12th IEEE International Conference on Control & Automation (ICCA)
Kathmandu, Nepal, June 1-3, 2016

A Binocular Vision-based Measuring System for UAVs Autonomous Aerial Refueling

Yimin Deng, Ning Xian, and Haibin Duan, *Senior Member*, *IEEE*

***Abstract*—In this paper, we present the systematic design and implementation of a binocular vision-based measuring system for autonomous aerial refueling. The hardware configuration of the verification platform is presented, and vision algorithms including feature extraction and pose estimation are employed for estimating the relation between two rotorcrafts. To verify the autonomous aerial refueling of unmanned aerial vehicles, a binocular vision system and an on-board data processing computer are utilized to provide the real-time pose information. Series of experiments are conducted to demonstrate the feasibility and effectiveness of the overall platform.**

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), especially small-size UAVs, have been widely utilized to execute various missions in both military and civilian domains. Owing to its vertical takeoff-and-landing, hovering, and maneuvering capabilities, an unmanned rotorcraft equipped with a vision system can easily perform a wide range of tasks, such as target detection [1] and visual tracking [2]. Many research efforts have also been devoted to formulation [3]-[4], vision-based navigation [5], air combat [6], and so on.

To extend UAV's operational radius, Autonomous Aerial Refueling (AAR) has aroused much interest [7]. To obtain the accurate pose information between UAVs is the premise of AAR. Vision-based measuring system is enabling UAVs to implement pose estimation during the docking phase. A CCD camera with an infra-red filter is used to identify the LEDs on a drogue [8]. The VisNav sensor has been adopted for vision navigation and pose estimation [9]. A boom and receptacle AAR system is developed using the visual snake optical sensor [10]. A hardware-in-loop simulation platform is established on the basis of computer vision [11]-[12]. A real-time visual sensing system is designed for the problem of semi-autonomous docking within aerial refueling for UAVs [13].

* This work was partially supported by National Natural Science Foundation of China under grant #61425008 and #61333004, and Aeronautical Foundation of China under grant #2015ZA51013.

Yimin Deng is with Bio-inspired Autonomous Flight Systems (BAFS) Research Group, Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electrical Engineering, Beihang University, Beijing, China (e-mail: ymdeng@buaa.edu.cn).

Ning Xian is with Bio-inspired Autonomous Flight Systems (BAFS) Research Group, Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electrical Engineering, Beihang University, Beijing, China (e-mail: xianning@buaa.edu.cn).

Haibin Duan is with Bio-inspired Autonomous Flight Systems (BAFS) Research Group, Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electrical Engineering , Beihang University, Beijing, China (tel: +86-10-8231-7318; e-mail: hbduan@buaa.edu.cn).

Figure 1. Experimental validation platform of AAR

However, most of the works only concentrate on vision systems for UAVs, such as hardware construction, digital simulation environment and vision algorithms. Some works are not suitable for real-time outdoor applications of UAVs. Although the target detection and pose estimation has already been studied in a number of literatures, the implementation of vision-based measurement for AAR between UAVs remains a challenge in a real environment. This motivated us to establish the outdoor experimental platform to verify the binocular vision-based measuring system for AAR, as shown in Fig. 1.

In this paper, The implementation of a comprehensive real-time on-board vision system is presented. The on-board vision system includes a binocular hardware system and real-time vision algorithms. In this platform, a rotorcraft form XAIRCRAFT Technology, X650 Pro, is used as the receiver UAV. Several red circles with the equal diameter of 2cm are pasted on the top to represent those markers around the receptacle for marker detection. Another rotorcraft, Spreading Wings S1000 from DJI technology Inc, is used as the tanker UAV, which is equipped with the binocular vision–based measuring system. Real-time vision algorithms, such as image processing, feature extraction, pose estimation and communicating, are developed and running on the on-board vision computer. The configuration of the overall platform is presented in Fig. 2. The overall platform includes these following parts: a Tanker UAV platform, a receiver UAV platform, a binocular visual sensor, a vision processing module, a ground station, and video and data links.

This paper is structured as follows: Sections II and III introduce the development of binocular vision system and hardware configuration of the platform, respectively. The experimental results are shown in Section IV. The concluding remarks are given in the final section.

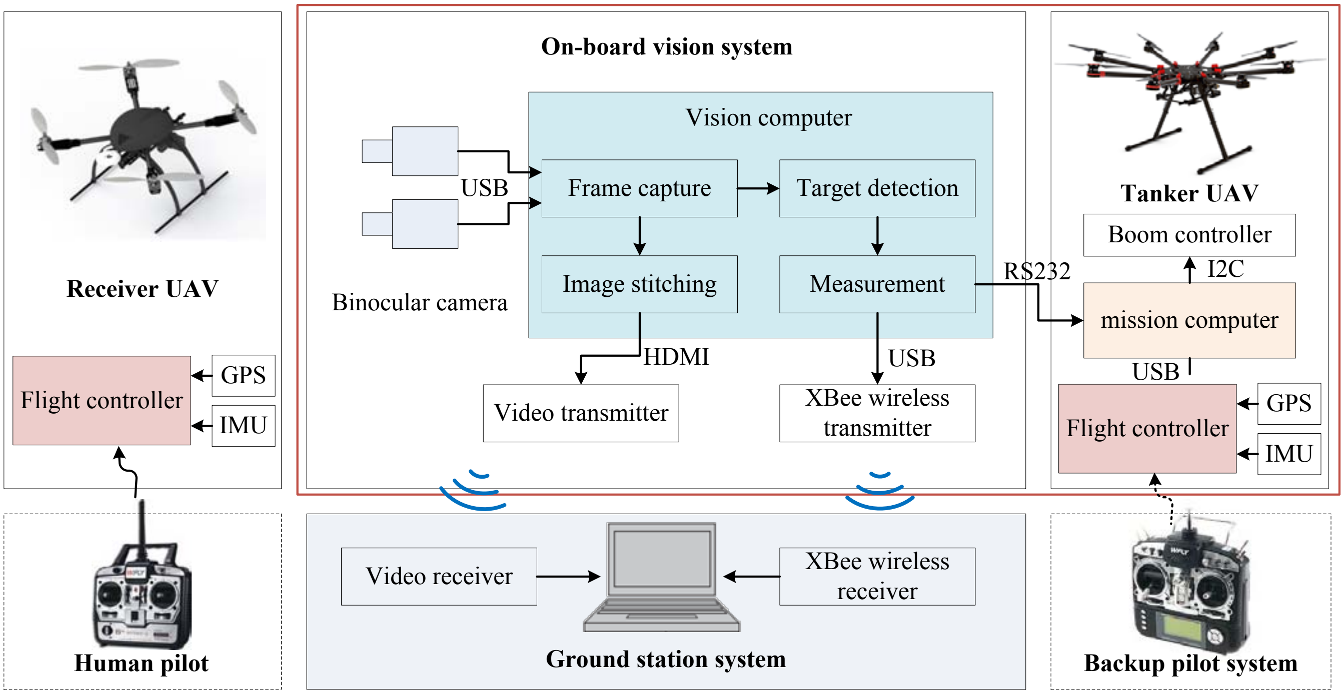


Figure 2. Configuration of the overall platform

II. Binocular Vision System

The process of calculating the relation is divided into two major steps, feature extraction and pose estimation. The goal of feature extraction is to generate the coordinates measured in pixels of labeled markers. Then the relation (rotational and translational) between the binocular camera system and the marker coordinate system is calculated using the pose estimation algorithm.

A. Feature Extration

As the initial relative distance between two UAVs is random in real applications, different strategies are required for cases at different distances. In general, the markers on the receiver UAV are hardly detected when the task begins at a far distance. To continuously estimate the position of the receiver UAV, two strategies of feature extraction are adopted for the long distance and short distance.

(1) **Saliency detection**. At a long distance, the saliency detection method is utilized to extract the regions of receiver UAV. In general, the selective region is unique within the image and stands out as the salient object. This process is able to guarantee the continuous localization when the markers are too small to be detected. When the receiver UAV approaches, saliency detection can also remove potential disturbance in images, thus it can improve the accuracy of marker detection.

Numerous works implement selective visual attention from a bottom-up perspective [14]. To avoid the large computational time in real applications, a global saliency measure based on the spectral residual [15] is adopted to favor regions with an unique appearance within the entire image. This saliency detection method is to detect the redundant part of image's log spectrum. Amplitude and phase are computed by log spectrum, and spectral residual is obtained by subtracting the average log spectrum. Based on the saliency analysis, the spectral residual corresponds to the salient information, while the average log spectrum represents the general shape. The saliency map S of the input image I is calculated as

$$S = g * \mathcal{F}^{-1}[\exp(\mathcal{R}(I) + \mathcal{P}(I))]^2 \qquad (1)$$

where $\mathcal{F}$ is the Fourier Transform, $\mathcal{R}(I)$ and $\mathcal{P}(I)$ are the spectral residual and the phase spectrum, and g is a Gaussian filter. Salient regions are segmented by applying a defined threshold. Thereofore, the centroid of the most salient region is estimated as the position of receiver UAV.

(2) **Marker detection**. When the receiver UAV approaches to the distance where the markers are capable to be seen clearly, the marker detection process is implemented. In the project, color features of those markers on receiver UAV are the main characteristics for feature extraction. Image frames captured from the binocular camera system are mapped from the RGB space to the HSV space. Intuitively, the HSV space is better in describing color features as it can effectively reduce the impact from light intensity. The hue and saturation channels are selected for the threshold segmentation to achieve binary images. The undesired noise in binary images can be excluded with morphology methods such as the erosion and dilation operators. After the morphological operators, the quantity of connected regions is counted as the number of feature points. The centroid of each connected region is also calculated as the image coordinate of each feature point.

B. Pose Estimation

To estimate the relative position and attitude of the relative UAV, extracted feature points are matched to actual markers to obtain the index number of each feature point. The Munkres algorithm [16] is employed to calculate the Euclidean distance matrix of two point sets transformed to the same image coordinates. Thus the problem of point matching is converted to a classical assignment task.

After indexing these feature points, binocular pose estimation algorithm based on the LHM algorithm [17] is presented to obtain the relative position. The binocular LHM algorithm extends the original one by minimize the object-space collinearity errors of two cameras. We define the left camera frame as the unified coordinate frame of the binocular camera system. The rotation and translation from the right camera to the left camera are denoted by $[R^{rl} \quad t^{rl}]_{3\times4}$, which can be acquired from the camera calibration procedure. Given several points $P_i (i=1,...n)$ in the 3D reference frame and the normalized coordinate $p_i^c (i \in \{1,...,n\}, c \in \{l,r\})$ of labeled feature point i of camera c, the binocular pose estimation problem is converted into a problem of minimizing the quadratic sum of the linear errors in object space:

$$\begin{aligned}[R,t] &= \arg\min \sum_{c=1}^{2}\sum_{i=1}^{n}\left\|e_i^c\right\|^2 \\ &= \arg\min \sum_{c=1}^{2}\sum_{i=1}^{n}\left\|(I-V_i^c)(RP_i+t-t^c)\right\|^2\end{aligned} \tag{2}$$

where $V_i^c = \dfrac{R^c p_i^c (R^c p_i^c)^T}{(R^c p_i^c)^T R^c p_i^c}$, $[R^c \quad t^c] = \begin{cases}[I \quad 0_{3\times1}], c=l \\ [R^{rl} \quad t^{rl}], c=r\end{cases}$, and $p_i^c = [u_i^c \quad v_i^c \quad 1]^T$.

The solution can be obtained by the use of the same iteration method as what is implemented in the original LHM algorithm. In general, the binocular LHM algorithm can converge within a small iteration beginning with any range of initial conditions.

C. Software Program

As shown in Fig. 3, the entire software program is divided into two main threads through multi-thread processing. Measuring algorithms are implemented in Thread 1, such as detecting features and estimating the relative distance between UAVs. Two detection methods are switched according to the distance threshold thr. In this platform, we set $thr = 3m$ to ensure the stability of maker detection algorithm. Other algorithms such as image stitching and data transmitting are processed in Thread 2, which is mainly for communicating with the flight controller and the ground station. To obtain an integral image for display in the ground station, the image stitching algorithm is utilized based on the homography matrix of the two cameras. Two threads share the data through the read-write lock.

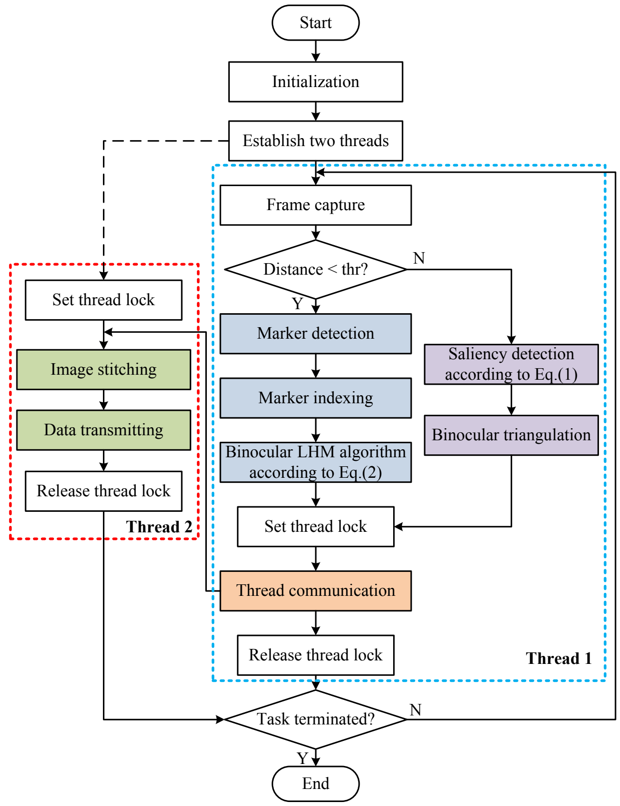


Figure 3. Flow chat of the on-board software program

III. HARDWARE CONFIGURATION

In Fig.4, the hardware configuration of the tanker UAV system consists of five main parts: a UAV platform, a visual sensor, a vision processing module, a ground station, and video and data links.

A. UAV Platform

The Spreading Wings S1000 from DJI technology Inc has been used as the tanker UAV. Its Maximum takeoff weight is about 11kg with its own weight of 4kg, thus this platform can easily carry the equipment. It has 15 minutes endurance equipped with a 6S 15000mAh battery. Its diameter is approximately 150cm and all eight arms can be completely folded down. Instead of its own controller, Pixhawk Autopilot from 3DRobotics Inc is adopted. The platform can be flown by a human pilot via an RC transmitter but it also accepts control signals from on-board computer.

B. Visual Sensor: Binocular Camera System

A visual sensor is utilized on board to obtain in-flight visual information. The binocular camera system consists of two same color video cameras. The selection is the Mercury camera from Daheng IMAVISION, which has a compact size equipped with Mini USB 2.0 interface and has a gross weight at about around 100g using a 12mm lens. It can provide a resolution of up to 1292×964 pixels and 40deg field of view. The frame rate is 30 frames per second (FPS), which is generally higher than that of vision algorithms (around 10 FPS).

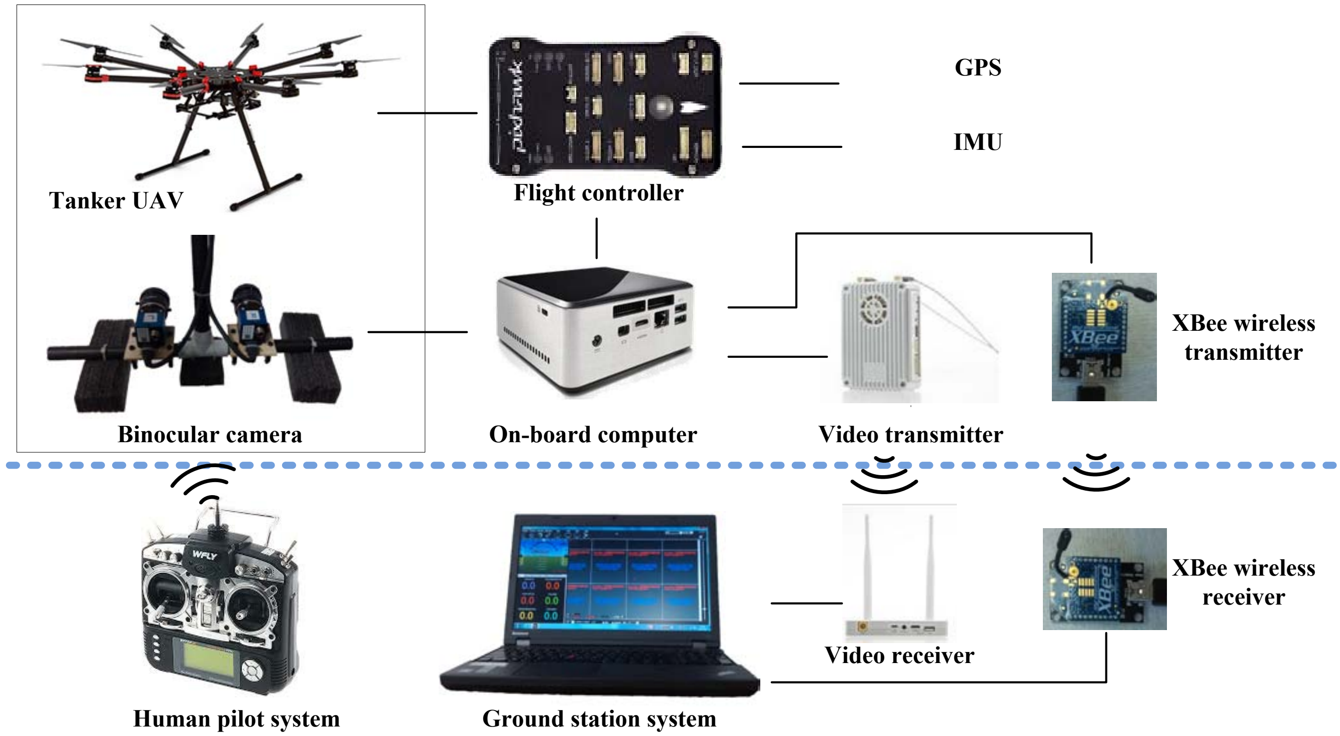


Figure 4. The tanker UAV system subcomponents and interconnections between them

Note that in this platform, other sensors such as GPS and IMU are also integrated. However, these devices provide worse accuracy than the binocular camera system, especially for close autonomous aerial refueling of UAVs.

C. Vision Processing Module: On-board Computer

The digitalized visual signals provided by the binocular camera system are transferred to the on-board computer. A separated on-board mini PC, Intel NUC, is employed to process the digitalized video signal and execute the vision algorithms. The core of the on-board PC is a 4th generation Intel Core i5 processor running at 1.3GHz. Equipped with a compact solid state disk, it weights around 650g. This vision computer coordinates the overall vision system, such as image processing, feature extraction, pose estimation and communicating with the flight controller and the ground station.

D. Wireless Data Link and Video Link

The video captured by the on-board camera is transmitted and displayed in a ground station using the DJI Lightbridge 2.4G HD digital video downlink. It can offer 1080p video data transmission from up to 1.7km away, while the gross weight of its air system is only 70g. The air system is connected to the on-board computer by HDMI. The ground system can be connected to a monitor, or transmit and display the video in the ground station through a video capture board. To transmit the position and pose data, a pair of XBee Modules from Digi International are employed based on UDP protocol in the platform.

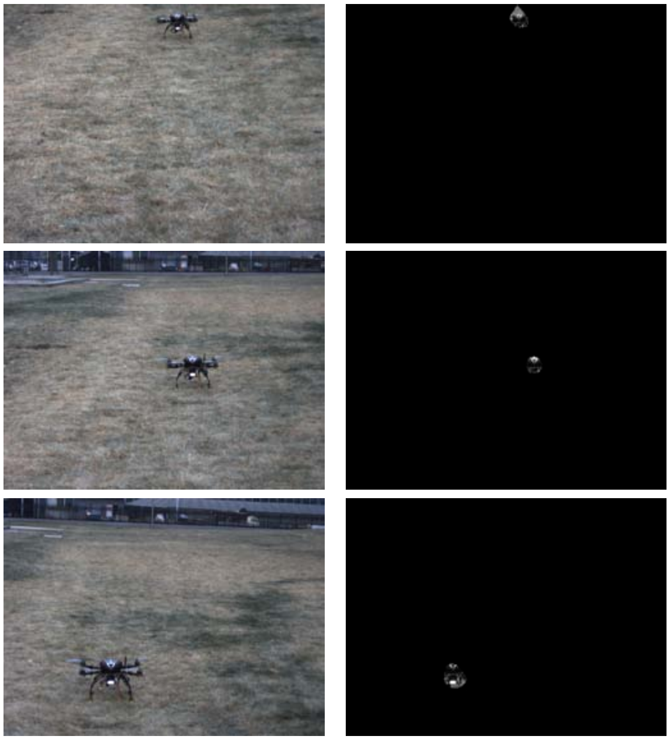
E. Ground Station

In the platform, the ground station displays the video captured by the on-board camera system and the position and pose data. It can provide ground operators with clear visualization during flight tests, and help operators to grasp the real-time information of the UAVs.

IV. Experimental Results

Series of experiments are conducted to verify the autonomous aerial refueling platform based on the binocular vision measuring system. During these tests, the receiver UAV with the markers is initialized at a random position. If the moving receiver UAV enters into the view of the vision system, the UAV or markers would be identified in the video sequence by the binocular vision measuring system automatically. Then the estimated position, pose data and stitched image are transmitted to the ground station for display.

The experimental results of saliency detection are shown in Fig.5. These results indicate that our algorithms could effectively detect and identify the target in the video sequence in the presence of the disturbance. The relative distance between two UAVs varies from around 11 meters to 5 meters. After calculating the centroid of saliency region of each camera, the relative distance is estimated using the triangulation method. Test result is shown in Fig.6. As there are some measuring errors and disturbance in Fig.6, the saliency based estimation method shall apply to the cases in a far distance.



(a) (b)

Figure 5. Results of saliency detection. (a) Original images of the receiver UAV from the distance to the near. (b) Saliency maps

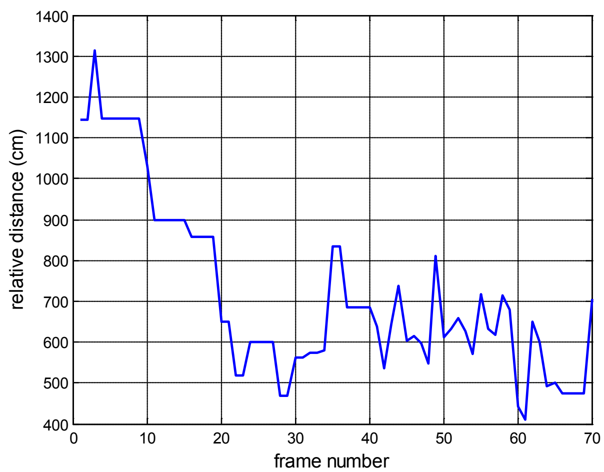
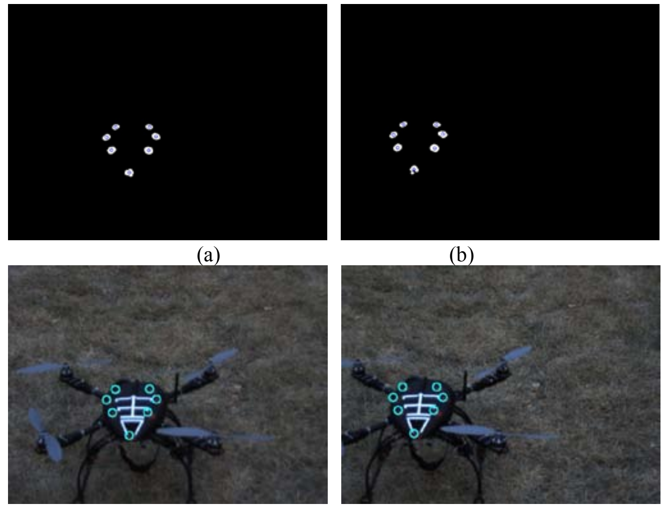


Figure 6. Test result of the relative distance estimation based on saliency detection

When the receiver UAV approaches to the place where those markers can be seen clearly, the marker detection algorithm and the binocular LHM algorithm are implemented. Estimation results are presented through the projective transformation. All markers are projected to the image frame based on the estimated pose information. The pose information is estimated accurately if all projective circles coincide with the red markers in the video sequence. From the experimental results shown in Fig.7 and Fig.8 we can see that, our platform could effectively detect all markers and provide accurate pose information in the presence of the disturbance.



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

Figure 7. Results of marker detection. (a)-(b) Segmentation results of left camera and right camera, respectively. (c)–(d) Corresponding projective results based on the estimated pose information

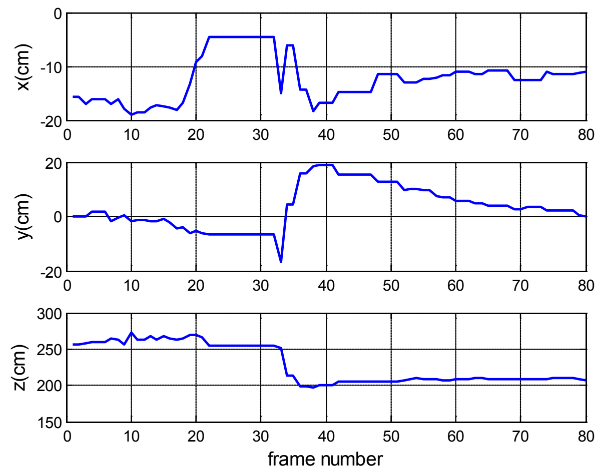


Fig. 8. Test result of the relative distance estimation based on the binocular LHM algorithm.

Note that the relative distance estimation varies as the drift of the hovering rotorcraft exists. Projective results based on the estimated pose information demonstrate the feasibility and effectiveness of the platform.

V. Conclusion

This paper presented the comprehensive implementation of the binocular vision-based measuring system for autonomous aerial refueling of UAVs, including hardware verification platform and vision algorithms. Series of tests were conducted to verify the platform. The experimental results indicate that the platform is not only capable to automatically detect the target but also capable to estimate the relative pose information in the video sequence. The efficiency of the developed platform for AAR could be achieved. Future research will focus on improving the measurement accuracy of the vision-based measuring system and implementation on vision-based automatic docking on a moving platform in a real environment.

REFERENCES

[1] H. B. Duan and L. Gan, "Elitist chemical reaction optimization for contour-based target recognition in aerial images," *IEEE Transactions on Geoscience and Remote Sensing*, vol.53, no.5, pp.2845-2859, 2015.

[2] F. Lin , X. X. Dong , B. M. Chen and K. Y. Lum, "A robust real time embedded vision system on an unmanned rotorcraft for ground target following," *IEEE Transactions on Industrial Electronics*, vol.59, no.2, pp.1038-1049, 2012.

[3] H. B. Duan, Q. N. Luo, G. J. Ma, and Y. H. Shi, "Hybrid particle swarm optimization and genetic algorithm for multi-UAVs formation reconfiguration," *IEEE Computational Intelligence Magazine*, vol.8, no.3, pp.16-27, 2013.

[4] F. Lin, K. M. Peng, X. X. Dong, S. Y. Zhao, and B. M. Chen, "Vision-based formation for UAVs," *IEEE International Conference on Control and Automation*, Taichung, Taiwan, 2014, pp.1375-1380.

[5] S. Hrabar, G. S. Sukhatme, P. Corke, K. Usher, and J. Roberts, "Combined optic-flow and stereo-based navigation of urban canyons for a UAV," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Albert, Canada, 2005, pp.3309-3316.

[6] H. B. Duan, X. X. Wei, and Z. N. Dong, "Multiple UCAVs cooperative air combat simulation platform based on PSO, ACO, and game theory," *IEEE Aerospace and Electronic Systems Magazine*, vol.28, no.11, pp.12-19, 2013.

[7] P. R. Thomas, U. Bhandari, S. Bullock, T. S. Richardson, and J. L. Du Bois, "Advances in air to air refuelling," *Progress in Aerospace Sciences*, vol.71, pp.14-35, 2014.

[8] L. Pollini, M. Innocenti, and R. Mati, "Vision algorithms for formation flight and aerial refueling with optimal marker labeling," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, San Francisco, California, AIAA 2005-6010, 2005.

[9] J. Valasek, K. Gunnam, J. Kimmett, M. D. Tandale, and J. L. Junkins, "Vision based sensor and navigation system for autonomous aerial refueling," *Journal of Guidance, Control, and Dynamics*, vol.28, no.5, pp.979-989, 2005.

[10] J. Doebbler, T. Spaeth, and J. Valasek, "Boom and receptacle autonomous air refueling using visual snake optical sensor," *Journal of Guidance, Control, and Dynamics*, vol.30, no.6, pp.1753-1769, 2007.

[11] H. B. Duan and Q. F. Zhang, "Visual measurement in simulation environment for vision-based UAV autonomous aerial refueling," *IEEE Transactions on Instrumentation and Measurement*, vol.64, no.9, pp.2468-2480, 2015.

[12] H. B. Duan, Q. F. Zhang, Y. M. Deng, and X. Y. Zhang, "Biologically eagle-eye-based autonomous aerial refueling for unmanned aerial vehicles," *Chinese Journal of Scientific Instrument*, vol.35, no.7, pp.1450-1458, 2014.

[13] R. V. Dell'Aquilla, G. Campa, M. R. Napolitano, and M. Mammarella, "Real-time machine-vision-based position sensing system for UAV aerial refueling," *Journal of Real-Time Image Processing*, vol.1, no.3, pp.213-224, 2007.

[14] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.35, no.1, pp.185-207, 2013.

[15] X. D. Hou and L. Q. Zhang, "Saliency detection: a spectral residual approach," *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, 2007, pp.1-8.

[16] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol.5, no.1, pp.32-38, 1957.

[17] C. P. Lu, G. D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, no.6, pp.610-622, 2000.