor equipped with the Pixhawk autopilot consisting of an IMU and a programmable ARM-Cortex-M4 microcontroller. The main computation unit onboard is an Intel NUC with a 3.1 GHz Core i7 processor with 16 GB of RAM and a 256 GB SSD. A high frame rate camera PS3-Eye is used to capture real-time video sequence. ROS tool is the interface middleware for all the running algorithms. In order to avoid the time delay of the information interaction between quadrotor and images, all apps are operated on the onboard computer.

Before the target-tracking system is run, a hovering experiment is implemented to verify the accuracy of position estimator for UAV. As shown in Fig. 14, the values of x, y, z axis are stable around a constant and the error is in an acceptable range.

In the experiment of tracking a ground moving target with UAV, a small red robot is selected as the target and the position is published to ROS topics by target-tracking algorithm, as is shown in Fig. 15. The coordinate of the target is based on the 2-dimensional images. The target-tracking trajectory is collected and described in Fig. 16 and Fig. 17 shows the position of UAV at x, y, z axis. The process of tracking a ground moving target by the quadrotor is reproduced in these figures. The quadrotor can fly to a desired altitude and tracks the target at this fixed altitude. At 102 second, the ground target moved quickly to escape from the video sequence. Then, the tracking strategy turns into target-lost mode, the quadrotor can increase its altitude to search the target and get the target back immediately. When the target stays at a fixed area, the quadrotor can fly to the above of the target and decreases its altitude to initial position at 1.5m, and then, hover at current position. According to the results of extensive experiments, the performance of the target-tracking system can be demonstrated stably.

Fig. 15 The coordinate of target in 2-dimensional images.

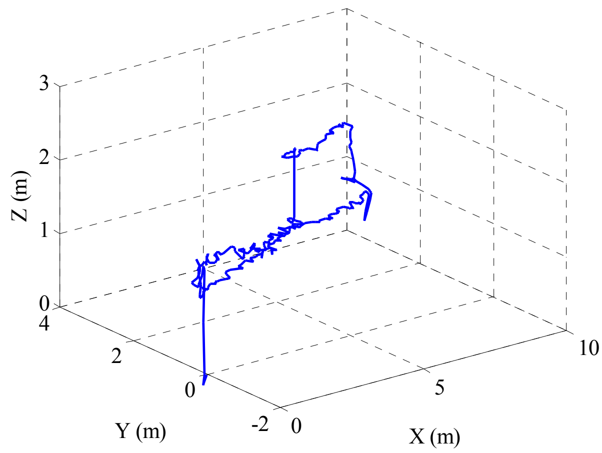

Fig. 16 The target-tracking trajectory of the system.

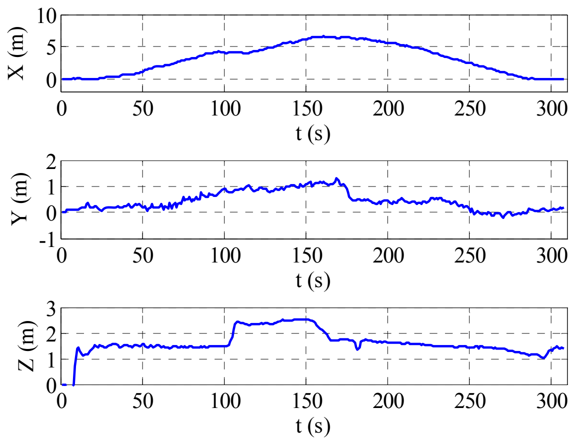

Fig. 17 The position of UAV at x, y, z axis.

VII. Conclusion and Future Work

In this paper, we have proposed a complete control strategy for target-tracking system with UAV. Three different control modes were discussed respectively. A robust target recognition and tracking method has been used to obtain the relative position of the ground moving target. A vision-based position estimator was applied to determine the real-time position of the quadrotor. A large number of experiments were successfully performed in indoor and outdoor environments which show that the quadrotor was stable and can follow the ground moving target real-timely in the target-tracking system. So, the control strategy presented is a feasible and stable scheme to the field of target-tracking. Future work will focus on the control algorithms of UAV to increase its robustness and accuracy when executes a mission in complex environment. Also, the target-tracking algorithm will be studied by combining the computer vision with machine learning.

References

[1] S. Gil, R. Milanese, and T. Pun, "Comparing features for target tracking in traffic scenes," *Pattern Recognition*, vol. 29, no. 8, pp. 1285-1296, 1996.

[2] A. Dukeman, and M. Anderson, "A CSP solution to multi-camera surveillance and target tracking," *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 97-102, 2011.

[3] D. Q. Abell, *Mobile Tracking and Location Awareness in Disaster Relief and Humanitarian Assistance Situations*, Thesis Collection, 2012.

[4] H. Choi, and Y. Kim, "UAV guidance using a monocular-vision sensor for aerial target tracking," *Control Engineering Practice*, vol. 22, no. 1, pp. 10-19, 2014.

[5] B. R. Trilaksono, et al, "Hardware-in-the-loop simulation for visual target tracking of octorotor UAV," *Aircraft Engineering and Aerospace Technology*, vol. 83, no. 6, pp. 407-419, 2011.

[6] J. E. Gomez-Balderas, et al, "Tracking a ground moving target with a quadrotor using switching control," *Journal of Intelligent and Robotic Systems*, vol. 70, no. 1-4, pp. 65-78, 2013.

[7] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1409-1422, 2011.

[8] S. Lin, A. G. Matthew, and J. L. Andrew, "Real-time 6DoF deck pose estimation and target tracking for landing an UAV in a cluttered shipboard environment using on-board vision," *IEEE International Conference on Mechatronics and Automation*, 2015.

[9] Z. Wang, C. Li, and Q. Liu, "Improved global optical flow estimation with mean shift algorithm for target detection," *International Journal of Wireless and Mobile Computing*, vol.8, no. 4, pp. 323-328, 2015.

[10] T. Bengtsson, T. Mckelvey, and K. Lindström, "Optical flow estimation on image sequences with differently exposed frames," *Optical Engineering*, vol. 54, no. 9, 2015.

[11] C. S. Yoo, and I. K. Ahn, "Low cost GPS/INS sensor fusion system for UAV navigation," *IEEE Digital Avionics Systems Conference*, vol. 2, pp. 8.A.1 - 8.1-9, 2003.

[12] W. Ding, J. Wang, and A. Almagbile, "Adaptive filter design for UAV navigation with GPS/INS/optic flow integration," *IEEE International Conference on Electrical and Control Engineering*, pp. 4623-4626, 2010.

[13] A. Soloviev, "Fusion of inertial, optical flow, and airspeed measurements for UAV navigation in GPS-denied environments," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 7332, pp. 733202-733212, 2009.

[14] A. R. Vetrella, et al, "Cooperative UAV navigation based on distributed multi-antenna GNSS, vision, and MEMS sensors," *International Conference on Unmanned Aircraft Systems*, 2015.

[15] S. Lynen, M. W. Achtelik, et al, "A robust and modular multi-sensor fusion approach applied to MAV navigation," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol.11, pp. 3923-3929, 2013.

[16] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO: fast semi-direct monocular visual odometry," *IEEE International Conference on Robotics and Automation*, pp. 15-22, 2014.

[17] R. Mur-Artal, J. M. Montiel, and J. D. Tardos, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147-1163, 2015.

[18] P. Castillo, A. Dzul, and R. Lozano, "Real-time stabilization and tracking of a four-rotor mini rotorcraft," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 4, pp. 510-516, 2004.

[19] F. Furrer, M. Burri, M. Achtelik, et al, *RotorS—A Modular Gazebo MAV Simulator Framework. Robot Operating System (ROS)*, Springer International Publishing, 2016.