Research on Aerial Robot Based on Visual Servo

Tianchen Deng

1 Harbin institute of technology, Harbin, Heilongjiang, 150001, China

Corresponding author’s e-mail: +1170400421@stu.hit.edu.cn, dtc99926@163.com

Abstract: Towards target grasping by an aerial manipulator, an aerial manipulator system is presented, for which a separated control strategy is adopted. In this paper, a suitable adaptive visual servo controller is proposed based on the original aerial manipulator system, and a visual servo controller is designed based on the depth independent image Jacobian matrix. Thus, the grasping motion of the manipulator is controlled by the visual servo of the monocular camera, which can greatly enhance the functionality and speed of the aerial manipulator system. Furthermore, a unique single degree of freedom joint collocation camera is designed to enable the camera to lock the target object at any time during the motion of aircraft. Finally, this paper designs a unique state machine system, which can make the control system switch into any state.

1. Introduction

In recent years, with the continuous improvement of navigation and control technology, aerial robots have been successfully applied in the aerial photography, disaster site survey (flood, earthquake), wide area environmental modeling and so on. A new robot system with operational ability has gradually become a research hot spot in the field of aircraft-----aerial manipulator. Usually, the aerial manipulator is composed of rotor flying robot and multi-link manipulator. The brand-new aerial robot equips the ability of 3D flight movement and operation, and has a wide range of application prospects, such as the collection of material samples in dangerous environment, the mobile capture of ground movement or flight target, the autonomous transfer of dangerous goods (explosive), and the contact measurement of environmental information, etc.

Towards potential applications, the international research team has designed a variety of flight manipulator systems. Paper[1] introduces a kind of flying manipulator prototype which can collect tree crown samples. In the disaster rescue scene, Paper [2] introduces a dual-arm flight manipulator system which towards the closing operation of pipeline valve. For contact-oriented tasks, a flight manipulator system with compliant contact characteristics is designed in Paper[3-4]. Paper[5] introduces a flight manipulator system which can pull out the long rod fixed on the ground, which is composed of unmanned helicopter and light industrial manipulator. A flight manipulator system composed of a four-rotor aircraft and a one-degree-freedom manipulator can maneuver to grab a stationary target. There is no doubt that the stable flight control of the flight manipulator and the high precision control at the end are the key to the completion of various operational tasks. Rotor aerial robot system has the characteristics of high dimension, under-actuation and nonlinearity. When it is combined with the multi-link manipulator, the dimension of the whole aerial manipulator system is greatly increased, and the coupling between the state variables becomes extremely complex, which makes the motion control part a difficult point in this direction. The research on dynamic modeling and nonlinear control of aerial manipulator system has been greatly carried out.
After fully investigating the results of institutions in this field, this paper proposed an adaptive visual servo controller based on the original aerial manipulator. At the same time, the visual servo controller is designed based on the depth independent image Jacobian matrix. The controller uses adaptive algorithm to estimate the unknown position of feature points and we use Lyapunov theory to prove the progressive stability of the controller. Thus, the grasping of the manipulator is controlled only by the visual servo of the monocular camera, which greatly enhances the functionality and speed of the aircraft.

2. System modeling
The aerial manipulator system is shown in figure 1, Σ₁, Σ₂ and Σ₃ are inertial coordinates, aircraft coordinate system and end-effector coordinate system. Among them, The coordinate axis x₂ and y₂ of the Σ₂ point to the head of the flying robot and the right side of the body respectively. Point O is the origin of Σ₂ coordinates system, And it coincides with the barycenter of the aircraft. A point c is the centroid of the entire flight manipulator system. When the arm moves, point c position will change. roc is a vector from point o towards point c.

![Figure 1. Frames of the aerial manipulator system.](image)

\[ P_2 = [x, y, z]^T, \quad v_2 = [v_x, v_y, v_z]^T \] are the absolute position and velocity of the aerial robot in the Σ₁, respectively. The attitude of the aerial robot is represented by z-y-x Euler angle, that is, \( \Phi_2 = [\phi, \theta, \psi]^T \).

\[ R_2 = \begin{bmatrix} \cos \psi \cos \theta & -\sin \psi \cos \phi + \cos \psi \sin \theta \sin \phi & \sin \psi \sin \phi + \cos \psi \sin \theta \cos \phi \\ \sin \psi \cos \theta & \cos \psi \cos \phi + \sin \psi \sin \theta \sin \phi & -\cos \psi \sin \phi + \sin \psi \sin \theta \cos \phi \\ -\sin \theta & \sin \theta \sin \phi & \sin \theta \cos \phi \end{bmatrix} \] (1)

The position and attitude of the end-effector in the Σ₁ are \( p_3 \) and \( \Phi_3 \), respectively. The relationship of the position and attitude of the aerial robot are as follows:

\[ \begin{cases} \phi_i = R_2 \phi_{23} \\ p_3 = p_2 + R_2 p_{23} \end{cases} \] (2)

\( ^2 \mathbf{p}_{23} \) and \( ^2 \Phi_{23} \) are the position and attitude of the end-effector relative to the aerial robot in the Σ₂, respectively. The velocity and angular velocity of the end-effector are expressed as \( \mathbf{v}_3 \) and \( \omega_3 \) in the Σ₁, respectively. The relationship between the \( \mathbf{v}_3 \) and \( \omega_3 \) and the velocity and angular velocity of the six-rotor aerial robot can be calculated through the differential of formula (2). And they are as follows:

\[ \begin{cases} \mathbf{v}_3 = v_2 + R_2 (\omega_3 \times \mathbf{p}_{23} + \omega_{23}) \\ \omega_3 = R_2 (\omega_2 \times \omega_{23}) \end{cases} \] (3)
where $\omega_2$ is the angular velocity of the aerial robot in the $\Sigma_2$. $\nu_{23}$ and $\omega_{23}$ are the velocity and angular velocity of the end of the manipulator relative to the six-rotor flying robot in the $\Sigma_2$, respectively.

3. System control

3.1. Path Generator

In order to generate the proper position reference of the manipulator base when the air manipulator catches the moving target object, we define a safe operating space $\omega$, in which the air manipulator can grasp the target object and keep a safe distance between the target object and the aircraft.

![Figure 2: the safe operating space $\omega$.](image)

As shown in the figure, spatial $\omega$ is the space around the target object. The space is a part of the space($s_{\text{min}}$ and $s_{\text{max}}$) between two concentric spheroids, the center of which is the point T, also it is the center of the target object. The radius of them are $l_{\text{min}}$ and $l_{\text{max}}$. The region is also confined to two conical surfaces coaxial with the cylindrical target object. The cone angle of the cone are $\beta_{\text{min}}$ and $\beta_{\text{max}}$.

Hence, the $\omega$ in space will satisfy the following inequalities:

$$l_{\text{min}} \leq \|r_T\| \leq l_{\text{max}}$$

$$\beta_{\text{min}} \leq \{r_{Zt}, r_T\} \leq \beta_{\text{max}}$$

(4)

$r_T$ represents a vector from point $t$ to point $p$ which is any point in space $\Omega$. $r_{Zt}$ is the coordinate axial quantity of $\Sigma_T$. $\langle r_{Zt}, r_T \rangle$ represents the angle between $r_{Zt}$ and $r_T$. The above two inequalities limit the distance and direction between aircraft and target object, respectively. $l_{\text{min}}$ and $l_{\text{max}}$, $\beta_{\text{min}}$ and $\beta_{\text{max}}$ are determined by actual offline testing.

In order to hold the aerial manipulator in the safe operating space $\omega$ during the grip process, the position reference of the aircraft is set according to the current position and target position. We define $r_{Tpb}$ as a vector from the target point to the current position of the aircraft, $r_{Tpb-d}$ a vector from the target point to the desired position. $r_{Zt}$ are perpendicular to $r_{Zt}$ and $r_{Tpb}$. If we rotate $r_{Tpb}$ around the cone, we can get vector $r'_{Tpb}$. The ideal position of the UAV can be determined by the following formula: ($R(\alpha_\text{d})$ is the rotation matrix representing the desired attitude of the trajectory)[6]

$$\begin{align*}
\eta_1 &= \frac{-r_{Zt} \times r_{Tpb}}{\|r_{Tpb}\|} \\
\eta_2 &= \frac{R(\alpha_\text{d})(v_{Tpb} - (r_{Zt} \times r_{Tpb} \cdot r_{Zt})/\|r_{Zt}\| \times r_{Tpb} - (r_{Zt} \times r_{Tpb} \cdot r_{Zt})/\|r_{Zt}\|)}{\|r_{Tpb}\|} \\
\eta_{\text{d}} &= \frac{r_{Zt} \times r_{Tpb}}{\|r_{Tpb}\|} \\
\eta_{\text{o}} &= \frac{r_{Zt} \times r_{Tpb}}{\|r_{Tpb}\|} \\
\rho_{d} &= \rho + \eta_{\text{d}}
\end{align*}$$

(5)
3.2. Visual servo control in non-contact mode

In non-contact state, the core of the controller design is that the camera can lock the position of the object at any time, so that the aircraft can follow the object farther away and finally move to the appropriate workspace. The visual field of the object will not be lost during the aircraft movement. The diagram shows the visual tracking system and partial coordinate system studied in this paper.

The rotation angle of the joint with single degree of freedom is $\theta_x$, $\theta_y$ is changed in every time during servo tracking of visual system, and its value is obtained by joint angle sensor. At the same time, in order to ensure that the target can be kept in the center of the camera's field of vision as much as possible, a PD controller is designed to perform image servo and control the single-degree-of-freedom joint driven by the micro steering gear, so that the camera can be aligned with the tracking target in real time. The following figure shows the servo tracking control block diagram of the visual system.

$$
\begin{align*}
    e_y(t) &= y^* - y(t) \\
    \Delta e_y(t) &= e_y(t) - e_y(t-1) \\
    \theta(t) &= k_{p1}e_y(t) + k_{d1}\Delta e_y(t)
\end{align*}
$$

$$
\begin{align*}
    e_x(t) &= x^* - x(t) \\
    \Delta e_x(t) &= e_x(t) - e_x(t-1) \\
    \psi(t) &= k_{p2}e_x(t) + k_{d2}\Delta e_x(t)
\end{align*}
$$

(6)
\( e_x(t), e_y(t) \) —— The error in the vertical and horizontal coordinate in the \( t \) moment, \( t-1 \) represents the previous moment;

\( k_{p1}, k_{p2} \) —— Proportional parameter;

\( k_{d1}, k_{d2} \) —— Differential parameter

\( \theta(t) \) —— Control output of joint angle of single degree of freedom at \( t \) moment;

\( \psi(t) \) —— Control output of aircraft yaw angle at \( t \) moment

### 3.3. Visual servo control in contact mode

In order to design an image-based visual servo controller, the most important point is to estimate \( b_x \), the position of uncalibrated feature points in base coordinate system. Here we design a controller and use online adaptive algorithm to estimate \( b_x \). To simplify the analysis, only one feature point is considered here, but the controller is also suitable for extending to multiple feature points[8]. The current position of the feature point in the image plane is defined as \( y(t) \), and \( y_d \) the desired position. The position error \( \Delta y(t) \) of the feature point in the image plane can be expressed as:

\[
\Delta y(t) = y(t) - y_d
\]  

Suppose that the camera on the manipulator is a pinhole camera, which satisfies the principle of perspective projection imaging. It is known that the internal parameter matrix of the camera is \( \Omega \), the coordinate of the feature point in the image plane is \( y(q(t)) \) and the depth of the feature point in the camera coordinate system is \( z(q(t)) \), then the projection of the feature point in the image plane can be expressed as follows:

\[
\begin{bmatrix}
y(q(t)) \\
z(q(t))
\end{bmatrix} = \frac{1}{z(q(t))} \begin{bmatrix}
\Omega & m \\
0 & 1
\end{bmatrix} \begin{bmatrix}
x(t) \\
1
\end{bmatrix}
\]  

\( M \) is called the perspective projection matrix, which only depends on the camera's internal parameter matrix \( \omega \) and the external parameter matrix \( ^eT_e \). For the eye-hand system, because the internal parameter matrix and the external parameter matrix are constant matrix, the perspective projection matrix \( M \) is also a constant matrix in the form of: \( M = \Omega ^eT_e \)

Rewriting formula, we can get \( y(q(t)) \), the coordinate of the feature point in the image plane. Similarly, the depth \( z(q(t)) \) of the feature point in the camera coordinate system can be expressed. Then, we do differential on both sides of the equation, we can get the two-dimensional velocity \( y'(q(t)) \) of the feature point in the camera coordinate system and the depth variable \( z(q(t)) \) of the feature point in the image plane.

And then, we can get \( ^e x(t) \), the position of the feature points in the end-effector coordinate system:

\[
^e x(t) = ^eR_e(q(t))^e x + ^e p_e(q(t))
\]

\[
^e x(t) = ^e v(t) - ^e w(t) \times ^e x(t)
\]

\[
= [-I_{3x3} \text{skew} (^eR_e(q(t))^e x + ^e p_e(q(t)))]
\]

\[
\begin{bmatrix}
^eR_e(q(t)) & 0 \\
0 & ^eR_e(q(t))
\end{bmatrix} \begin{bmatrix}
v(t) \\
w(t)
\end{bmatrix}
\]  

And then, we can get the relationship between the depth change rate of the feature point in the camera coordinate system and the linear velocity \( v(t) \) and the angular velocity \( w(t) \) of the manipulator end-effector. And the two-dimensional velocity \( y'(q(t)) \) of the feature point in the image plane and the linear velocity \( v(t) \) and angular velocity \( w(t) \) of the manipulator end actuator are as follows:

\[
\hat{y}(q(t)) = \frac{1}{z(q(t))} A(y(t), q(t)) \begin{bmatrix}
v(t) \\
w(t)
\end{bmatrix}
\]
Set \( J_T(q(t)) = \frac{1}{z(q(t))} A(y(t), q(t)) \)

We can get \( A(y(t), q(t)) = (P - y(t)m_y^T) \begin{bmatrix} -I_{3 \times 3} & \text{skew}(R_y(q(t))^T x + \dot{p}_y(q(t))) \\ 0_{3 \times 3} & 0 \\ 0_{3 \times 3} & R_y(q(t)) \end{bmatrix} \) \( \text{(10)} \)

\( J_T(q(t)) \) is called the image Jacobian matrix, which represents the differential transformation relationship between the three-dimensional end-effector coordinate system and the two-dimensional image coordinate system. \( A(y(t), q(t)) \) is called the depth-independent image Jacobian matrix. The difference from the image Jacobian matrix \( J_T(q(t)) \) is that it does not contain depth information \( z(q(t)) \) of feature points and is therefore a linear function of uncalibrated position \( b_x \).

Based on the traditional PD feedback control method with gravity compensation, the following visual servo controller based on kinematics is designed:

\[
\dot{q}(t) = -J_T(q(t)) \dot{A}^T(y(t), q(t)) K \Delta y(t) - \frac{1}{2} J_T(q(t)) \ddot{b}^T(q(t)) \Delta y(t) K \Delta y(t)
\]

\( \dot{q}(t) \) represents input joint angular velocity, \( J \) represents the jacobian matrix, the \( k_1 \) is the control parameter. If we can get the estimate of the position of the feature points in the base coordinate system, the estimate of the depth independent image Jacobian matrix can be calculated.

Using Lyapunov’s theory, we can prove that, in the image-based visual servo controller designed above, the adaptive algorithm is applied to the unknown position of the feature point in the base coordinate system. When the time tends to infinity, the position error of the feature point in the image plane will approach zero, that is,

\[
\lim_{t \to \infty} \Delta y(t) = 0
\]

3.4 Robot trajectory planning

Because each joint of the manipulator is continuously controllable and easy to use mathematical formula to describe the trajectory of joint space, so this paper adopts trajectory planning based on joint space, which is beneficial to the subsequent trajectory generation. Trajectory description commonly used in robots is a bit point-to-point interpolation trajectory, cubic polynomial curve trajectory, five polynomial curve trajectory, spline curve (such as b spline) trajectory and other [7]. The general cubic polynomial curve has the advantages of less parameters, continuous position and velocity, but because of the limitation of the state quantity described and the discontinuity of acceleration, it is easy to mutate at the initial point and the end point, and the flexibility of the description is poor. B spline curve has good smoothness, the acceleration of cubic B spline curve is continuous, but the calculation is relatively large, and the control point parameters need to be determined. At the same time, it is difficult to reflect the advantage of B spline control on collision avoidance constraint in joint space trajectory planning. The five-time polynomial curve has the advantages of continuous position, velocity and acceleration. Based on the above reasons, this paper uses five polynomial curves to describe the joint motion of the flying manipulator.

The joint space trajectory function is expressed as a five-order polynomial equation:

\[
\theta(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + c_5 t^5
\]

where six of these parameters can be determined by six constraints:

\[
\begin{cases}
\theta(0) = \theta_0, & \dot{\theta}(0) = 0, & \ddot{\theta}(0) = 0, \\
\theta(tf) = \theta_f, & \dot{\theta}(tf) = 0, & \ddot{\theta}(tf) = 0
\end{cases}
\]

The trajectory planning of the manipulator can be realized by these six constraints.
3.5 State machine control

During the programming of the control system, the flag bit instruction signal of the system is defined as the bool value. After each instruction is completed, the system will change the value of the flag bit. At the same time, we can judge whether the control system has completed the instruction according to the value of the flag bit signal obtained, and then we can realize the switching between the various states of the UAV.

Next, we introduce the state of the control system corresponding to each instruction:

Search instructions —— the UAV is just taking off, when the camera on the UAV has not detected the target (too far away, etc.), and the UAV will perform the search task in the specified area until the camera can detect the target (according to the actual test, image detection is normal, about 30 meters can identify the target).

Tracking instructions —— at this time the camera has detected the target object, at this time in the non-contact mode control, the use of pd control to make the camera aligned with the object, do not lose the field of vision. At the same time, the constraint function designed before is used to make the UAV move into a safe workspace.

Grab instruction —— at this time the UAV has moved into a safe workspace, so that the UAV is in a hovering state, and control the manipulator to grab objects.

4. Conclusion

In this paper, based on the original operational aerial robot, an adaptive visual servo controller is proposed. At the same time, a visual servo controller is designed based on depth independent image Jacobian matrix. The controller uses adaptive algorithm to estimate the unknown position of feature points in base coordinate system, and Lyapunov theory to prove the progressive stability of the controller. Thus, the grasping of the manipulator is controlled only by the visual servo of the monocular camera, which greatly enhances the function and speed of the UAV. However, at present, it is only a conjecture of some solutions to the existing problems in this field. Whether it can be implemented requires some practical verification (simulation or physical). In general, if the accuracy and rapidity of image recognition can be improved, the visual method should be the future development trend and improve the functionality of UAV.

References

[1] Kutia J R, Stol K A, Xu W L. Initial flight experiments of a canopy sampling aerial manipulator[C]//International Conference on Unmanned Aircraft Systems. Piscataway, USA: IEEE, 2016: 1359-1365.

[2] Korpela C, Orsag M, Oh P. Towards valve turning using a dualarm aerial manipulator[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2014: 3411-3416.

[3] Bartelds T, Capra A, Hamaza S, et al. Compliant aerial manipulators: Toward a new generation of aerial robotic workers[J]. IEEE Robotics and Automation Letters, 2016, 1(1): 477-483.

[4] Kamel M, Alexis K, Siegwart R. Design and modeling of dexterous aerial manipulator[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2016: 4870-4876.

[5] Huber F, Kondak K, Krieger K, et al. First analysis and experiments in aerial manipulation using fully actuated redundant robot arm[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2013: 3452-3457.

[6] G. Zhang et al., "Grasp a Moving Target from the Air: System & Control of an Aerial Manipulator," 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, 2018, pp. 1681-1687.
[7] Ata A A. Optimal Trajectory Planning of Manipulators: A Review [J]. Journal of Engineering Science and Technology, 2007, 2(1): 32-54.

[8] Xinwu Liang, Hesheng Wang, Yun-Hui Liu, Weidong Chen, Jie Zhao. A unified design method for adaptive visual tracking control of robots with eye-in-hand/fixed camera configuration [J]. Automatica, 2015, 59.