sensors

MDPI

Article

Monocular Visual Position and Attitude Estimation Method of a Drogue Based on Coaxial Constraints

Kedong Zhao *, Yongrong Sun, Yi Zhang and Hua Li

Navigation Research Center, College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; sunyr@nuaa.edu.cn (Y.S.); zynuaa@nuaa.edu.cn (Y.Z.); huali@nuaa.edu.cn (H.L.)

* Correspondence: zhaokedong1993@nuaa.edu.cn

Abstract: In aerial refueling, there exists deformation of the circular feature on the drogue's stabilizing umbrella to a certain extent, which causes the problem of duality of position estimation by a single circular feature. In this paper, a monocular visual position and attitude estimation method of a drogue is proposed based on the coaxial constraints. Firstly, a procedure for scene recovery from one single circle is introduced. The coaxial constraints of the drogue are proposed and proved to be useful for the duality's elimination by analyzing the matrix of the spatial structure. Furthermore, we came up with our method, which is composed of fitting the parameters of the spatial circles by restoring the 3D points on it, using the two-level coaxial constraints to eliminate the duality, and optimizing the normal vector of the plane where the inner circle is located. Finally, the effectiveness and robustness of the method proposed in this paper are verified, and the influence of the coaxial circle's spatial structure on the method is explored through simulations of and experiments on a drogue model. Under the interference of a large amount of noise, the duality elimination success rate of our method can also be maintained at a level that is more than 10% higher than others. In addition, the accuracy of the normal vector obtained by the fusion algorithm is improved, and the mean angle error is reduced by more than 26.7%.

Keywords: monocular vision; aerial drogue; position and attitude estimation; circular feature; duality

check for updates

Citation: Zhao, K.; Sun, Y.; Zhang, Y.; Li, H. Monocular Visual Position and Attitude Estimation Method of a Drogue Based on Coaxial Constraints. *Sensors* **2021**, *21*, 5673. https://doi.org/10.3390/s21165673

Academic Editors: Umar Iqbal and Michael J. Korenberg

Received: 12 July 2021
Accepted: 18 August 2021
Published: 23 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Aerial refueling technology can significantly extend the endurance of aircraft, which is of great importance in strategic or tactical aviation operations [1]. It is well known at present that there are two kinds of aerial refueling system: the probe-and-drogue refueling system pioneered by Flight Refueling Ltd. [2] and the Flying Boom refueling system developed by Boeing [3]. Due to the economy and flexibility of the probe-and-drogue refueling method and its adoption by many countries, its use is advisable for autonomous aerial refueling [4]. In a docking situation, the principal issue is how to accurately and quickly measure the relative position and attitude between the probe (from the receiving aircraft) and the drogue (from the tanker) during the end game [4–6]. Monocular vision is the most popular and fastest-growing method and has been extensively studied by researchers owing to its feasibility, ease of calibration, low cost, passive nature, and large effective field of view [7,8]. Therefore, it is a common navigation method for the AAR docking stage [1,4,5].

Typical geometric features used for pose estimation include points, lines, and circles in vision-based relative pose measurements. Whether in theoretical research [9–12] or application [13–16], scholars at home and abroad have done a lot of research. Compared with other features, circular features are common and widespread in the aviation field. Circular features have a high anti-interference ability with respect to occlusion owing to their geometric characteristics. It is an accepted and advantageous method in aeronautics

and astronautics to obtain the navigation information by using the circular features of the target to accomplish the task of the aircraft and spacecraft [17,18].

The shape of a drogue is a typical body of revolution (BOR) that contains many coaxial circular features. However, with a change in aircraft speed, the stabilizing umbrella of the drogue is in different opening and closing states, which leads to the deformation of the circular features to some extent during the refueling process. It is difficult to obtain in real time the effective radius of circular features for monocular vision pose estimation. Only the shape of the circle near the oil joint is fixed, which is called the 'inner circle' of the drogue in this paper. The 5 degrees of freedom (DOF) pose of the drogue can be estimated by the inner circle.

Nonetheless, there exists the problem of a dual solution for the pose estimation by a single circular feature, that is, there are two sets of the circle center's positions and the normal vectors of the circle plane [19]. If and only if the optical center and the target circle form a cone is there a unique solution.

To eliminate the duality, in References [18,20] the authors proposed to construct a reference point in a plane or space and used the invariance in the Euclidean distance to the circle center as a constraint to select the correct pose. However, this method relies on a known point feature; in Reference [21], the constraint angle from a motion reconstruction was utilized to eliminate the duality, that is, the space angle does not change with the motion of the rigid body, but it requires dense point features in the scene for reconstruction. For the case of multiple circular features, the authors suggested that the pose estimation of the target be constrained by multiple parallel circular or cylindrical features [19,22]. The implicit condition is that the radius of all coaxial circles is known, which is not satisfied in the drogue case. The features of a double planar circle were applied to estimate the pose of the target in Reference [23] and calibrate the camera parameters in References [24,25]. In addition, stereo vision was introduced to the pose estimation by providing additional information to the scene reconstruction in References [26,27].

All of the methods mentioned above are limited by specific requirements and special conditions, such as the prerequisite of all information about the coaxial circles, the introduction of auxiliary equipment other than a monocular camera, the existence or installation of special structures that can be recognized, or the need for many features for reconstruction. Notwithstanding, it is not possible or at least difficult to install an artificial mark on the drogue and introduce other auxiliary equipment into the airborne monocular vision system for visual navigation in AAR or an Actively Stabilized Refueling Drogue System (ASRDS).

Aiming at solving the problem that only a single circular feature known in advance can be directly used for pose estimation, while the others are deformed to a certain extent during air refueling, a monocular visual position and attitude estimation method of a drogue based on coaxial constraints is proposed in this paper. It makes full use of multiple circular features of the drogue itself, and the proposed coaxial constraints can effectively eliminate the duality of solutions. Moreover, the accuracy of the target's normal vector, optimized by fusing multiple circular features, is greatly improved.

The paper is organized as follows. Section 2 presents the method for scene recovery from a circular feature. In Section 3, we propose the coaxial constraints and prove them. The position and attitude estimation algorithm based on the coaxial constraints is presented in Section 4. Section 5 presents the simulations and experiments conducted to evaluate the proposed methodology. Section 6 presents the conclusions.

2. Scene Recovery from a Circle

2.1. Projection of a Circle

Define the world coordinate system, camera coordinate system, and image coordinate system of the object as $O_w - X_wY_wZ_w$, $O_c - X_cY_cZ_c$, and $O_I - UV$, respectively. Let $\tilde{\mathbf{x}}^{(w)} = (x_w, y_w, z_w, 1)^T$ denote the 3D homogeneous coordinates of a 3D point in the world coordinate system, and let $\tilde{\mathbf{x}}^{(I)} = (u, v, 1)^T$ be the homogeneous coordinates of its projection

in the image coordinate system. The projection from the world coordinate system to the image can be described as:

$$z_c\tilde{\mathbf{x}}^{(I)} = \mathbf{K}\mathbf{x}^{(c)} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\tilde{\mathbf{x}}^{(w)} \quad with \quad \mathbf{K} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where z_c is a scale factor (with a projection depth of $\mathbf{x}$); $\mathbf{K}$ is the camera intrinsic matrix, with the focal length (f_u, f_v) and the principle point (u_0, v_0); $\mathbf{R}$ and $\mathbf{t}$ are the rotation matrix and the translation vector from the world coordinate system to the camera coordinate system, respectively; and $\mathbf{x}^{(c)} = (x_c, y_c, z_c)^T \in \mathbb{R}^3$ denote the non-homogeneous coordinates of the corresponding 3D point in the camera coordinate system.

Without loss of generality, all the discussions in this paper assume that the camera intrinsic matrix $\mathbf{K}$ has been calibrated. We establish a world coordinate system where the x–y plane lies on the plane of the circle and the z-axis of the coordinate system is perpendicular to the plane and faces away from the camera. A point $\tilde{\mathbf{x}}_p{}^{(w)} = (x_p, y_p, z_p, 1)^T$ $with$ $z_p = 0$ on the circle P of which the center homogeneous coordinates are $(x_o, y_o, 0, 1)^T$ with a radius of r_0 can be expressed by the following formula.

$$\begin{cases} (\tilde{\mathbf{x}}_p{}^{(xy)})^T\mathbf{P}\tilde{\mathbf{x}}_p{}^{(xy)} = 0 \quad with \quad \tilde{\mathbf{x}}_p{}^{(xy)} = (x_p, y_p, 1)^T \quad and \quad z_p = 0 \\ \mathbf{P} = \begin{bmatrix} 1 & 0 & -x_o \\ 0 & 1 & -y_o \\ -x_o & -y_o & {x_o}^2 + {y_o}^2 - {r_0}^2 \end{bmatrix} \end{cases} \tag{2}$$

Combining Equation (1) with Equation (2), the projection of the circle in the image coordinate system is derived as:

$$\begin{cases} (\tilde{\mathbf{x}}^{(I)})^T\mathbf{Q}\tilde{\mathbf{x}}^{(I)} = 0 \\ \mathbf{Q} = \mu\big(\mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}\big)^{-T}\mathbf{P}\big(\mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}\big)^{-1} \end{cases} \tag{3}$$

Denote $\mathbf{H} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$. $\mathbf{Q}$ can be rewritten as $\mathbf{Q} = \mu\mathbf{H}^{-T}\mathbf{P}\mathbf{H}^{-1}$, where μ is a scale factor and equal to ${z_c}^2$. $\mathbf{r}_1$ and $\mathbf{r}_2$ are the first and second columns, respectively, of the rotation matrix $\mathbf{R}$, and $\mathbf{t}$ is the translation vector.

Since $\mathbf{Q}$ is a symmetric matrix, it forms an elliptical cone by the optical center of the camera and the spatial circle under the conditions of the principles and the technique of imaging. Without considering the degradation that occurs when the projection of the circle becomes a line, the projection of the circle is an ellipse corresponding to the intersection of the image plane and the elliptical cone. The projection of the spatial circle on the image plane satisfies the ellipse constraint, which is consistent with the conclusion [25].

2.2. Position and Attitude Estimation from a Circle

The solution of a pose estimation problem from one circle is the position of the circle center $\mathbf{O}^{(c)} = (x, y, z)^T$ and the normal vector $\mathbf{n}^{(c)} = (n_1, n_2, n_3)^T$ of the plane where the circle is located in the camera coordinate system due to the rotational symmetry of the circle feature. The normal vector is equivalent to the direction of the z-axis of the world coordinate system in the camera coordinate system.

When the camera intrinsic matrix $\mathbf{K}$ has been calibrated, it is easy to transfer the image coordinates $\tilde{\mathbf{x}}^{(I)}$ with $\mathbf{K}^{-1}$. For the convenience of derivation, let $\mathbf{K} = \mathbf{I}_3 = diag(1, 1, 1)$. Then, the matrix $\mathbf{H}$ can be rewritten and the inverse of the matrix can be derived as:

$$\begin{cases} \mathbf{H} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} = \mathbf{R}\big[\mathbf{e}_1, \mathbf{e}2, \mathbf{R}^T\mathbf{t}\big] \\ \mathbf{H}^{-1} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{s}]\mathbf{R}^T \end{cases} \tag{4}$$

where $\mathbf{e}_1 = (1,0,0)^T$, $\mathbf{e}_2 = (0,1,0)^T$, and $\mathbf{s} = (s_1, s_2, s_3)^T$. $\mathbf{R}$, $\mathbf{t}$, and $\mathbf{s}$ follow the rule $\mathbf{R}^T\mathbf{t} = {s_3}^{-1}\begin{bmatrix} -s_1 \\ -s_2 \\ 1 \end{bmatrix}$ with $s_3 \neq 0$. Combined with Equation (4), it can be derived from Equation (3) that:

$$\mathbf{Q} = \mu\mathbf{R}\mathbf{M}\mathbf{R}^T \quad \text{with} \quad \mathbf{M} = \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \\ s_1 & s_2 & {s_1}^2 + {s_2}^2 - (s_3 r_0)^2 \end{bmatrix} \tag{5}$$

Under perspective geometry, the parameter matrix of a projected ellipse and the parameter matrix of the spatial circle are similar. According to Reference [28], the analytical solutions of position and attitude can be settled based on the theorem that if two matrices are similar, they have the same eigenvalues. Four sets of circle centers and normal vectors can be obtained by the Eigen-decomposition of the matrix $\mathbf{Q}$, as:

$$\begin{aligned} \mathbf{o}^{(c)} &= -\omega_1 r\sqrt{-\tfrac{\lambda_3(\lambda_1-\lambda_2)}{\lambda_1(\lambda_1-\lambda_3)}}\mathbf{v}_1 - \omega_2 r\sqrt{-\tfrac{\lambda_1(\lambda_3-\lambda_2)}{\lambda_3(\lambda_3-\lambda_1)}}\mathbf{v}_3 \\ \mathbf{n}^{(c)} &= -\omega_1\sqrt{-\tfrac{\lambda_1-\lambda_2}{\lambda_1-\lambda_3}}\mathbf{v}_1 - \omega_2\sqrt{-\tfrac{\lambda_3-\lambda_2}{\lambda_3-\lambda_1}}\mathbf{v}_3 \end{aligned} \tag{6}$$

where $\omega_1 = \pm 1, \omega_2 = \pm 1$. λ_i are the eigenvalues of the matrix $\mathbf{Q}$ and $\mathbf{v}_i$ are the corresponding eigenvectors. According to the constraint that the z-axis of the camera coordinate system points to the target, two sets of solutions can be eliminated. The remaining two sets of solutions are the duality problems for the position and attitude estimation by a single circular feature under normal circumstances.

3. The Coaxial Constraints on a Drogue

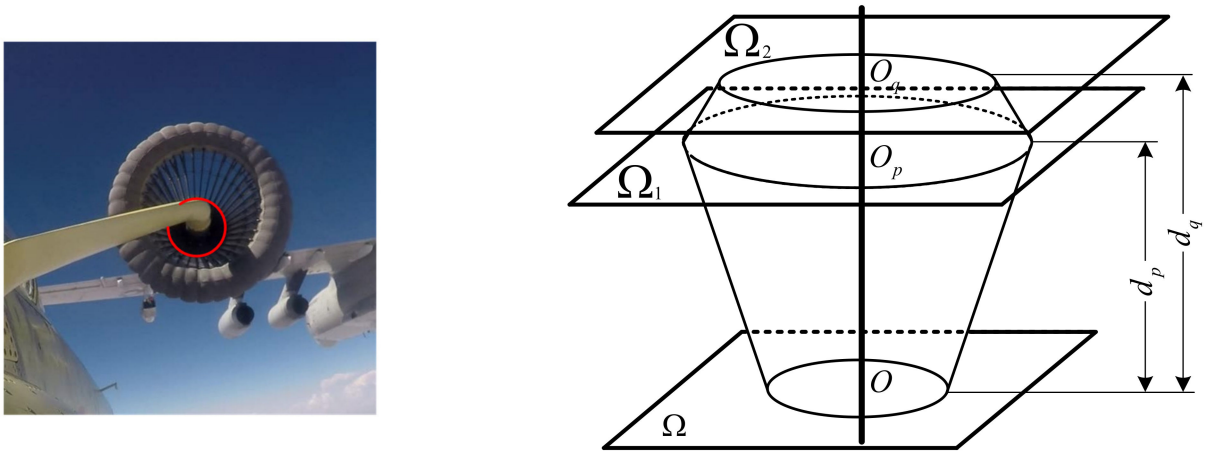
3.1. Analysis of a Drogue

During the docking stage of autonomous aerial refueling, the control system that controls the motion state of the tanker and receiver, such as position, attitude, and speed, is extremely dependent on the relative pose between the drogue and the receiver obtained by the visual navigation system. In addition, the actively stabilized refueling drogue system (ASRDS) for stabilizing the drogue also relies on visual navigation to monitor the position and attitude between the drogue and the probe in real time.

In these scenes, the target with which visual navigation needs to deal is the drogue. The shape of the drogue shown in Figure 1 is approximately a BOR, and its outer contour is the surface of revolution (SOR), which contains many coaxial circular features. As the motion state of the tanker changes, the stabilizing umbrella of the drogue transitions between different opening and closing states. As a result, the actual size of most circular features is not fixed and its reference value cannot be obtained in advance. In this case, only the radius of the circle marked with the solid red line in Figure 1 is fixed, which is defined as the inner circle of the drogue in this paper and can be used directly to yield the analytical solution of position and attitude.

Nevertheless, the planes on which the circular features are located remain parallel, and the distance between the planes changes little. The vectors connecting the centers of circles are also collinear with the normal vector of each circle on the whole. Therefore, the coaxial constraints are proposed as follows:

1. There exists a structure of spatial circles on several planes parallel to the plane of the inner circle.
2. The centers of multiple spatial circles are collinear and the vectors composed of the circles' center are collinear with the normal vector of the inner circle.



(**a**) (**b**)

Figure 1. This is a figure of a drogue for aerial refueling. (**a**) The drogue during aerial refueling seen from the perspective of the receiver; (**b**) a side view of the drogue model.

3.2. Proof

In this section, we prove that the constraints proposed in Section 3.1 can be adopted to eliminate the duality of the estimation of position and attitude. According to the two sets of solutions calculated by the inner circle and the distance between the planes of circles, two corresponding spatial structures can be restored in combination with the image. Denote Ω_j^1 and Ω_j^2 the j-th plane parallel to the inner circle corresponding to the two sets of solutions, and denote $\widehat{O_j^1}$ and $\widehat{O_j^2}$ the intersection curves between the space structures and the j-th plane parallel to the inner circle. If it can be proved that the space structure recovered from the false solution does not meet the coaxial constraints, that is, $\widehat{O_j^1}$ and $\widehat{O_j^2}$ are not circular at the same time, and the vectors composed of the circles' center are not collinear with the normal vector of the inner circle, so the coaxial constraints can be used to eliminate the duality.

Without special instructions, the coordinates below are all in the camera coordinate system. Suppose that two sets of 5 DOF poses of the inner circle are calculated as $(\mathbf{O}^1, \mathbf{n}^1)$ and $(\mathbf{O}^2, \mathbf{n}^2)$ by Section 2.2. Among them, $\mathbf{O}^1$ and $\mathbf{O}^2$ are the candidate coordinates of the inner circle center, while $\mathbf{n}^1$ and $\mathbf{n}^2$ are the candidate normal vectors of the inner circle plane.

Two rotation matrices, denoted $\overline{\mathrm{R}}_1$ and $\overline{\mathrm{R}}_2$, were constructed by $\mathbf{n}^1$ and $\mathbf{n}^2$, respectively, which make the z-axis of the camera coordinate system parallel to the normal vector $\mathbf{n}^1$ or $\mathbf{n}^2$, and the z-axis coordinate of the target is positive. The coordinate system transforms as follows:

$$\begin{cases} \mathbf{X}_1 = \overline{\mathrm{R}}_1 \mathbf{X} \\ \mathbf{X}_2 = \overline{\mathrm{R}}_2 \mathbf{X} \end{cases} \tag{7}$$

where $\mathbf{X}$ are the coordinates in the original camera coordinate system; $\mathbf{X}_1$ and $\mathbf{X}_2$ are in the transformed camera coordinate system. The relationships between them are as follows:

$$\mathbf{X}_1 = \overline{\mathrm{R}}_1 \overline{\mathrm{R}}_2^T \mathbf{X}_2 = \overline{\mathrm{R}} \mathbf{X}_2 \tag{8}$$

In accordance with the cause of duality, the camera's optical center and the inner circle, denoted O, constitute an oblique cone space in the two transformed coordinate systems. The matrix expression of the oblique cone surface is given as follows:

$$\mathbf{X}_i{}^T \mathbf{\Phi}_i \mathbf{X}_i = \mathbf{X}_i{}^T \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ \alpha & \beta & \gamma \end{bmatrix} \mathbf{X}_i = 0, i = 1, 2 \tag{9}$$

That is, the two transformed coordinate systems are constrained to the following formula:

$$\begin{cases} \mathbf{X}_1^T \mathbf{\Phi}_1 \mathbf{X}_1 = 0 \\ \mathbf{X}_2^T \mathbf{\Phi}_2 \mathbf{X}_2 = 0 \end{cases} \tag{10}$$

Substituting Equation (8) into Equation (10), we can obtain:

$$\mathbf{\Phi}_2 = \overline{\mathrm{R}}^T \mathbf{\Phi}_1 \overline{\mathrm{R}} \tag{11}$$

If the hypothesis that the dual solutions of the circle O make the two space structures of coaxial circle P satisfy the coaxial constraint (1) is not null, the camera's optical center and circle P also form an oblique cone in the two transformed coordinate systems, $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$, transformed by $\overline{\mathrm{R}}_1$ and $\overline{\mathrm{R}}_2$, that satisfies:

$$\mathbf{\Psi}_2 = \overline{\mathrm{R}}^T \mathbf{\Psi}_1 \overline{\mathrm{R}} \tag{12}$$

From the coaxial constraint (2), $\mathbf{\Psi}_1$, $\mathbf{\Phi}_1$ and $\mathbf{\Psi}_2$, $\mathbf{\Phi}_2$ satisfy the following formula:

$$\begin{cases} \mathbf{\Phi}_1 = \mathbf{\Psi}_1 - \begin{bmatrix} 0 & 0 & (k-1)\alpha_{q1} \\ 0 & 0 & (k-1)\beta_{q1} \\ (k-1)\alpha_{q1} & (k-1)\beta_{q1} & \delta\gamma \end{bmatrix} = \mathbf{\Psi}_1 - \mathbf{\Delta}_1 \\ \mathbf{\Phi}_2 = \mathbf{\Psi}_2 - \begin{bmatrix} 0 & 0 & (k-1)\alpha_{q2} \\ 0 & 0 & (k-1)\beta_{q2} \\ (k-1)\alpha_{q2} & (k-1)\beta_{q2} & \delta\gamma \end{bmatrix} = \mathbf{\Psi}_2 - \mathbf{\Delta}_2 \end{cases} \tag{13}$$

where $k \neq 1$ is the ratio of the z-axis coordinates of circles O and P in the two transformed coordinate systems.

$$k = \frac{z_{q1}}{z_{p1}} = \frac{z_{q2}}{z_{p2}} = \frac{z_q}{z_q + z_j^{(w)}} \quad with \quad z_{q1} = \left(\mathbf{O}^1\right)^T \cdot \mathbf{n}^1, z_{q2} = \left(\mathbf{O}^2\right)^T \cdot \mathbf{n}^2 \tag{14}$$

$z_j^{(w)}$ is the z-axis coordinate of P in the world coordinate system. From Equation (6), we can obtain: $z_{q1} = z_{q2} = z_q$. Denote

$$\begin{aligned} \overline{\mathrm{R}} &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \\ \mathbf{\Delta}_1 &= \begin{bmatrix} 0 & 0 & \alpha_1 \\ 0 & 0 & \beta_1 \\ \alpha_1 & \beta_1 & \gamma \end{bmatrix} \\ \mathbf{\Delta}_2 &= \begin{bmatrix} 0 & 0 & \alpha_2 \\ 0 & 0 & \beta_2 \\ \alpha_2 & \beta_2 & \gamma \end{bmatrix} \end{aligned} \tag{15}$$

Combining Equation (11) with Equation (14), it can be inferred that:

$$\mathbf{\Delta}_1 \overline{\mathrm{R}} = \overline{\mathrm{R}} \mathbf{\Delta}_2 \tag{16}$$

Substituting Equation (15) into Equation (16), Equation (16) can be written as:

$$\begin{aligned} &\begin{bmatrix} \alpha_1 r_{31} & \alpha_1 r_{32} & \alpha_1 r_{33} \\ \beta_1 r_{31} & \beta_1 r_{32} & \beta_1 r_{33} \\ \alpha_1 r_{11} + \beta_1 r_{21} + \gamma r_{31} & \alpha_1 r_{12} + \beta_1 r_{22} + \gamma r_{32} & \alpha_1 r_{13} + \beta_1 r_{23} + \gamma r_{33} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_2 r_{13} & \beta_2 r_{13} & \alpha_2 r_{11} + \beta_2 r_{12} + \gamma r_{13} \\ \alpha_2 r_{23} & \beta_2 r_{23} & \alpha_2 r_{21} + \beta_2 r_{22} + \gamma r_{23} \\ \alpha_2 r_{33} & \beta_2 r_{33} & \alpha_2 r_{31} + \beta_2 r_{32} + \gamma r_{33} \end{bmatrix} \end{aligned} \tag{17}$$

Analyzing the formula above, there exist the following situations:

1. $r_{13} = r_{31} = 0, r_{23} = r_{32} = 0$
   In this case, r_{33} must be 1. There is a rotation transformation around the z-axis between the two coordinate systems with $\overline{\mathrm{R}}$ as the rotation matrix, which means that the position and the normal vector of the circle are not changed. In other words, $(\mathbf{O}^1, \mathbf{n}^1)$ and $(\mathbf{O}^2, \mathbf{n}^2)$ are equal.
2. Other Situations

Comparing the left and right sides of Equation (17), we can obtain $\alpha_1 = \alpha_2 0$, $\beta_1 = \beta_2 0$, and $\gamma 0$, where means possibly equal. Combining Equation (13) and Equation (17), we can obtain $\alpha_{q1} = \alpha_{q2}, \beta_{q1} = \beta_{q2}$. If the intersecting curve is of the same radius, the following equation must exist.

$$\left(\alpha_{q1}{}^2 + \beta_{q1}{}^2 - \gamma_{q1}\right) z_{q1}{}^2 = r_z{}^2 = \left(\alpha_{q2}{}^2 + \beta_{q2}{}^2 - \gamma_{q2}\right) z_{q2}{}^2 \tag{18}$$

From Equation (18), we can obtain $\gamma_{q1} = \gamma_{q2}$. It can also be inferred that $(\mathbf{O}^1, \mathbf{n}^1)$ are equal to $(\mathbf{O}^2, \mathbf{n}^2)$.

In summary, the space structure recovered from the false analytical solution, which is caused by the duality of the pose estimation based on one circular feature, does not meet the coaxial constraints. Thereby, the coaxial constraints proposed by Section 3.1 can be utilized to eliminate the duality.

4. The Position and Attitude Estimation Method Based on the Coaxial Constraints

The Pseudo-algorithm of the position and attitude estimation method based on the coaxial constraints is shown in Algorithm 1.

The algorithm uses the parameters of the ellipse, such as center coordinates, the major axis, the minor axis, the inclination angle, the radius of the inner circle, and the z-axis coordinates of the inner coaxial circles in the world coordinate system as input.

First of all, the matrix expression of ellipses from the input parameters of the ellipses and the position and attitude with the known condition of the inner circle are calculated by the algorithm in Section 2.2. Then, two corresponding spatial structures are restored using the method described in Section 4.1. The method described in Section 4.2 is used to eliminate the duality. Finally, the normal vector of the target is optimized by fusing multiple circular features in the spatial structure.

4.1. Spatial Structure Recovery

The spatial relationship of a 3D point and circle corresponding to the duality is shown in Figure 2. Take the coaxial circle corresponding to $(\mathbf{O}^1, \mathbf{n}^1)$ as an example, which is on the plane parallel to the inner circle, and the z-axis coordinate is $z_j^{(w)}$ in the world coordinate system. The restored 3D point $\mathbf{x}_*^{(c)} = \left(x^*, y^*, z^*\right)^T$ on the coaxial circle is the intersection of $\overrightarrow{\mathbf{O}_c\mathbf{U}_*}$ (the vector connecting the camera's optical center $\mathbf{O}_c$ and the imaging point $\mathbf{U}_*$) with the space plane, where the coaxial circle is located. $\mathbf{x}_*^{(c)}$ satisfies the following:

$$\begin{cases} z^* \tilde{\mathrm{U}}_* = \mathbf{K}\mathbf{x}_*^{(c)} \\ \left(\mathbf{n}^1\right)^T \mathbf{x}_*^{(c)} - d = 0 \end{cases} \tag{19}$$

where $d = \frac{\left\|(\mathbf{n}^1)^T \mathbf{O}^1\right\|}{\left\|\mathbf{n}^1\right\|} + z_j^{(w)}$, according to the relationship between the two planes. It can be inferred that:

$$\mathbf{x}_*^{(c)} = \frac{d}{\left(\mathbf{n}^1\right)^T \mathbf{K}^{-1} \tilde{\mathrm{U}}_*} \mathbf{K}^{-1} \tilde{\mathrm{U}}_* \tag{20}$$

Algorithm 1: Position and Attitude Estimation Algorithm Based on the Coaxial Constraints

Input:

Image: *The parameters of ellipses have been fitted and matched*

r: *Radius of inner circle*

$z_j^{(w)}$**:** *The z-axis coordinates of coaxial circles' center*

Output:

O: *The position of drogue in the camera coordinate system*

***n*:** *The normal vector of drogue in the camera coordinate system*

function COAXIALCIRCLEPOSE(***Image***, ***r***, $z_j^{(w)}$)

Transform the parameters of ellipses into matrix expression

Calculate two sets of position and attitude of the inner circle $(\mathbf{O}^1, \mathbf{n}^1)$,$(\mathbf{O}^2, \mathbf{n}^2)$

for $j =$ ***1*** to ***J*** do

Restore the spatial structures of $(\widehat{O}_j^1, \widehat{O}_j^2)$ *by* $\mathbf{x}_*^{(c)} = \frac{d}{(\mathbf{n}^T)\mathbf{K}^{-1}\tilde{\mathbf{U}}_*}\mathbf{K}^{-1}\tilde{\mathbf{U}}_*$

Fit $\left\{\mathbf{O}_j^1, r_j^1, \Delta r_j^1\right\}$ *and* $\left\{\mathbf{O}_j^2, r_j^2, \Delta r_j^2\right\}$ *from* $(\widehat{O}_j^1, \widehat{O}_j^2)$

Calculate $(\mathbf{O}_j^1, \mathbf{n}_j^1)$ *and* $(\mathbf{O}_j^2, \mathbf{n}_j^2)$ *from* $\begin{cases} r_j^1, & if \quad \Delta r_j^1 < \Delta r_j^2 \\ r_j^2, & if \quad \Delta r_j^1 > \Delta r_j^2 \end{cases}$

end for

Calculate ϕ^1 *and* ϕ^2 *by*

$$\phi^i = \sum_{j=1}^{J} \left(\frac{\Delta r_j^i}{r_j^1 + r_j^2}\right) i = 1, 2$$

Eliminate duality by

$$(\mathbf{O}, \mathbf{n}) = \begin{cases} \left(\mathbf{O}^1, \mathbf{n}^1\right), & if \quad \phi^1 < \phi^2 \\ \left(\mathbf{O}^2, \mathbf{n}^2\right), & if \quad \phi^1 > \phi^2 \end{cases}$$

Eliminate duality by

$$(\mathbf{O}, \mathbf{n}) = \begin{cases} \left(\mathbf{O}^1, \mathbf{n}^1\right), & if \quad \left(\mathbf{n}_j^1 \quad \text{or} \quad \mathbf{n}_j^2\right)^T \mathbf{n}^1 \approx 1 \\ \left(\mathbf{O}^2, \mathbf{n}^2\right), & if \quad \left(\mathbf{n}_j^1 \quad \text{or} \quad \mathbf{n}_j^2\right)^T \mathbf{n}^2 \approx 1 \end{cases}$$

Optimize normal vector by

$$\hat{\mathbf{n}} = \underset{\zeta}{\operatorname{argmin}} \left(\sum_{j=1}^{J} \left\| \mathbf{n}_j^{(c)} - \zeta \right\|^2 + \left\| \mathbf{n}^{(c)} - \zeta \right\|^2 \right) \quad \text{with} \quad \zeta^T \zeta = 1$$

Obtain translation vector by $\hat{\mathbf{t}} = \mathbf{O}$

result $\leftarrow$ $\hat{\mathbf{n}}, \hat{\mathbf{t}}$

return *result*

end function

Select N points on the curve uniformly and restore the coordinates of the 3D points on the plane Ω_1 in terms of Equation (20). Denote the N 3D points, the center of the coaxial circle, and the midpoint of any two different points as $\mathbf{x}_*^{(c)}{}_i = \left(x_i^*, y_i^*, z_i^*\right)^T$, $\mathbf{O}_* = \left(x_o^*, y_o^*, z_o^*\right)^T$, and $\mathbf{x}_*^{(c)}{}_{ij} = \left(\frac{x_i^*+x_j^*}{2}, \frac{y_i^*+y_j^*}{2}, \frac{z_i^*+z_j^*}{2}\right)^T$, respectively, where $i, j = 1, 2, \ldots N$. Since the center of a circle must be on the intersection line of the mid vertical plane of any two points, the vector $\overrightarrow{\mathbf{O}_1\mathbf{x}_*^{(c)}{}_{ij}}$ should be perpendicular to $\overrightarrow{\mathbf{x}_*^{(c)}{}_i\mathbf{x}_*^{(c)}{}_j}$. This means that:

$$\left(x_j^* - x_i^*, y_j^* - y_i^*, z_j^* - z_i^*\right)\left(\frac{x_i^* + x_j^*}{2} - x_o^*, \frac{y_i^* + y_j^*}{2} - y_o^*, \frac{z_i^* + z_j^*}{2} - z_o^*\right)^T = 0 \tag{21}$$

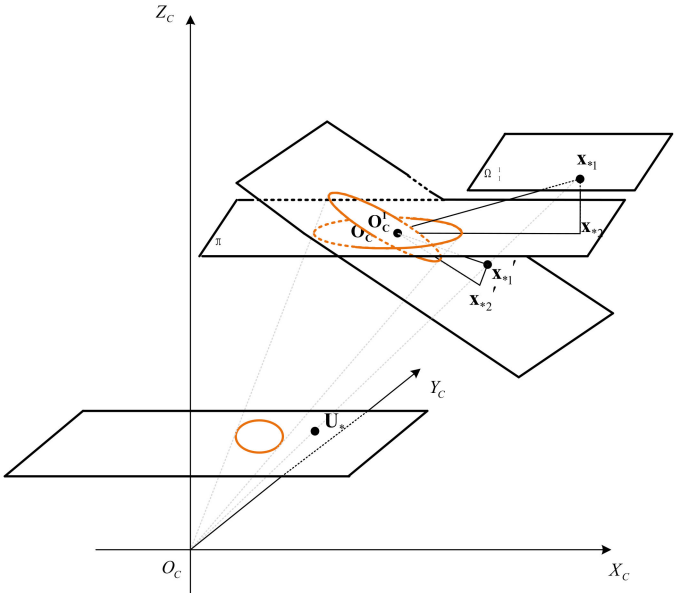


Figure 2. This is a figure of the spatial relationship of a point and circle corresponding to the duality.

After sorting out the equation, we can obtain

$$\Delta x_{ij}^* \cdot x_o^* + \Delta y_{ij}^* \cdot y_o^* + \Delta z_{ij}^* \cdot z_o^* = c_{ij} \tag{22}$$

where $\Delta x_{ij}^* = x_i^* - x_j^*$, $\Delta y_{ij}^* = y_i^* - y_j^*$, $\Delta z_{ij}^* = z_i^* - z_j^*$, and $c_{ij} = \frac{(x_i^*)^2+(y_i^*)^2+(z_i^*)^2-(x_j^*)^2-(y_j^*)^2-(z_j^*)^2}{2}$. N-1 linearly independent vectors, denoted $\left(\Delta x_{ij}^*, \Delta y_{ij}^*, \Delta z_{ij}^*\right)^T$, can be obtained from the N 3D points, and the following formula can be obtained:

$$\begin{bmatrix} \Delta x_{12}^* & \Delta y_{12}^* & \Delta z_{12}^* \\ \Delta x_{23}^* & \Delta y_{23}^* & \Delta z_{23}^* \\ \cdots & \cdots & \cdots \\ \Delta x_{N-1N}^* & \Delta y_{N-1N}^* & \Delta z_{N-1N}^* \end{bmatrix} \begin{bmatrix} x_o^* & y_o^* & z_o^* \end{bmatrix}^T = \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_{N-1} \end{bmatrix} \tag{23}$$

Rewrite Equation (23) as:

$$\mathbf{AO}_* = \mathbf{C} \tag{24}$$

Since the circle's center must be on Ω_1, which can be represented by $\mathbf{n}_1$ (the normal vector) and d (the intercept), we can define the error equation as:

$$f(\mathbf{O}_*) = \|\mathbf{AO}_* - \mathbf{C}\|^2 + \lambda(\mathbf{BO}_* - 1) \quad with \quad \mathbf{B} = \mathbf{n}_1/d \tag{25}$$

The coordinates of the circle's center are the least square solution of the above equation.

$$\begin{bmatrix} \mathbf{O}_* \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A}^T\mathbf{C} \\ 1 \end{bmatrix} \tag{26}$$

Furthermore, the radius and the roundness of the circle are obtained by Equation (27). We define the roundness of the circle, denoted Δr^*, as the average roundness of the points.

$$\begin{cases} l_i = \sqrt{\left(x_i^* - x_o^*\right)^2 + \left(y_i^* - y_o^*\right)^2 + \left(z_i^* - z_o^*\right)^2} \\ r = \frac{1}{N}\sum_{i=1}^{N} l_i \\ \Delta r = \sum_{i=1}^{N} |l_i - r| \end{cases} \tag{27}$$

The radius and the roundness of other coaxial circles can also be calculated.

4.2. Elimination of Duality

As the name suggests, the two-level coaxial constraints algorithm consists of two parts. The first level is the coaxial constraint (1) in Section 3.1, that is, the elimination of the duality by using the roundness of the spatial structure. After restoring the spatial structures by Section 4.1, we can obtain the radius and roundness of the coaxial circles, denoted r_j^1, r_j^2 and $\Delta r_j^1, \Delta r_j^2$, respectively, corresponding to the two pose solutions of the inner circle. If there are J coaxial circles, the definition is as follows:

$$\phi^i = \sum_{j=1}^{J} \left(\frac{\Delta r_j^i}{r_j^1 + r_j^2} \right) \quad with \quad i = 1, 2 \quad and \quad j = 1, 2 \dots J \tag{28}$$

Then, the first-level constraint can be expressed as the following equation.

$$(\mathbf{O}, \mathbf{n}) = \begin{cases} \left(\mathbf{O}^1, \mathbf{n}^1\right) & if \quad \phi^1 < \phi^2 \\ \left(\mathbf{O}^2, \mathbf{n}^2\right) & if \quad \phi^1 > \phi^2 \end{cases} \tag{29}$$

The second-level constraint is the coaxial constraint (2) in Section 3.1, that is, the normal vectors of multiple planes where the circles are located are parallel. Using the radius calculated by Equation (27), two sets of positions and attitudes of multiple coaxial circles, denoted $\left(\mathbf{O}_j^1, \mathbf{n}_j^1\right)$ and $\left(\mathbf{O}_j^2, \mathbf{n}_j^2\right)$, respectively, can be solved for. The second-level coaxial constraint can be written as:

$$(\mathbf{O}, \mathbf{n}) = \begin{cases} \left(\mathbf{O}^1, \mathbf{n}^1\right) & if \quad \left(\begin{array}{ccc} \mathbf{n}_j^1 & \text{or} & \mathbf{n}_j^2 \end{array} \right)^T \mathbf{n}^1 \approx 1 \\ \left(\mathbf{O}^2, \mathbf{n}^2\right) & if \quad \left(\begin{array}{ccc} \mathbf{n}_j^1 & \text{or} & \mathbf{n}_j^2 \end{array} \right)^T \mathbf{n}^2 \approx 1 \end{cases} \tag{30}$$

After the coaxial constraints are applied, the normal vectors of the planes where multiple circles are located can also be obtained.

4.3. Fusion of Multiple Circular Features

After the two-level coaxial constraints described in the previous subsection are applied, the center position of the inner circle and the normal vectors of multiple planes can be obtained. In the pose fusion problem of multiple coaxial circles with a known radius, there are the following relationships among the circles.

$$\mathbf{O}_j^{(c)} = {z_j}^{(w)} \mathbf{n}^{(c)} + \mathbf{t}^{(c)} \tag{31}$$

where ${z_j}^{(w)}$ is the z-axis coordinate of O_j in the world coordinate system. Let $\mathbf{O}_{jk}^{(c)} = \mathbf{O}_j^{(c)} - \mathbf{O}_k^{(c)} \quad with \quad j, k \in \{1, 2 \dots J\}$, where $\mathbf{O}_j^{(c)}$ and $\mathbf{O}_k^{(c)}$ are the centers of the coaxial circles obtained by Section 4.2, then

$$\mathbf{O}_{jk}^{(c)} = (z_j^{(w)} - z_k^{(w)}) \mathbf{n}^{(c)} \tag{32}$$

In Reference [29], it was shown that the pose fusion algorithm for multiple coaxial circles with a known radius is equivalent to solving the following equation.

$$\hat{\mathbf{n}} = \underset{\zeta}{\operatorname{argmin}} \left(\sum_{j=1}^{J} \left\| \mathbf{n}_j^{(c)} - \zeta \right\|^2 + \sum_{j,k \in \{1,2 \dots J\}} \left\| \frac{\mathbf{O}_{jk}^{(c)}}{z_j^{(w)} - z_k^{(w)}} - \zeta \right\|^2 \right) \quad with \quad \zeta^T \zeta = 1 \tag{33}$$

The analysis of Equation (6) shows that the radius estimation error has little influence on the error of $\mathbf{n}_j^{(c)}$ but some influence on the error of $\mathbf{O}_j^{(c)}$. When the value of $\left| z_j^{(w)} - z_k^{(w)} \right|$

is small, the vector $\frac{\mathbf{O}_{jk}^{(c)}}{z_j^{(w)}-z_k^{(w)}}$ has a little error. Accordingly, we use the following formula to solve $\hat{\mathbf{n}}$.

$$\hat{\mathbf{n}} = \underset{\zeta}{\operatorname{argmin}}\left(\sum_{j=1}^{J}\left\|\mathbf{n}_j^{(c)}-\zeta\right\|^2+\left\|\mathbf{n}^{(c)}-\zeta\right\|^2\right) \quad \text{with} \quad \zeta^T\zeta = 1 \tag{34}$$

where $\mathbf{n}_j^{(c)}$ and $\mathbf{n}^{(c)}$ are the normal vectors of the coaxial circle and the inner circle solved by Section 4.2. Considering that the error of $\mathbf{O}_j$ is a little larger, the solution for the position of the inner circle is selected as the corresponding $\hat{\mathbf{t}}$.

So far, the position and the normal vector of the target with coaxial constraints have been solved.

5. Results and Analysis

5.1. Evaluation Indices

The simulations of and experiments on the drogue model were designed to verify the effectiveness of the proposed algorithm. The performance of the algorithm was evaluated in terms of two aspects: the duality elimination success rate and the accuracy of the normal vector of the target.

Assuming that the position and the normal vector of the target calculated by our algorithm were $(\hat{\mathbf{n}}, \hat{\mathbf{t}})$ and the benchmarks we set were $(\mathbf{R}_g, \mathbf{t}_g)$, the error was defined to be composed of the angle error $\Delta\theta$ (the angle between the solved normal vector and the ground truth $\mathbf{r}_3$, where $\mathbf{r}_3$ is the third column of the matrix $\mathbf{R}_g$) and the translation error $\Delta\mathbf{t}$.

$$\begin{aligned}\Delta\theta &= \arccos\left(\hat{\mathbf{n}}^T\mathbf{r}_3\right)\\ \Delta\mathbf{t} &= \left\|\hat{\mathbf{t}}-\mathbf{t}_g\right\|\end{aligned} \tag{35}$$

The evaluation criterion of the successful elimination of the duality was that the angle error and the translation error of $(\hat{\mathbf{n}}, \hat{\mathbf{t}})$ were less than what was eliminated.

In this paper, the success rate of duality elimination was defined as the percentage of frames with successful elimination of the duality with respect to the total number of image frames.

5.2. Simulations

The simulations were carried out on a simulation system based on MATLAB, in which the resolution was 1600 × 1200 pixels, the pixel size was 0.01 × 0.01 mm, the focal distance was 8 mm, and the principal point of the photograph was (800.00, 600.00).

To verify the anti-interference ability and the robustness to relative distances of the algorithm, we carried out comparative simulations under different conditions, including noise in the image feature (an error in fitting the feature's parameters), noise in the spatial structure (deformation of the drogue's frame), and different relative distances between the camera and the target.

The world coordinate system was established with the inner circle's center as the origin and the z-axis perpendicular to the plane of the inner circle and facing away from the camera.

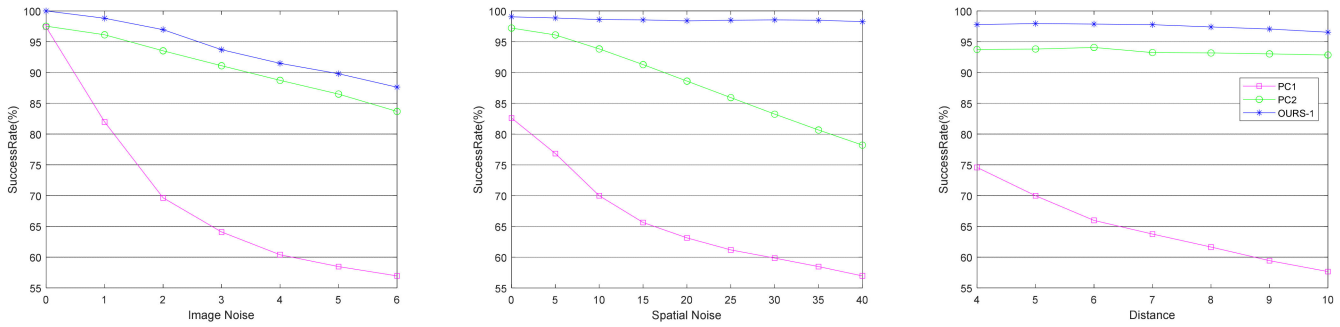
The parameters of the target were set according to the measurements of the actual drogue. The radius of the inner circle was $r = 125.0$ mm, and the coordinates of the inner circle's center were $(0, 0, 0)$; the radius of the circle O_1 was $r_1 = 385.0$ mm, and the coordinates of the circle's center were $(0, 0, -210)$; the radius of the circle O_2 was $r_2 = 320.0$ mm, and the coordinates of the circle's center were $(0, 0, -310)$. Only O_2 was used in the duality elimination simulations.

1. Elimination of Duality

- Noise in the Image Feature

The z-axis coordinate of the target in the camera coordinate system was 5 m, and the x-axis coordinate, y-axis coordinate, heading, pitch, and roll were set as a random value

within the range of ±1.5 m, ±1.5 m, ±30°, ±30°, and ±15°, respectively. We added white gaussian noise with a mean value of 0 and a variance of 10 to the center coordinates of O_2, and noise with a mean value of 0 and a variance of 20 to r_2. The noise in the image features was set as white gaussian noise with a mean value of 0 and a variance of interval 1 from 1 to 6. According to the above simulation conditions, images of the target were generated, and the features in the image with the noise disturbance were extracted to solve for the position and the normal vector. Each test generated 2000 simulated images to run the algorithm, and the results of each simulation condition are the average of 10 tests. The results are shown in Figure 3.



(**a**) (**b**) (**c**)

Figure 3. This is a figure showing the success rate of the algorithm under different simulation conditions. The curves corresponding to the algorithm in References [18,20] are marked as PC1 and PC2. The coordinates of the two points were set to (300, 300, 0) and (150, 150, −500), and gaussian noise with a mean of 0 and a variance of 5 was added to the position. (**a**) The success rate of the algorithms under different levels of noise in the image; (**b**) The success rate of the algorithms under different levels of noise in the spatial structure; (**c**) The success rate of the algorithms under different *z*-axis coordinates of the targets.

- Noise in the Spatial Structure

The setting of the target's pose in the camera coordinate system was the same as above. We added white gaussian noise with a mean of 0 and a variance of 2 to the image features. At this time, the noise in the coordinates of O_2 was set as white gaussian noise with a mean value of 0 and a variance of interval 5 from 0 to 40. The noise in r_2 was twice the noise in O_2. Under the above conditions, each test generated 2000 simulated images to run the algorithm, and the results of each simulation condition are an average of 10 tests. The results are shown in Figure 3b.

- Different *z*-axis Coordinates of Targets

We added white gaussian noise with a mean of 0 and a variance of 2 to the image features, white gaussian noise with a mean value of 0 and a variance of 20 to O_2, and white gaussian noise with a mean value of 0 and a variance of 40 to r_2. The settings of the *x*-axis coordinate, *y*-axis coordinate, and attitude were the same as above, but the *z*-axis coordinates were 4–10 m with an interval of 1. Under the above conditions, each test generated 2000 simulated images to run the algorithm, and the results of each simulation condition are an average of 10 tests. The results are shown in Figure 3c.

To verify the effectiveness of our algorithm and consider its applicability to drogues for air refueling, we selected the algorithms in References [18,20] for comparison, which are the curves marked as PC1 and PC2, respectively, in Figure 3.

The results in Figure 3a show that the proposed algorithm for duality elimination has a good anti-interference ability against the noise in the image features. It also has excellent performance when there is a large disturbance. That is, when the variance is 5, the success rate can be higher than 90%. The algorithm in Reference [20] also has a high degree of robustness to the noise in the image features, but it is about 3% lower than ours.

The results in Figure 3b demonstrate that the proposed algorithm is surprisingly robust to the noise in the spatial structure. Under noise with a variance of 40, the algorithm in Reference [18] almost fails, but our success rate is higher than 98.5%, which also has obvious advantages compared with the 78.20% of the algorithm in Reference [20].

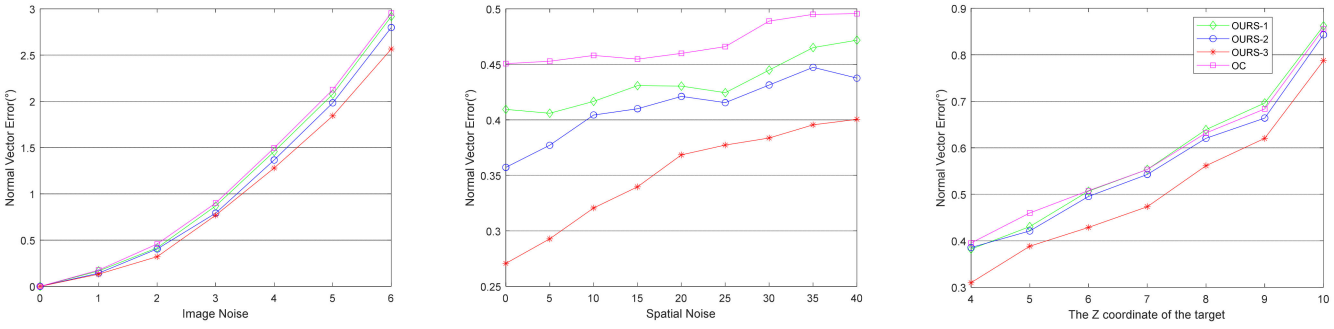
This shows that the performance of our algorithm remains pretty consistent with changes in the radius of the coaxial circle and the distance between the planes caused by the opening and closing of the drogue's stabilizing umbrella during aerial refueling.

It is worth mentioning that there is about a 30% difference between the results of the different simulations using the algorithms in References [18,20]. This means that the two algorithms are sensitive to the pose of the target.

The results in Figure 3c show that our algorithm is effective when different z-axis coordinates of the target within 10 m are used. Moreover, it can obtain a higher success rate (about 6% higher) than the algorithm in [20] under the same conditions.

2. Accuracy of the Normal Vector

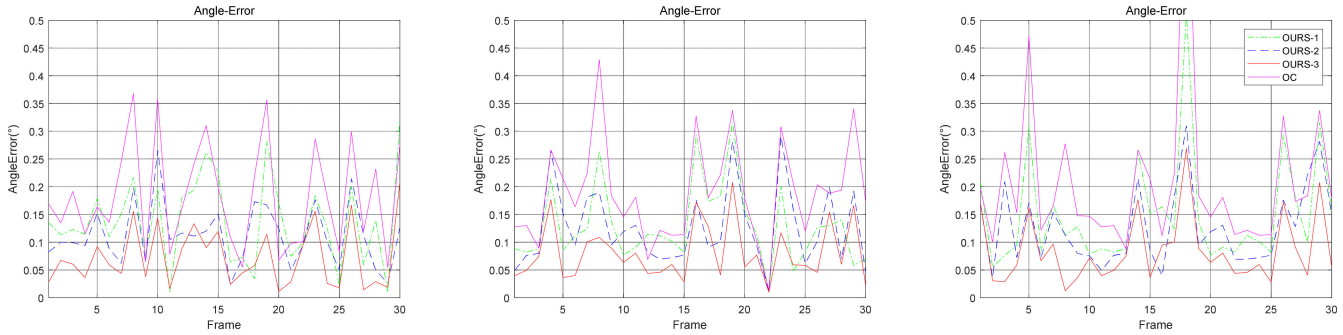
We also evaluated the accuracy of the normal vector solved by the algorithm for the fusion of multiple circular features under the same simulation conditions as above. We compared the accuracy of the normal vectors solved by the different algorithms in simulations. The results are shown in Figures 4 and 5.



(**a**) (**b**) (**c**)

Figure 4. This is a figure about the angle error of the normal vector under different simulation conditions. The normal vector that only eliminates duality by O_1 and O_2 is marked as OC, and the normal vectors solved by the fusion algorithm with O_1, O_2 and both O_1 and O_2 are marked as OURS-1, OURS-2, and OURS-3, respectively. (**a**) The angle error of the normal vector under different levels of noise in image features; (**b**) The angle error of the normal vector under different levels of noise in the spatial structure; (**c**) The angle error of the normal vector under different z-axis coordinates of the targets.

Comparing the curves marked OURS-1, OURS-2, OURS-3, and OC in Figure 4a–c, it is noticeable that fusing multiple coaxial circles can effectively improve the accuracy of the normal vector. The mean value of the curves for each simulation is shown in Table 1. From the curves OC and OURS-3, the data show that our fusion algorithm with both O1 and O2 can effectively improve the accuracy of the target's normal vector. The advantages are quite obvious since the angle errors of the normal vector solved by the fusion algorithm with both O1 and O2 achieve good results. In addition, the angle errors corresponding to the simulations of different levels of image feature noise and spatial structure noise and different z-axis coordinates are reduced by 18.4%, 25.6%, and 15.4%, respectively, compared with the result that only eliminates duality by O1 and O2. Once again, these results show that our algorithm is robust against the noise in the image features and spatial structure and is effective when different z-axis coordinates of the target within at least 10 m are used. The normal vector solved by the fusion algorithm with O1 and O2 (marked as OURS-1 and OURS-2, respectively) can also improve the accuracy. However, using two coaxial circles is much better than using one. Moreover, the results for O1 and O2 are different in the case of a single coaxial circle. This is discussed in the next section.



(**a**) (**b**) (**c**)

Figure 5. This is a figure showing some examples of simulations for the angle error of the normal vector. (**a**) The angle error of the normal vector when the variance in the image feature noise is 2; (**b**) The angle error of the normal vector when the variance in the spatial structure noise is 20; (**c**) The angle error of the normal vector at z-axis coordinates of 6 m.

Table 1. This is a table about the mean error of the normal vector in the curves under different simulation conditions.

Curves	Corresponding to Simulation (a)	Corresponding to Simulation (b)	Corresponding to Simulation (c)
Ours-1	0.13	0.13	0.15
Ours-2	0.11	0.12	0.12
Ours-3	0.07	0.08	0.08
OC	0.17	0.19	0.20

Notice that the curve of OC is under the curves of OURS-1 in Figure 4c when the z-axis coordinates of the target are greater than 8 m. The results here are due to the success rate of duality elimination with two coaxial circles being higher than that with one coaxial circle and the pose estimation algorithm in Section 2.1 providing a highly accurate solution with less image feature noise.

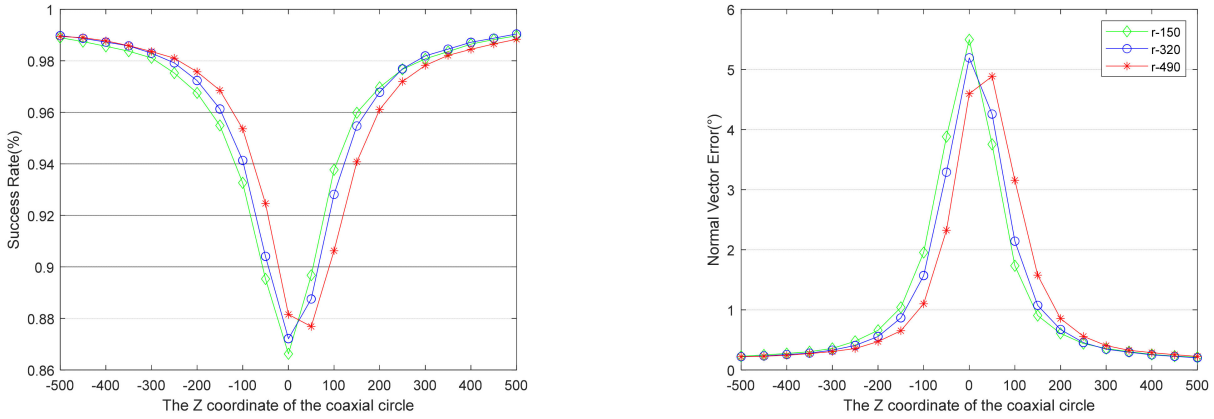
3. Influence of spatial structure on the algorithm's performance

The striking result that emerges from the curves of OURS-1 and OURS-2 in Figure 4 is that there are great differences between them. To verify the influence of different spatial structures of the coaxial circles on the algorithm, evaluation simulations were carried out.

White gaussian noise with a mean of 0 and a variance of 2 was added to the image features. Noise with a mean value of 0 and a variance of 20 was added to the position of O_1, and noise with a mean value of 0 and a variance of 30 was added to the position of r_1. The settings of the x-axis coordinate, y-axis coordinate, and attitude were the same as above.

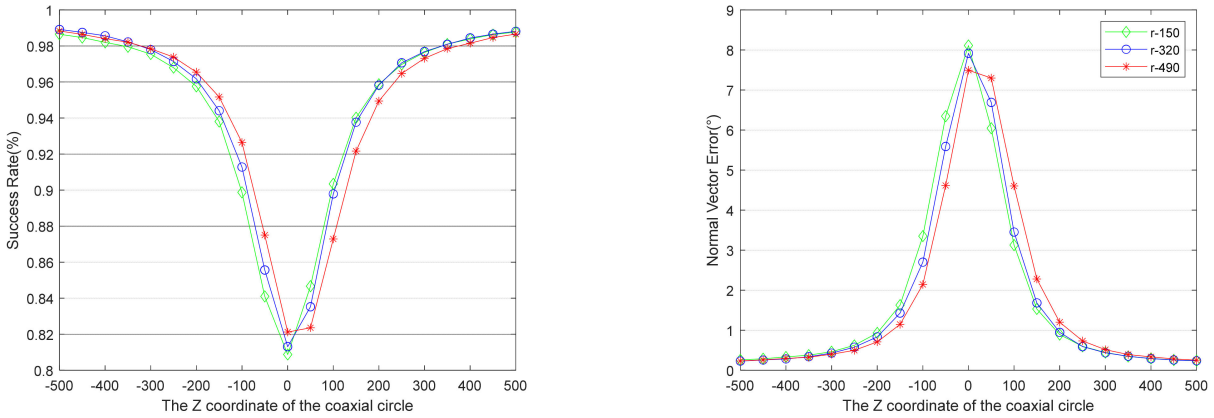
Under the above simulation conditions, comparative simulations of different z-axis coordinates of the target (5 m, 7 m, 9 m), different radii of the coaxial circles (150 mm, 320 mm, 490 mm), and different center coordinates (50 mm intervals within $\pm$500) were carried out. Each test generated 2000 simulated images to run the algorithm, and the results of each simulation condition are an average of 10 tests. The results are shown in Figures 6–8.

The relationship between the success rate and the target's spatial structure is presented in Figures 6a–8a, while the relationship between the accuracy of the normal vector and the target's spatial structure is presented in Figures 6b–8b. All the curves show that the absolute values of the coaxial circles' z-axis coordinates in the world coordinate system are smaller, and the performance of the algorithm is worse. Moreover, all the curves are not completely symmetrical when the z-axis coordinates of the target in the camera coordinate system are negative and positive. When the absolute values of the z-axis coordinates of the coaxial circle with the same radius in the world coordinate system are equal, the negative one is better than the positive one, e.g., the points with abscissa of -100 and 100 on the curve marked r-490 in Figures 6–8.



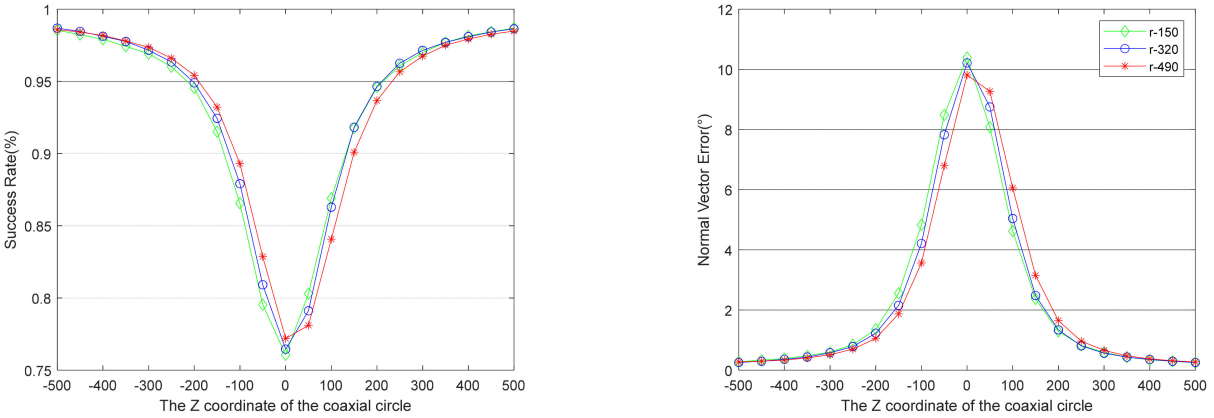
(**a**) (**b**)

Figure 6. This is a figure about the relationship between the performance of the proposed algorithm and the target's spatial structure when the *z*-axis coordinate of the target is 5 m. The curves of the different radii are marked r-150, r-320, and r-490, respectively. (**a**) The curves of the success rate. (**b**) The curves of the normal vector's angle error.



(**a**) (**b**)

Figure 7. This is a figure about the relationship between the performance of the proposed algorithm and the target's spatial structure when the *z*-axis coordinate of the target is 7 m. (**a**) The curves of the success rate. (**b**) The curves of the normal vector's angle error.



(**a**) (**b**)

Figure 8. This is a figure about the relationship between the performance of the proposed algorithm and the target's spatial structure when the *z*-axis coordinate of the drogue is 9 m. (**a**) The curves of the success rate. (**b**) The curves of the normal vector's angle error.

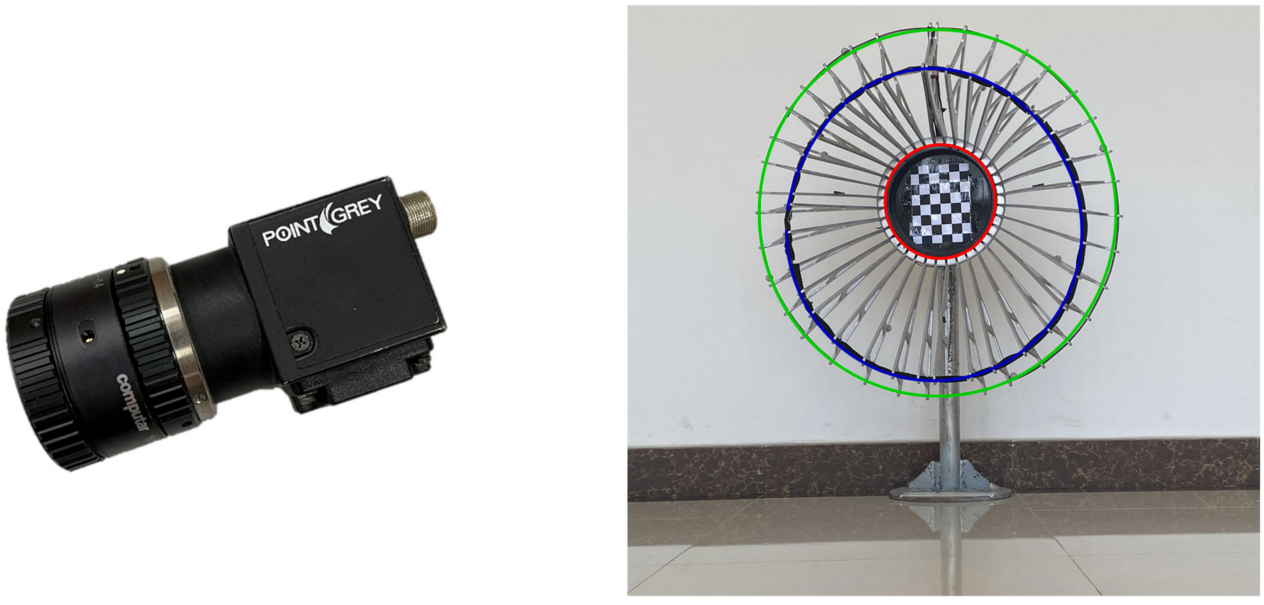
More interestingly, Figures 6–8 reveal that the relationship between the performance of the proposed algorithm and the radii of the coaxial circles is the opposite when the z-axis coordinates of the coaxial circle in the world coordinate system are positive and negative. When the z-axis coordinates of the coaxial circle are negative, the coaxial circle with a radius of 490 mm (r-490) has a better success rate and the accuracy of the normal vector is higher. In contrast, the coaxial circle with a radius of 150 mm (r-150) has better performance when the z-axis coordinates of the coaxial circle are positive.

All in all, the results reveal a relationship between the performance of the proposed algorithm and the spatial structure of the target, and the algorithm's performance is not consistent when the z-axis coordinates of the drogue in the camera coordinate system are different. Coaxial circles with a reasonable structure could greatly improve the performance of the algorithm. The following conclusions can be drawn by comparing the curves in the figure:

- The farther away the plane of the coaxial circle is from the plane of the inner circle, the better the performance of the algorithm is. It would be a good choice to select a circular feature in the plane that is more than twice the radius of the inner circle away from the plane of the inner circle. (When the two planes are very close, the false solutions of the two circles are also very similar, which leads to a reduction in the success rate of duality elimination.)
- When the distance between the plane of the coaxial circle and the plane of the inner circle is the same, the closer the coaxial circle plane is to the camera, the better the performance of the algorithm is.
- When the plane of the coaxial circle is closer to the camera than the inner circle, the larger the radius of the coaxial circle is, the better the performance of the algorithm is.
- When the plane of the coaxial circle is farther away from the camera than the inner circle, the smaller the radius of the coaxial circle is, the better the performance of the algorithm is.

5.3. Experiments on the Drogue Model

The proposed algorithm was tested on a sequence of 49 images captured by the Point Grey BFLY-U3-23S6C-C camera shown in Figure 9a with the drogue model shown in Figure 9b. The image resolution was set to 960 × 600 pixels. The focal length of the camera was calibrated as follows: fu = 1343.44, fv = 1347.53, principle point u_0 = 493.53, v_0 = 289.09.



(**a**) (**b**)

Figure 9. This is a figure about the experimental equipment. (**a**) A photo of the camera and the lens used for image capture. (**b**) A photo of the drogue model.

The drogue model was made according to a real drogue, in which the diameter of the inner circle marked with a red solid line is 250 mm. The origin of the world coordinate system of the target was located at the center of the inner circle, and the z-axis was perpendicular to the plane and faced away from the camera.

The coordinates of the other coaxial circles were $O_p = (0, 0, -210)$ and $O_q = (0, 0, -310)$. To simulate the deformation of the drogue during aerial refueling, an image of the drogue model was captured with some stretching and compression.

The ground truth corresponding to each frame was obtained by fusing and optimizing the pose of the target with a chessboard in multiple images according to the method described in [13].

Experiments were carried out with circle O_p, circle O_q, and both, which are marked O1, O2, and O3, respectively. The success rate of duality elimination is presented in Table 2, which shows that our algorithm for duality elimination has high practicability since it succeeded in all the images captured in the experiment.

Table 2. The success rate of the algorithm with different coaxial circles.

Circles	Success Rate (%)
O1	100
O2	100
O3	100

The angle error of the normal vector is shown in Figure 10, where the curve marked OC describes the normal vector's angle error with the duality eliminated by both O_p and O_q without fusion. The mean value of the curves O1, O2, O3, and OC is 0.11°, 0.10°, 0.08°, and 0.15°, respectively.

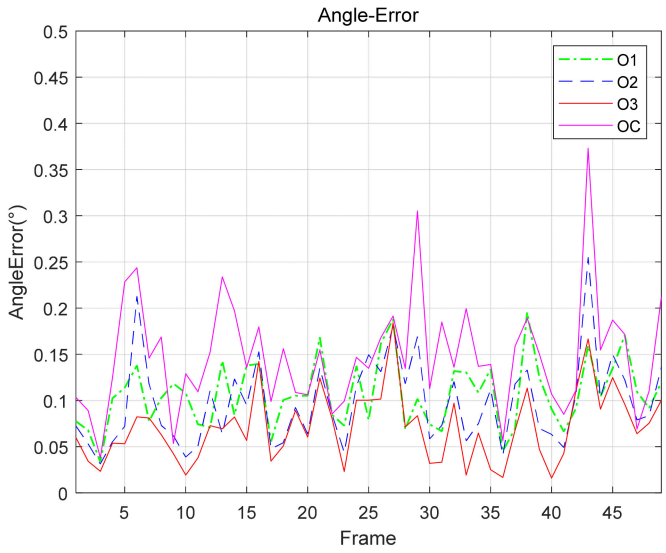


Figure 10. This is a figure of the angle error of the normal vector corresponding to the captured images.

Furthermore, to test the computational efficiency of the algorithm, points of interval 10 from 20 to 60 were selected to recover the spatial structures by Section 4.1. The case of one coaxial circle was compared with the case of two coaxial circles.

The relationship between the computation time and the number of points for the recovery of spatial structures is presented in Table 3. As the number of points increases, the computation time of the algorithm increases. In the experiments corresponding to Table 2 and Figure 10, 30 points were selected for the spatial structure recovery of each circle, and the corresponding computation time is shown (red font) in Table 3. The case of

one coaxial circle is 1.2 ms while that of two coaxial circles is 2.1 ms. This result shows that our algorithm owns a desirable real-time performance.

Table 3. This is a table of the relationship between the computation time and the number of points for spatial structure recovery.

Number of Points for Each Circle	Computation Time of 1 Coaxial Circle (ms)	Computation Time of 2 Coaxial Circles (ms)
20	0.9	1.3
30	1.2	2.1
40	1.4	2.6
50	1.6	3.1
60	1.9	3.8

Overall, these results suggest that our proposed algorithm is capable of solving the problem of estimating the pose of the drogue, which means that it can effectively eliminate duality and simultaneously improve the accuracy of the target's normal vector in real time.

6. Conclusions

In summary, we have proposed a monocular visual position and attitude estimation method of a drogue based on coaxial constraints, which can effectively eliminate the duality in pose estimation by a single circular feature and greatly improve the accuracy of the optimized normal vector. The effectiveness and robustness of the method were verified by simulations and experiments. Furthermore, we established a basis for engineering through a number of simulations and experiments. To our knowledge, no one has so far explored the relationship between the algorithm's performance and the target's spatial structures.

Author Contributions: K.Z. contributed to the development of the methodology, performed the experiments, and wrote the paper; Y.S. contributed to the project's administration, provided supervision, and reviewed this paper; and Y.Z. and H.L. collected and preprocessed the data for the experiments. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61533008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, J.; Zhen, L. Tracking and position of drogue for autonomous aerial refueling. In Proceedings of the 2018 IEEE 3rd Optoelectronics Global Conference (OGC), Shenzhen, China, 4–7 September 2018; pp. 171–175. [CrossRef]
2. Latimer-Needham, C.H. Apparatus for Aircraft-Refueling in Flight and Aircraft-Towing. U.S. Patent Application No. 2,716,527, 30 August 1955.
3. Clifford, J.L. Aircraft Interconnecting Mechanism. U.S. Patent Application No. 2,663,523, 12 December 1953.
4. Maa, Y.; Zhao, R.; Liu, E.; Zhang, Z.; Yan, K. A novel autonomous aerial refueling drogue detection and pose estimation method based on monocular vision. *Measurement* **2019**, *136*, 132–142. [CrossRef]
5. Valasek, J.; Gunnam, K.; Kimmett, J.; Tandale, M.D.; Junkins, J.L.; Hughes, D. Vision-based sensor and navigation system for autonomous air refueling. *J. Guid. Control Dyn.* **2005**, *28*, 979–989. [CrossRef]
6. Martinez, C.; Richardson, T.; Thomas, P.; du Bois, J.L.; Campoy, P. A vision-based strategy for autonomous aerial refueling tasks. *Robot. Auton. Syst.* **2013**, *61*, 876–895. [CrossRef]
7. Lin, S.; Garratt, M.A.; Lambert, A.J. Monocular vision-based real-time target recognition and tracking for autonomously landing an UAV in a cluttered shipboard environment. *Auton. Robot.* **2017**, *41*, 881–901. [CrossRef]
8. Yang, Y.; Shen, Q.; Li, J.; Deng, Z.; Wang, H.; Gao, X. Position and Attitude Estimation Method Integrating Visual Odometer and GPS. *Sensors* **2020**, *20*, 2121. [CrossRef] [PubMed]
9. Lepetit, V.; Moreno-Noguer, F.; Fua, P. EPnP: An Accurate O(n) Solution to the PnP Problem. *Int. J. Comput. Vis.* **2009**, *81*, 155–166. [CrossRef]

10. Hesch, J.A.; Roumeliotis, S.I. A Direct Least-Squares (DLS) Method for PnP. In Proceedings of the IEEE International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 383–390.
11. Kneip, L.; Li, H.; Seo, Y. UPnP: An Optimal O(n) Solution to the Absolute Pose Problem with Universal Applicability. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014; Volume 8689, pp. 127–142.
12. Ferraz, L.; Binefa, X.; Moreno-Noguer, F. Very Fast Solution to the PnP Problem with Algebraic Outlier Rejection. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 501–508. [CrossRef]
13. Sun, Y.; Huang, B.; Yang, B.; Yao, S. Vision Calibration Method of Spatial Point Position Under Multi-View Scene Determination. *J. Nanjing Univ. Aeronaut. Astronaut.* **2015**, *47*, 343–347.
14. Barone, F.; Marrazzo, M.; Oton, C.J. Camera Calibration with Weighted Direct Linear Transformation and Anisotropic Uncertainties of Image Control Points. *Sensors* **2020**, *20*, 1175. [CrossRef] [PubMed]
15. Qin, T.; Li, P.; Shen, S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Trans. Robot.* **2018**, *34*, 1004–1020. [CrossRef]
16. Zhao, J.; Hu, Y.; Tian, M. Pose Estimation of Excavator Manipulator Based on Monocular Vision Marker System. *Sensors* **2021**, *21*, 4478. [CrossRef] [PubMed]
17. Huang, B.; Sun, Y.R.; Zeng, Q.H. Real-time drogue detection and template tracking strategy for autonomous aerial refueling. *J. Real-Time Image Process.* **2020**, *17*, 437–446. [CrossRef]
18. Miao, X.; Zhu, F.; Ding, Q.; Hao, Y.; Wu, Q.; Xia, R. Monocular vision pose measurement based on docking ring component. *Guangxue Xuebao/Acta Opt. Sin.* **2013**, *33*, 123–131. [CrossRef]
19. Lu, J.; Xia, M.; Li, W.; Guo, W.-P.; Yang, K.-C. 3-D location estimation of underwater circular features by monocular vision. *Optik* **2013**, *124*, 6444–6449. [CrossRef]
20. Zhang, L.M.; Zhu, F.; Hao, Y.M.; Wang, M.Y. Pose measurement based on a circle and a non-coplanar feature point. *Guangzi Xuebao/Acta Photonica Sin.* **2015**, *44*, 112002. [CrossRef]
21. Zhang, L.; Huang, X.; Feng, W.; Hu, T.; Liang, S. Solution of duality in pose estimation of a single circle utilizing constraint angles motion reconstruction. *Guangxue Xuebao/Acta Opt. Sin.* **2016**, *36*, 115002. [CrossRef]
22. Gurdjos, P.; Sturm, P.; Wu, Y. Euclidean structure from N $\geq$ 2 parallel circles: Theory and algorithms. In *Computer Vision—ECCV 2006*; Leonardis, A., Bischof, H., Pinz, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3951, pp. 238–252.
23. Huang, B.; Sun, Y.; Zeng, Q. General fusion frame of circles and points in vision pose estimation. *Optik* **2018**, *154*, 47–57. [CrossRef]
24. Huang, H.; Zhang, H.; Cheung, Y.-m.; IEEE. The Common Self-polar Triangle of Concentric Circles and Its Application to Camera Calibration. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 4065–4072. [CrossRef]
25. Zhao, Z.; Liu, Y. Applications of projected circle centers in camera calibration. *Mach. Vis. Appl.* **2010**, *21*, 301–307. [CrossRef]
26. Ma, W.; Sun, S.; Song, J.; Li, W. Circle Pose Estimation Based on Stereo Vision. *Appl. Mech. Mater.* **2012**, *263–266*, 2408–2413. [CrossRef]
27. Liu, Z.Y.; Liu, X.; Duan, G.F.; Tan, J.R. Precise pose and radius estimation of circular target based on binocular vision. *Meas. Sci. Technol.* **2019**, *30*, 14. [CrossRef]
28. Huang, B.; Sun, Y.; Zhu, Y.; Xiong, Z.; Liu, J. Vision pose estimation from planar dual circles in a single image. *Optik* **2016**, *127*, 4275–4280. [CrossRef]
29. Huang, B. *Research on Close-range High-accuracy Vision Navigation Technology for UAV Aerial Refueling*; Nanjing University of Aeronautics and Astronautics: Nanjing, China, 2019.