Binocular Pose Estimation for UAV Autonomous Aerial Refueling via Brain Storm Optimization

Cong Zhang
School of Automation Science and Electrical Engineering
Beihang University(BUAA)
Beijing, China
970846714@qq.com

Xiaobin Xu
School of Automation Science and Electrical Engineering
Beihang University(BUAA)
Beijing, China
xiaobinxu@buaa.edu.cn

Yuhui Shi, IEEE Fellow
Department of Computer Science and Engineering
Southern University of Science and Technology
Shenzhen, China
shiyh@sustc.edu.cn

Yimin Deng
School of Automation Science and Electrical Engineering
Beihang University(BUAA)
Beijing, China
ymdeng@buaa.edu.cn

Cong Li
School of Automation Science and Electrical Engineering
Beihang University(BUAA)
Beijing, China
447020047@qq.com

Haibin Duan, *Senior Member, IEEE*
1. *Beihang University(BUAA)*
Beijing, China
2. *Pengcheng Laboratory*
Shenzhen, China
hbduan@buaa.edu.cn

***Abstract*—Autonomous aerial refueling (AAR) is a crucial technique of unmanned aerial vehicles (UAVs) to push the fuel limits and play a great role in both civilian and military domains. This paper presents an accurate and robust binocular pose estimation algorithms optimized by brain storm optimization (BSO), which is developed from a robust non-iterative solution of PnP (RPnP). In this algorithm, BSO is employed to select the best rotation axis in RPnP. A large quantity of contrastive simulation experiments has been conducted to verify the proposed algorithm. Furthermore, this work built an aerial verification platform for vision-based AAR. A tanker UAV and a receiver UAV were applied to implement AAR. The real-time visual measuring system includes feature extraction and pose estimation. Several state-of-the-art pose estimation algorithms and the proposed method (which refers to BSO-BPnP) have been tested in the aerial verification platform. Adequate comparative trials and detailed analyses are given in this paper.**

Keywords—autonomous aerial refueling (AAR), unmanned aerial vehicles (UAVs), pose estimation, brain storm optimization (BSO), binocular camera system

I. INTRODUCTION

Unmanned Aerial Vehicle (UAV) is a type of aircraft that does not require human pilot on board [1]. UAV shows immense potential in both civilian and military domains. Benefiting from its higher cost performance, light size, and maneuvering capabilities, UAV gradually replaces manned aircraft to execute various ordinary and sophisticated missions, such as border patrol [2], surveillance [3], forest firefighting [4, 5], and target detection and traction [6, 7], etc. The limited fuel has been a critical bottleneck for further development[8]. Autonomous aerial refueling (AAR) is a powerful way to solve this problem [9]. Common aerial refueling can be clasped into boom and probe-and-drogue [10]. However, a human operator is quite a challenging task.

A key technology of AAR is gaining the accurate pose information between receiver and tanker. Currently, inertial navigation system (INS) and global positioning system (GPS) are widely used for aircraft navigation[11]. However, INS is not suitable because the accumulation of errors. GPS-based measurements are ill-suited for docking maneuver in autonomous aerial refueling accurately. 3D flash lidar sensor is utilized to measure the relative position and attitude between aircraft for drogue tracking [12]. Boom approach using visual snake optical sensor was proposed in [13]. A real-time visual measurement in simulation environment is designed for semi-autonomous docking within aerial refueling [14]. Computer vision-based navigation could refrain from error accumulation and meet precision requirement. Therefore, visual measuring system is applied for pose estimation for AAR. BPnP is proposed based on the RPnP (a robust non-iterative solution of PnP) [22]. BSO is an innovative Swarm Intelligence (SI) algorithm proposed by Shi in 2009 [16], which is motivated by the brainstorming of human beings. Recently years, BSO has been proven feasible and effective in dealing with various optimization issues [17-19]. In this paper, BSO is implemented to find the optimal rotation axis, and then yield the pose of the binocular camera system.

This paper is organized as follows: Section II introduces the feature extraction of the receptacle. Section III presents the BPnP pose estimation method. BSO- BPnP is given in Section IV. Section V introduces the experimental platform configuration. The simulation results and outfield aerial tests are shown in Section VI.

II. FEATURE EXTRACTION OF THE RECEPTACLE

Feature extraction is a significant step in visual measurement. Firstly, the labeled markers are recognized from image sequences, and the pixel coordinates needs to be

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

resolved. Then, the pose estimation algorithm is applied to figure out the relation (rotational and translational) between camera coordinate system and receptacle coordinate system.

A. Marker Detection

Feature of color is a considerable feature to implement image segmentation. The Hue-Saturation-Intensity (HSV) space can weaken the influences of illumination intensity[14]. Therefore, Red-Green-Blue (RGB) images are mapped to HSV space. The conversion formulas are shown

$$h=\begin{cases} 0^\circ & \max=\min \\ 60^\circ\times\dfrac{g-b}{\max-\min}+0^\circ & \max=r, g\ge b \\ 60^\circ\times\dfrac{g-b}{\max-\min}+360^\circ & \max=r, g<b \\ 60^\circ\times\dfrac{b-r}{\max-\min}+120^\circ & \max=g \\ 60^\circ\times\dfrac{r-g}{\max-\min}+240^\circ & \max=b \end{cases} \tag{1}$$

$$s=\begin{cases} 0 & \max=0 \\ \dfrac{\max-\min}{\max} & \max\ne 0 \end{cases} \tag{2}$$

$$v=\max \tag{3}$$

where h, s, v are the hue channel, saturation channel and intensity channel, the red channel, green channel and blue channel are indicated by r, g, b, max and min are the maximum and minimum of color space respectively. h and s are selected for the threshold segmentation. The obtained binary image with noisy point will execute morphological operation. Every connected area of adequate size counts as a feature point. The focuses of feature points are calculated as the image coordinate.

B. Point Matching

Point matching is an important procedure for pose estimation. The mapping relation between image coordinates and world coordinates of actual markers around the receptacle in the receiver UAV need be determined. Assume that detected feature point set is $P=\{p_1,p_2,\cdots,p_n,\}$. The measured corresponding image coordinate of marker j is $p_j=(u_j,v_j)$. The projective point set $\hat{P}=\{\hat{p}_1,\hat{p}_2,\cdots,\hat{p}_n,\}$ is obtained from the priori pose information. $\hat{p}_j=\{\hat{u}_j,\hat{v}_j\}$ is the calculated image coordinate according perspective projection model.

The similarity degree of two points set is defined as a cost function. All the Euclidean distances between two points set reaches the minimum when the two points set matches correctly. A matrix Err is shown as

$$Err=\begin{bmatrix} d(\hat{p}_1,p_1) & d(\hat{p}_1,p_2) & \cdots & d(\hat{p}_1,p_n) \\ d(\hat{p}_2,p_1) & d(\hat{p}_2,p_2) & \cdots & d(\hat{p}_2,p_n) \\ \vdots & \vdots & \vdots & \vdots \\ d(\hat{p}_m,p_1) & d(\hat{p}_m,p_2) & \cdots & d(\hat{p}_m,p_n) \end{bmatrix} \tag{4}$$

III. The BPnP Algorithm

The pose between camera coordinate and world coordinate from n correspondences is a "Perspective-n-Point problem" (PnP) [21]. BPnP originates from RPnP, which was proposed with the lowest computational complexity $\Theta(n)$ by Li et.al [22]. In this section, BPnP will be introduced.

A. Problem Description

The rotation and translation from the left camera to the right camera are denoted by $[R^{rl}\quad t^{rl}]_{3\times4}$. $P_i\left(i=1,...,n,n\ge3\right)$ are points in a 3D reference frame. their corresponding projections on normalized left image plane $p_i^l\left(i=1,...,n\right)$ and right image plane $p_i^r\left(i=1,...,n\right)$ are shown in Fig. 1.

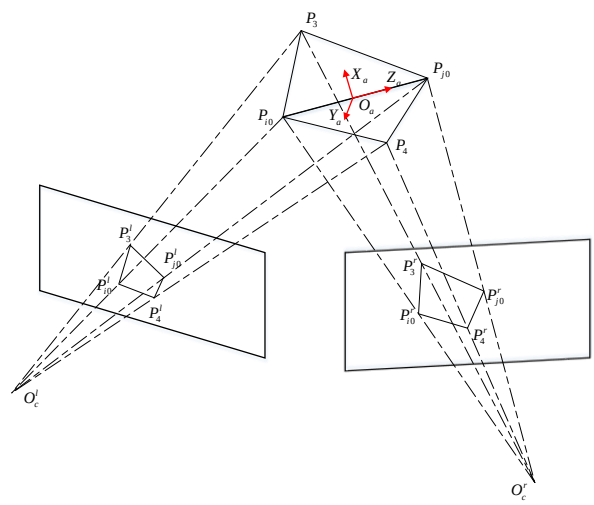


Fig. 1. The 3D reference points and their projections on the binocular camera image plane.

The task is to estimate the pose information between two cameras: $[R^l\quad t^l]_{3\times4}$ and $[R^r\quad t^r]_{3\times4}$. If the transformation relation among two camera frames and the 6DOF of left camera are known, the right camera frame can be got based on the principle of stereo vision as follows

$$\begin{bmatrix} R^r & t^r \\ 0^{\mathrm{T}} & 1 \end{bmatrix}=\begin{bmatrix} R^{rl} & t^{rl} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} R^l & t^l \\ 0^{\mathrm{T}} & 1 \end{bmatrix} \tag{5}$$

B. Selection of a Rotation Axis

BPnP begins with an edge $\overline{P_{i0}P_{j0}}$ selected from all the edges $\left\{\overline{P_iP_j}\mid i>j, i\in\{1,...,n\}, j\in\{1,...,n\}\right\}$ in the n-point set

Identify applicable funding agency here. If none, delete this text box.

as a rotation axis [15]. The edge with longest projection length $\|p_i p_j\|$ is less affected by noise. Then, BSO is employed to select the best rotation axis.

A new coordinate frame $O_a X_a Y_a Z_a$ related to selected edge $\overline{P_{i0}P_{j0}}$ is established. The original point O_a is in the middle of $\overline{P_{i0}P_{j0}}$, and Z_a axis points to $\overrightarrow{P_{i0}P_{j0}}$. Then, the n reference points are divided into $(n-2)$ subsets, which contain three points $\{P_{i0}P_{j0}P_k \mid k \neq i0, k \neq j0\}$. According to 3 point constraint, $(n-2)$ polynomials are got as follows

$$\begin{cases} f_1(x) = \alpha_1 x^4 + \beta_1 x^3 + \chi_1 x^2 + \delta_1 x + e_1 = 0, \\ f_2(x) = \alpha_2 x^4 + \beta_2 x^3 + \chi_2 x^2 + \delta_2 x + e_2 = 0, \\ \ldots, \\ f_{n-2}(x) = \alpha_{n-2} x^4 + \beta_{n-2} x^3 + \chi_{n-2} x^2 + \delta_{n-2} x + e_{n-2} = 0 \end{cases} \tag{6}$$

After finding the minima of the square sum of the polynomials, the rotation axis is determined by

$$Z_a = \frac{\overrightarrow{P_{i0}P_{j0}}}{\|P_{i0}P_{j0}\|} \tag{7}$$

C. Derivation of Rotation and Translation

When the Z_a axis is determined, the rotation matrix from $O_a X_a Y_a Z_a$ to the camera $O_c X_c Y_c Z_c$ can be denoted as

$$R^l = Rrot(Z_a, \alpha) = \begin{bmatrix} r_1 & r_4 & r_7 \\ r_2 & r_5 & r_8 \\ r_3 & r_6 & r_9 \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

where R is an orthogonal rotation matrix, and it has the same elements as Z_a in its third column, $rot(Z_a, \alpha)$ denotes a rotation of α degree around Z_a axis. The transformation relation between the 2D projections of left and right image plane is obtained by

$$\begin{cases} z_i^l p_i^l = R^l P_i + t^l \\ z_i^r p_i^r = R^{rl} R^l P_i + R^{rl} t^l + t^{rl} \end{cases} \tag{9}$$

where z_i^l and z_i^r are the projection depth of ith reference point whose 3D coordinates are $P_i = [X_i \quad Y_i \quad Z_i]^{\mathrm{T}}$ in $O_a X_a Y_a Z_a$, $p_i^l = (u_i, v_i, 1)^T$ and $p_i^r = (u_i^r, v_i^r, 1)^T$ are the normalized coordinates in the left and right image plane, respectively. Rewrite the (9) to a $4n \times 6$ homogenous linear equation system shown by (10).

If the pose of the left camera by 3D alignment scheme [24] is got, the pose of the right can be obtained based on (5).

$$[A_{4n\times1} \; B_{4n\times1} \; C_{4n\times4}][\cos\alpha \quad \sin\alpha \quad t_x^l \quad t_y^l \quad t_z^l \quad 1]^T = 0 \tag{10}$$

where $[\cos\alpha \quad \sin\alpha \quad t_x^l \quad t_y^l \quad t_z^l \quad 1]^T$ is obtained by Singular Value Decomposition (SVD) [23], $C_{4n\times4} = [C_1 \quad C_2 \quad C_3 \quad C_4]$,

$$A_{4n\times1} = \begin{bmatrix} u_1^l X_1 r_3 - Y_1 r_4 - X_1 r_1 + u_1^l Y_1 r_6 \\ v_1^l X_1 r_3 - Y_1 r_5 - X_1 r_2 + v_1^l Y_1 r_6 \\ u_1^r X_1 d_3 - Y_1 d_4 - X_1 d_1 + u_1^r Y_1 d_6 \\ v_1^r X_1 d_3 - Y_1 d_5 - X_1 d_2 + v_1^r Y_1 d_6 \\ \ldots \\ u_n^l X_n r_3 - Y_n r_4 - X_n r_1 + u_n^l Y_n r_6 \\ v_n^l X_n r_3 - Y_n r_5 - X_n r_2 + v_n^l Y_n r_6 \\ u_n^r X_n d_3 - Y_n d_4 - X_n d_1 + u_n^r Y_n d_6 \\ v_n^r X_n d_3 - Y_n d_5 - X_n d_2 + v_n^r Y_n d_6 \end{bmatrix}, C_1 = \begin{bmatrix} -1 \\ 0 \\ u_1^r r_3^{rl} - r_1^{rl} \\ v_1^r r_3^{rl} - r_2^{rl} \\ \ldots \\ -1 \\ 0 \\ u_n^r r_3^{rl} - r_1^{rl} \\ v_n^r r_3^{rl} - r_2^{rl} \end{bmatrix},$$

$$B_{4n\times1} = \begin{bmatrix} Y_1 r_1 + u_1^l X_1 r_6 - u_1^l Y_1 r_3 - X_1 r_4 \\ Y_1 r_2 + v_1^l X_1 r_6 - v_1^l Y_1 r_3 - X_1 r_5 \\ Y_1 d_1 + u_1^r X_1 d_6 - u_1^r Y_1 d_3 - X_1 d_4 \\ Y_1 d_2 + v_1^r X_1 d_6 - v_1^r Y_1 d_3 - X_1 d_5 \\ \ldots \\ Y_n r_1 + u_n^l X_n r_6 - u_n^l Y_n r_3 - X_n r_4 \\ Y_n r_2 + v_n^l X_n r_6 - v_n^l Y_n r_3 - X_n r_5 \\ Y_n d_1 + u_n^r X_n d_6 - u_n^r Y_n d_3 - X_n d_4 \\ Y_n d_2 + v_n^r X_n d_6 - v_n^r Y_n d_3 - X_n d_5 \end{bmatrix}, C_2 = \begin{bmatrix} 0 \\ -1 \\ u_1^r r_6^{rl} - r_4^{rl} \\ v_1^r r_6^{rl} - r_5^{rl} \\ \ldots \\ 0 \\ -1 \\ u_n^r r_6^{rl} - r_4^{rl} \\ v_n^r r_6^{rl} - r_5^{rl} \end{bmatrix},$$

$$C_3 = \begin{bmatrix} u_1^l \\ v_1^l \\ u_1^r r_9^{rl} - r_7^{rl} \\ v_1^r r_9^{rl} - r_8^{rl} \\ \ldots \\ u_n^l \\ v_n^l \\ u_n^r r_9^{rl} - r_7^{rl} \\ v_n^r r_9^{rl} - r_8^{rl} \end{bmatrix}, C_4 = \begin{bmatrix} u_1^l r_9 Z_1 - r_7 Z_1 \\ v_1^l r_9 Z_1 - r_8 Z_1 \\ u_1^r d_9 Z_1 - d_7 Z_1 + u_1^r t_3^{rl} - t_1^{rl} \\ v_1^r d_9 Z_1 - d_8 Z_1 + v_1^r t_3^{rl} - t_2^{rl} \\ \ldots \\ u_n^l r_9 Z_n - r_7 Z_n \\ v_n^l r_9 Z_n - r_8 Z_n \\ u_n^r d_9 Z_n - d_7 Z_n + u_n^r t_3^{rl} - t_1^{rl} \\ v_n^r d_9 Z_n - d_8 Z_n + v_n^r t_3^{rl} - t_2^{rl} \end{bmatrix},$$

$$D = R^{rl} R = \begin{bmatrix} r_1^{rl} & r_4^{rl} & r_7^{rl} \\ r_2^{rl} & r_5^{rl} & r_8^{rl} \\ r_3^{rl} & r_6^{rl} & r_9^{rl} \end{bmatrix} \begin{bmatrix} r_1 & r_4 & r_7 \\ r_2 & r_5 & r_8 \\ r_3 & r_6 & r_9 \end{bmatrix} = \begin{bmatrix} d_1 & d_4 & d_7 \\ d_2 & d_5 & d_8 \\ d_3 & d_6 & d_9 \end{bmatrix}$$

IV. BSO Optimized BPnP

BSO algorithm is used to solve the rotation axis selection problem of BPnP. In this section, the detailed introduction and the structure diagram of BSO is presented.

The position of the ith individual is denoted by $X_i = [x_{i1}, x_{i2}, ..., x_{im}]$, where $m = 2$ is the dimension of each individual. $X_i = [x_{i1}, x_{i2}]$ represents the edge $\overline{P_{x_{i1}} P_{x_{i2}}}$, which is an essential variable. Firstly, n individuals are randomly generated within the solution space. Then each individual is evaluated by the fitness function which is the total error of BPnP. The optimization objective is obtaining the minimal total error of BPnP based on BSO.

A random selected cluster center is replaced by randomly generated individual with the possibility of p_{6a}. BSO selects one cluster of two with the possibility of p_{6b} randomly. Then select the center or other individuals of the chosen cluster(s) with the possibility of p_{6b3}. After that, the selected individual(s) can be updated based on (11) and (12).

$$x_{old} = \begin{cases} x_{ij}, & \text{one cluster} \\ w_1 x_{i1,j} + w_2 x_{i2,j} & \text{two cluster} \end{cases} \tag{11}$$

$$x_{new} = x_{old} + \xi N(\mu, \sigma) \tag{12}$$

where w_1 and w_2 are the weight of two clusters. $N(\mu, \sigma)$ is the Gaussian random value with mean μ and variance σ. ξ denotes the adjusting factor decreasing as the iteration goes.

$$\xi = r \log sig\left(\frac{Nc_{\max} - N_c}{K}\right) \tag{13}$$

where r is the random number within [0,1], $Nc_{\max}$ denotes the maximum number of iteration, N_c signifies the current iteration number and K adjusts the slope of the *logsig* function. After the new individual is generated, the crossover operation is conducted and the best individual is updated.

The process above repeats until all individuals are created and the iteration proceeds until the maximum number of iteration is met. Finally, the best individual of the algorithm is the optimal edge.

V. Using the Template

To validate the validity of the proposed method, the experimental platform configuration is prepared in this section.

Specifically, two hexrcopters are regarded as the receiver UAV and the tanker UAV in Fig.2. A visual measuring system is installed in the tanker. Real-time vision algorithms are developed and running on the on-board vision computer. The video sequence captured onboard is transmitted to a ground station by a digital image transmission. The data results are also displayed in the ground station. Therefore, we can implement real-time monitor from the view of the tanker. If an exception occurs, operator will take over the control power of UAVs. The data transmission is made up of two XBee modules based on UDP protocol.

Fig. 2. Experimental platform for autonomous boom type aerial refueling.

The schematic diagram of the aerial verification platform and the connection of each subsystem is presented in Fig. 3.

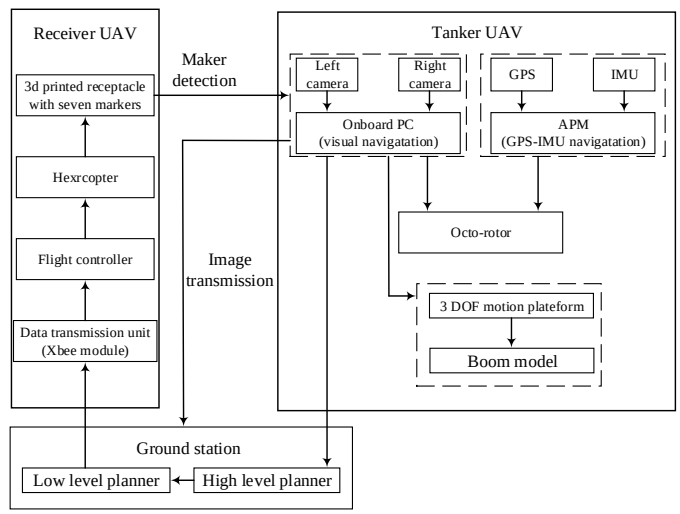


Fig. 3. System structure of the aerial verification platform.

VI. Experimental Results

In this section, two sets of experiments are designed. One is conducting simulation experiment to compare the robustness and accuracy of BSO-BPnP against that of some advanced pose estimation algorithms. The other type of experiment is executed to verify the designed visual measuring system for UAV autonomous aerial refueling.

A. Simulation Experiment of BSO-BPnP

The comparison algorithms are shown.

- LHM. A applicable iterative method proposed by Lu et al [25] in 2000.
- MLHM [26]. One of the best pose estimation methods improved from LHM by minimize the object-space collinearity errors of two cameras. MLHM can obtain the pose and position of a binocular camera system.
- EPnP. A efficient $O(n)$ noniterative solution of PnP by Lepetit [27]. The core concept is to express the n reference points as a weighted sum of four virtual control points.

- DLT [28]. A direct linear transformation method.
- HOMO. A homography method for planar targets [29].
- RP*n*P. A robust non-iterative PnP solution by Li [22].
- BP*n*P. A pose estimation method for binocular camera systems proposed by Gan et al [15], which can be used to estimate the poses of both cameras at the same time.
- BSO-BP*n*P. BSO algorithm is applied to select the best optimal rotation axis in BP*n*P. BSO-BP*n*P is suited to the pose estimation of a binocular camera system.

The experimental configuration is two virtual perspective cameras with 640×480 pixels size, 800 pixels focal length. The relation matrix between two cameras are given by

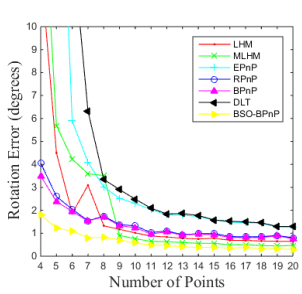
$$\begin{bmatrix} R^{rl} & t^{rl} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{13}$$

To prove the effectiveness of BSO-BP*n*P, three cases of the 3D points are generated randomly [17]. The reference points are in the range of $[-2,2]\times[-2,2]\times[4,8]$, $[-2,2]\times[-2,2]\times0$ and $[1,2]\times[1,2]\times[4,8]$ randomly. Besides, Gaussian noise with $\sigma=3$ is added to their projections on the two camera image planes. The parameters and their detail descriptions of BSO are expressed in Table I.

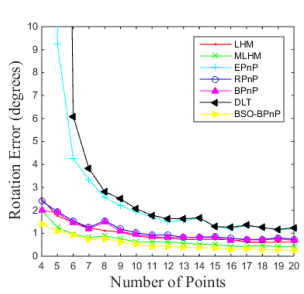
TABLE I. THE PARAMETERS OF BSO ALGORITHM

Parameter	Description	Value
m	Number of ideas	30
Nc_{max}	Maximum times of iteration	30
K	Number of clusters	3
P_{5a}	Probability to directly update a cluster center	0.2
P_{6b}	Probability to choose one cluster	0.8
P_{6b3}	Probability to select the center of the selected cluster	0.4
P_{6c}	Probability to select the centers of the two selected clusters	0.5

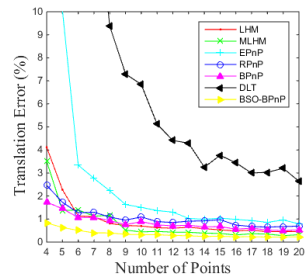
Each experiment is conducted for 100 times. The results of the three cases are recorded to make statistics, which are given from Fig.4 to Fig.6.



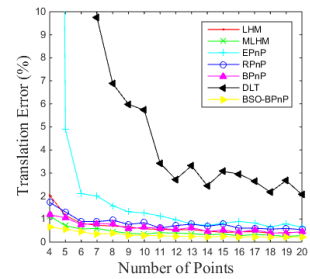
(a) Mean Rotation Error



(b)Median Rotation Error



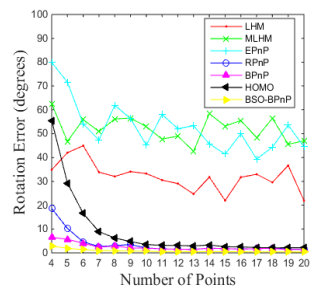
(c) Mean Translation Error



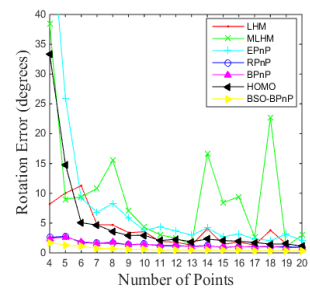
(d) Median Translation Error

Fig. 4. Experimental results in the ordinary 3D case. The number of points $n=4,\cdots,20$, and the level of Gaussian noise $\sigma=3$ pixels.

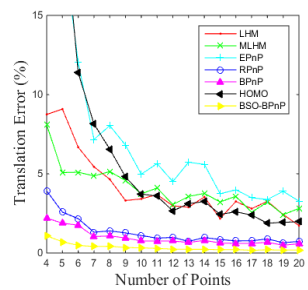
The mean and median rotation error, the mean and median translation error are presents respectively. Shown in Fig.4, in the ordinary 3D case, the relatively accurate results could be obtained when n was large enough. The best performance of the proposed BSO-BP*n*P emerge early among the 7 methods.



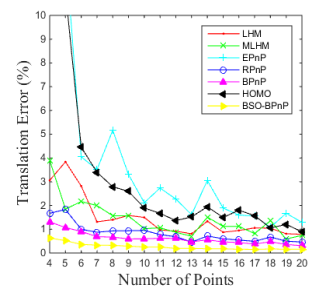
(a) Mean Rotation Error



(b)Median Rotation Error

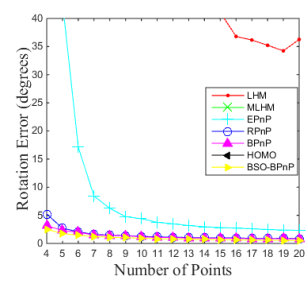


(c) Mean Translation Error

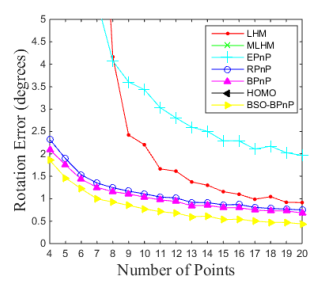


(d) Median Translation Error

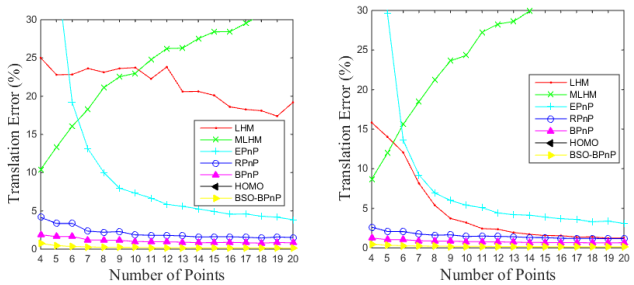
Fig. 5. Experimental results in the planar case. The number of points $n=4,\cdots,20$, and the level of Gaussian noise $\sigma=3$ pixels.



(a) Mean Rotation Error



(b)Median Rotation Error



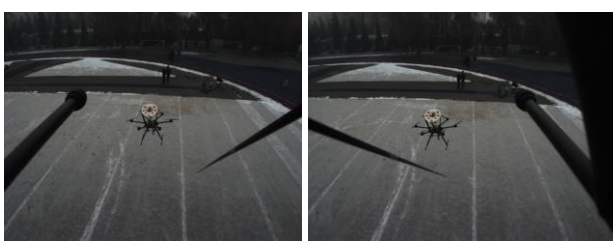
(c) Mean Translation Error (d) Median Translation Error

Fig. 6. Experimental results in the quasi-singular case. The number of points $n = 4, \cdots, 20$, and the level of Gaussian noise $\sigma = 3$ pixels.

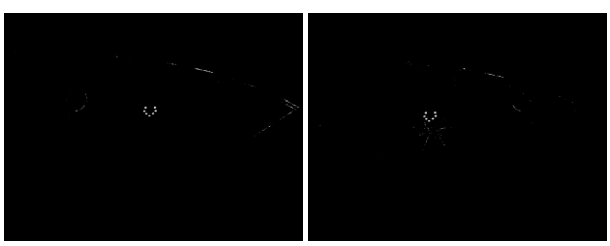
To coplanar PnP problem, Fig.5 presents the simulation results. LHM, MLHM and EPnP are not suitable for this scenario, while BSO-BPnP, BPnP and RPnP remain stable and highly accurate with the error mainly below five points. BSO-BPnP reaches the best results in the test obviously. As shown in Fig.6, EPnP, LHM and MLHM have bad performance in the quasi-singular case because of local minima problem. Besides, the error of HOMO was too large to make a curve in figure. BSO-BPnP obtains the most stable and accurate result in all of methods.

B. *Aerial Practical Experiment*

To verify the feasibility and effectiveness of the visual measuring system, outdoor aerial practical experiments are conducted. Two UAVs are placed at random positions. They take off automatically. The receiver navigated by GPS/INS unit flies to the scheduled position relative to the tanker. When the receiver comes into the view of the binocular camera system, the navigating mode is switched to the visual navigating model.



(a) Original images from left cameras (b) Original images from right cameras



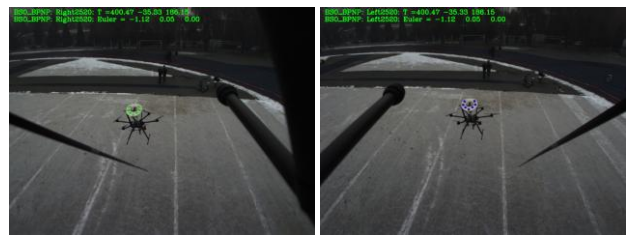
(c) - (d) Corresponding detection results based on HSV space segmentation

Fig. 7. Results of marker detection.

Fig.7 shows the experimental results of feature extraction. Seven red markers are detected successfully based on the color feature. The central coordinates of these markers are calculated. Fig. 8 gives the results of monocular vision algorithms with the same image. Fig. 9 illustrates the results of binocular vision methods for two corresponding images. The green circles are projected on the original images based on the measured pose. All projective circles coincide with the red markers in the image sequences. It is verify that the tested monocular and binocular vision algorithms are effective and feasible.

(a) RPnP (b) LHM (c) EPnP

Fig. 8. Results of monocular vision algorithms for a same image.



BSO-BPnP. (a) Left camera frame (b) Right camera frame

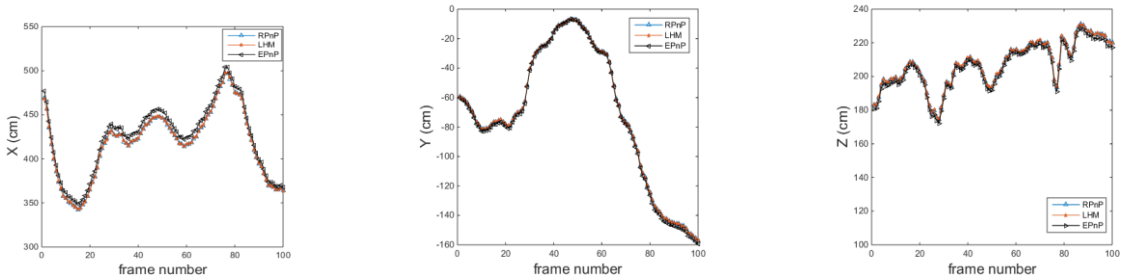
BLHM. (c) Left camera frame (d) Right camera frame

Absolute Orientation. (e) Left camera frame (f) Right camera frame

Fig. 9. Results of binocular vision algorithms for the two same images captured from left and right cameras simultaneously.

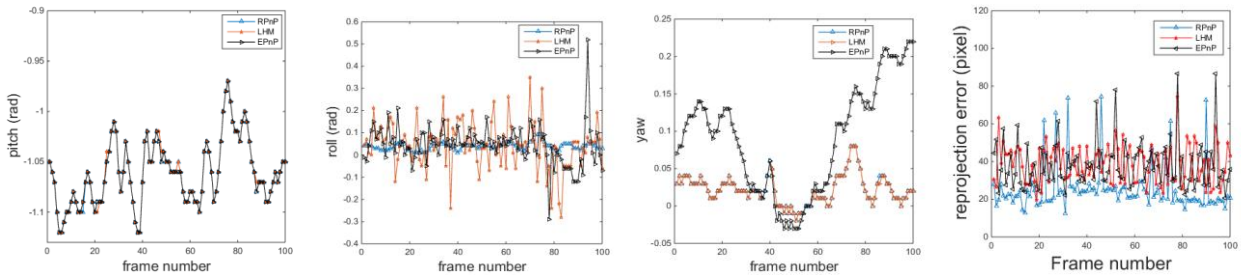
Fig. 10 shows the measured relative distance between two UAVs. RP*n*P, LHM and EP*n*P obtain similar results of relative distance estimation in three orthogonal directions. The relative pose is given in Fig.11 for monocular vision. The results are consistent in pitch angle estimation in Fig.11(a). RP*n*P has stable estimation in roll angle estimation, while LHM and EP*n*P show more fluctuation shown in Fig.11(b). As shown in Fig.11(c), RP*n*P and LHM acquire alike estimation results in yaw angle, and EP*n*P shows the trend of divergence. Since the actual value of pose cannot be obtained in outdoor, an alternative evaluation standard is proposed. The re-projected

center is close to the original markers if the measurements are exact. Fig.11(d) gives the re-projection error of monocular vision algorithms. RP*n*P shows the highest accuracy. LHM is the second. EP*n*P is most unstable. RP*n*P is the appropriate monocular vision algorithm for pose estimation among them.



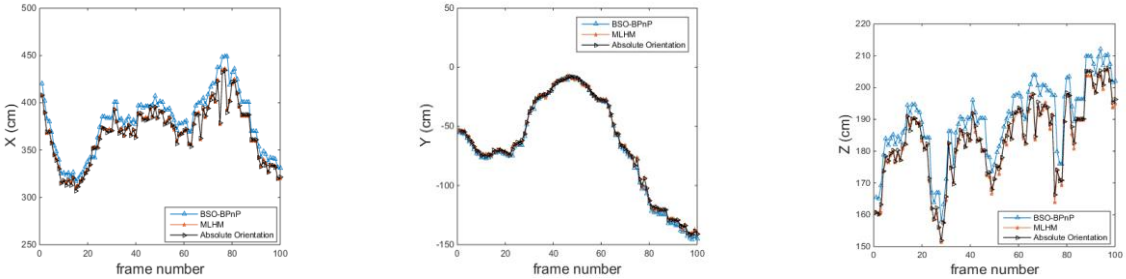
(a) Relative distance in X-axis (b) Relative distance in Y-axis (c) Relative distance in z-axis

Fig. 10. Experimental results of relative distance estimation based on monocular vision in three orthogonal direction between two UAVs during the flight.



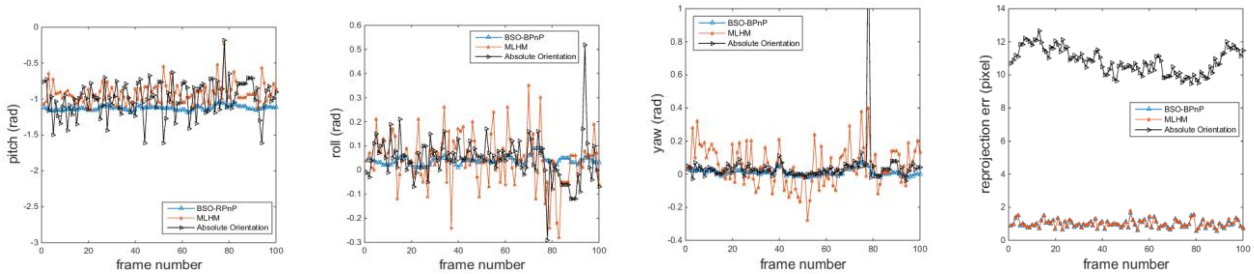
(a) Pitch angle estimation (b) Roll angle estimation (c) Yaw angle estimation (d)Comparision of the reprojection error of monocular vision algorithms

Fig. 11. Experimental results of relative distance estimation based on monocular vision in three orthogonal direction between two UAVs during the flight.



(a) Relative distance in X-axis (b) Relative distance in Y-axis (c) Relative distance in z-axis

Fig. 12. Experimental results of relative distance estimation based on monocular vision in three orthogonal direction between two UAVs during the flight.



(a) Pitch angle estimation (b) Roll angle estimation (c) Yaw angle estimation (d) Comparison of the reprojection error of monocular vision algorithms

Fig. 13. Experimental results of relative distance estimation based on monocular vision in three orthogonal direction between two UAVs during the flight.

The binocular vision algorithms BSO-BP*n*P, MLHM and absolute orientation are compared in Fig. 12 -Fig. 13. As shown in Fig.12, there are no significant differences in relative distance estimation among the three methods. However, BSO-BP*n*P is more stable than others in Fig. 13. Absolute orientation reaches outliers sometime. The comparison of the re-projective error of monocular vision algorithms is given in Fig.13(d). BSO-BP*n*P and MLHM have the re-projection error less than 2 pixels, which means these two methods could be available binocular algorithms in UAVs AAR.

VII. CONCLUTIONS

An accurate and robust binocular pose estimation algorithms optimized by BSO-BP*n*P is proposed. It is applied in vision-based UAVs autonomous aerial refueling. To verify the feasibility and effectiveness of BSO-BP*n*P, an aerial vision-based AAR platform established. Considerable simulation tests and outdoor flight tests are conducted. The proposed BSO-BP*n*P has shown great robustness and accuracy compared with some state-of-the-art pose estimation algorithms. BSO-BP*n*P is outclassed from other methods in three test cases in simulation environment. Besides, the flight tests indicates that the verification platform could detect the receptacle and implement relative distance and pose estimation with acceptable accuracy. LHM, MLHM, RP*n*P and the proposed BSO-BP*n*P show best performance so that they could be the alternative vision algorithms for UAVs ARR. In the future, we will focus on the improvement of the visual measurement in extreme conditions, such as target occlusion, strong exposure and lighting environment. Furthermore, we will improve the aerial verification platform and implement the docking phase successfully and steadily

REFERENCES

[1] S. P. Yeong, L. M. King and S. S. Dol, "A review on marine search and rescue operations using unmanned aerial vehicles," International J. of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering, vol. 9, pp. 390–393, 2015.

[2] A. R. Girard, A. S. Howell and J.K. Hedrick, "Border patrol and surveillance missions using multiple unmanned air vehicles," 43rd IEEE Conf. on Decision and Control, vol.1, pp. 620–625, 2004.

[3] F. Lin, X. Dong and B. M. Chen, "A robust real-time embedded vision system on an unmanned rotorcraft for ground target following," IEEE Trans. on Industrial Electronics, vol.59, pp. 1038–1049, 2012.

[4] A. Tsukerman, M. Weiss, T. Shima and D. Löbl, "Optimal rendezvous guidance laws with application to civil autonomous aerial refueling," J. of Guidance Control and Dynamics, vol. 41, pp. 1–8, 2018.

[5] M. Kumar, K. Cohen and B. HomChaudhuri, "Cooperative control of multiple uninhabited aerial vehicles for monitoring and fighting wildfires," J. of Aerospace Computing, Information, and Communication, vol. 8, pp. 1–16, 2011.

[6] C. P. C. Chanel, F. Teichteil-Königsbuch and C. Lesire, "Multi-target detection and recognition by UAVs using online POMDPs," AAAI Conf.e on Artificial Intel., pp. 1381–1387, 2013.

[7] F. Rafi, S. Khan and K. Shafiq, "Hybrid tests of contact events in air-to-air refueling," J. of Aircraft, vol. 55, pp. 1–12, 2018.

[8] J. P. Nalepka and J. L. Hinchman, "Automated aerial refueling: extending the effectiveness of unmanned air vehicles," AIAA Modeling and Simulation Technologies Conf. and Exhibit, pp. 15–18, 2005.

[9] P. R. Thomas, U. Bhandari and S. Bullock, "Advances in air to air refueling, " Progress in Aerospace Sciences, vol. 71, pp. 14–35, 2014.

[10] J. Valasek, K. Gunnam and J. Kimmett, "Vision-based sensor and navigation system for autonomous air refueling," Journal of Guidance, Control, and Dynamics, vol. 28, pp. 979–989, 2005.

[11] X. Wang, N. Cui and J. Guo, "INS/VisNav/GPS relative navigation system for UAV," Aerospace Science and Technology, vol. 28, pp. 242–248, 2013.

[12] C. I. Chen and R. Stettner, "Drogue tracking using 3D flash LIDAR for autonomous aerial refueling," SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, pp. 80370:1–11, 2011.

[13] J. Doebbler, T. Spaeth and J. Valasek, "Boom and receptacle autonomous air refueling using visual snake optical sensor," J. of Guidance, Control, and Dynamics, vol. 30, pp. 1753–1769, 2007.

[14] H. B. Duan and Q. F. Zhang, "Visual measurement in simulation environment for vision-based UAV autonomous aerial refueling," IEEE Trans. on Instr. and Measurement, vol.64, pp. 2468–2480, 2015.

[15] L. Gan and H. B. Duan, "Robust binocular pose estimation based on pigeon-inspired optimization," IEEE 10th Conference on Industrial Electronics and Applications (ICIEA), pp. 1043–1048, 2015.

[16] Y. H. Shi, "Brain storm optimization algorithm, Advances in Swarm Intelligence," Springer Berlin Heidelberg, pp. 303–309, 2011.

[17] S. Cheng, J. F. Chen, X. J. Lei and Y. H. Shi, "Locating multiple optima via brain storm optimization algorithms," IEEE Access, vol. 6, pp. 17039–17049, 2018.

[18] J. N. Li and H. B. Duan, "Simplified brain storm optimization approach to control parameter optimization in F/A-18 automatic carrier landing system," Aerospace Science and Technology, vol. 42, pp. 187–195, 2015.

[19] C. Li and H. B. Duan, "Information granulation-based fuzzy RBFNN for image fusion based on chaotic brain storm optimization," Optik-International Journal for Light and Electron Optics, vol. 126, pp. 1400–1406, 2015.

[20] M. Mammarella, G. Campa and M.R. Napolitano, "Comparison of point matching algorithms for the UAV aerial refueling problem," Machine Vision and Applications, vol. 21,pp. 241–251, 2010.

[21] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, vol. 24, pp. 381–395, 1981.

[22] S. Li, C. Xu and M. Xie, "A robust O (n) solution to the perspective-n-point problem," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 34, pp. 1444–1450, 2012.

[23] W. H. Press, "Numerical Recipes 3rd Edition: The Art of Scientific Computing," 2007.

[24] C. Y. Tsai, K. J. Hsu, and H. Nisar, "Efficient model-based object pose estimation based on multi-template tracking and PnP algorithms," Algorithms, vol.11, pp. 11–17, 2018.

[25] T. Hamel and C. Samson, "Riccati observers for the nonstationary PnP problem," IEEE Trans. on Automatic Control, vol. 63, pp. 726–741, 2018.

[26] Y. X. X, Y. L. J and C. Fang, "Generalized orthogonal iterative algorithm for pose estimation of multiple camera systems," Acta Optica Sinica, vol. 29, pp. 72–77, 2009.

[27] V. Lepetit, F. Moreno-Noguer and P. Fua, "EPnP: An accurate o (n) solution to the PnP problem," International J.of Computer Vision, vol. pp. 155–166, 2009.

[28] Y. I. Abdel-Aziz, H. M. Karara and M. Hauck, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," Photogrammetric Engineering & Remote Sensing, vol. 81, pp. 103–107, 2015.

[29] X. Yang, J. B. Guo, T. L. Xue and K. T. Cheng, "Robust and real-time pose tracking for augmented reality on mobile devices," Multimedia Tools and Applications, vol. 77, pp. 6607–6628, 2018.