

# Guidance and Navigation for UAV Airborne Docking

Daniel B. Wilson, Ali H. Göktoğan and Salah Sukkarieh  
Australian Centre for Field Robotics (ACFR)  
The University of Sydney, NSW, 2006, Australia  
{d.wilson,a.goktogan,s.sukkarieh} at acfr.usyd.edu.au

**Abstract**—Unmanned aerial vehicle (UAV) capability is currently limited by the amount of energy that can be stored onboard. Airborne docking, for mid-air refueling, is a viable solution that has been implemented with manned aircraft for decades, but has yet to be achieved with their unmanned counterparts. The prohibitive challenge is the highly accurate and reliable relative positioning performance that is necessary to dock with a small target, in the air, amidst external disturbances.

This paper presents a complete solution to airborne docking, which includes vision-aided unscented Kalman filters for leader-relative navigation and docking appendage motion estimation; and guidance that is suitable for all phases of the mission. The work concludes by demonstrating the proposed algorithms in what is thought to be the first UAV airborne docking.

## I. INTRODUCTION

Currently, unmanned aerial vehicles (UAV) are limited by the energy they can store onboard or gather from the environment. Airborne docking offers a solution by enabling in-flight refueling [22, 24], recharging and payload transfer. In addition to extended range and endurance, fuel savings of 30–40 % have been suggested for haulage flights, through the elimination of refueling stopovers [17]. But despite the clear value, the state-of-the-art in industry is still NASA/DARPA’s automated approach to docking with manned aircraft, in 2006 [8], and more recently, close formation with two Global Hawk UAVs in experiments by Northrop Grumman/DARPA [7].

Our work considers two small UAVs, flying in leader-follower formation, as shown in Figs. 1 and 2. The leader is towing a docking appendage, which is a cone shaped drogue, and the follower is tasked with docking its nose within the drogue. The vehicles communicate sensor and state data in real-time, but the leader’s intent is unknown. Knowledge of the intent would allow a more optimal and predictable path to rendezvous, as investigated in the author’s previous work [25], and allow faster leader dynamics, at closer separation. However, these advantages were considered insufficient to justify the additional communication and complexity.

Multiple phases of the mission exist, including a) rendezvous to a formation position to obtain visual observations, b) transition to a pre-contact position to obtain observations of the drogue and c) a final approach to docking. Although the techniques for autonomous docking are applicable to both large and small UAVs, the operational environment is very different. Larger vehicles operate in the benign upper atmosphere where wake turbulence is the main disturbance, whereas small



Fig. 1: The terminal approach to docking with a towed drogue.

UAVs must contend with wind and low altitude turbulence, which they are particularly susceptible to. Consequently, we adopt the rationale that scaling a solution up to larger vehicles is more feasible than the reverse.

The prohibitive challenge in airborne docking, is the highly accurate and reliable relative positioning performance that is necessary to dock with a small target, in the air, and amidst external disturbances. This paper contributes a complete solution to this challenge, which is experimentally verified in airborne docking experiments.

### A. Related Work

Airborne docking is a large, multifaceted problem, so a plethora of research exists. This review will examine the key work on relative navigation, drogue estimation and formation guidance, but, a more in-depth discussion can be found in [21]. Pioneering work at Georgia Tech [12, 23] demonstrated the first vision-only formation flight, using line of sight (LOS) measurements to a leader, from a monocular camera. Marker based, vision-only methods are also popular [6, 16, 22], but reliance on a field of view constrained sensor can be unreliable. Of particular note is the work by Darling [6] which demonstrated formation flight at 20 m separation, using a vision-only pose estimation algorithm. Various forms of the Kalman filter have been used to fuse onboard sensor data with vision-only pose estimates [4, 14] and with raw

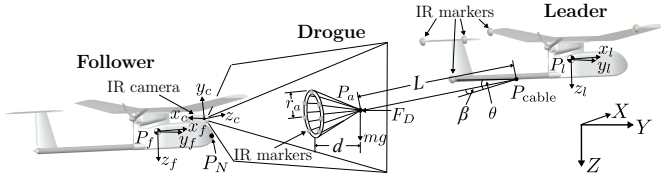


Fig. 2: Leader-follower problem setup and coordinate frames.

marker observations [10, 24]. Our work has similarities with the latter but differs by estimating the relative state directly, rather than a concatenated individual vehicle states. Further, vehicle attitudes are estimated separately, then formulated as inputs with process noise, to reduce dimensionality and avoid convergence to an ambiguous state.

Previous work on docking appendage estimation also suffers from unreliability due to the exclusive use of vision. Examples in [19, 29] tracked a drogue in manned refueling footage using *a priori* knowledge of the drogue characteristics. Monocular template matching and feature-based methods were demonstrated in [15] using a high-fidelity hardware-in-the-loop simulation. Our approach differs by using vision as an aid to correct the state of a derived dynamic model. Drogue aerodynamic parameters are also estimated to improve the open-loop accuracy when vision is unavailable.

The formation guidance literature focuses on pursuit based strategies such as proportional navigation [2] and LOS guidance [6, 20]. An implementation of vision-based LOS guidance for rendezvous with a towed drogue occurred in [18], where the authors report near misses of only a few metres. However, as noted in [18], these techniques become unstable at close range. Instead, we aim to minimise the error between the current state and a setpoint state, similarly to [3, 11]. Our approach differs by compensating for leader dynamics and proposing a terminal docking strategy that satisfies the physical constraints of the drogue.

As expected, the state-of-the-art resides within industry, where the most successful effort thus far is the NASA/DARPA Autonomous Airborne Refueling Demonstration project [8]. Although the aircraft were manned, they performed the final stage of docking automatically, using a relative GPS/INS system and a separate vision system to observe the drogue. So far, this is the only program to have demonstrated autonomous physical connection using refueling apparatus [21].

## B. System Architecture

The diagram in Fig. 3 provides an overview of the algorithm architecture, the information flow between the high-level modules, and section where each module appears in the paper. Each layer, indicated by the dashed line, builds upon the layer below. The level of the layer is analogous to the nonlinearity of the system dynamics, where higher layers have faster, more nonlinear dynamics.

The bottom layer addresses navigation and control for a single UAV, relative to the world frame, but is not addressed in this paper. The middle layer adds a second UAV and

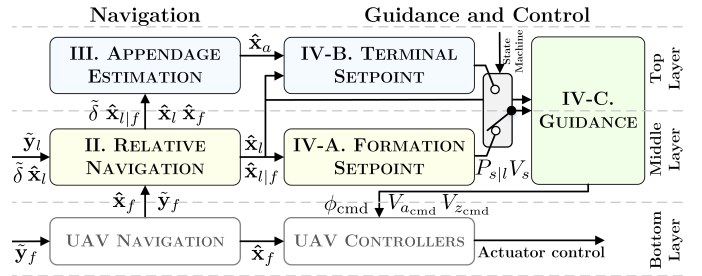


Fig. 3: The algorithm architecture and information flow.

is concerned with autonomous positioning, relative to the other aircraft. The relative navigation module in Section II estimates the relative state by aiding sensor and state data from both aircraft, with precise air-to-air LOS measurements. The formation setpoint algorithm in Section IV-A augments the commanded formation configuration to compensate for leader dynamics. The guidance in Section IV-C completes the loop by generating low level commands to minimise the error between the desired relative state and the estimated relative state.

The top layer adds a tethered drogue which has its own dynamics, relative to the leader. The state of the drogue is estimated in Section III, using the sensor and state data from each aircraft, a dynamic model and air-to-air LOS measurements. Using this estimate, the terminal setpoint algorithm in Section IV-B generates guidance setpoints that satisfy the physical constraints of the drogue, to guide the follower to docking. A state machine supervises the mission and determines which setpoints are used within the guidance module.

The complete solution is verified in airborne docking experiments in Section V, then conclusions are offered in Section VI.

## II. RELATIVE NAVIGATION

A vision-aided, relative navigation UKF estimates the relative state between two dynamic vehicles, operating in an outdoor environment with dynamic lighting and IR interference. In contrast to previous work, our approach employs vision as an aid for the onboard sensors, rather than as the primary sensor. This architecture allows accuracy to be maintained during momentary visual outages and to degrade gracefully during extended outages.

The relative state,  $\mathbf{x}_{l|f}$  in Eq. (1) comprises the relative position  $P_{l|f} = [X_{l|f} \ Y_{l|f} \ Z_{l|f}]^T$ , the relative velocity  $V_{l|f}$ , a pressure altitude bias  $h_B$  and attitude error quaternions  $\mathbf{q}_{l_B}$  and  $\mathbf{q}_{f_B}$  for the leader and follower respectively.  $P_{l|f}$  and  $V_{l|f}$  are expressed as the leader with respect to the follower, in the navigation frame.  $h_B$  is estimated to account for biases in the barometric sensing.  $\mathbf{q}_{l_B}$  accounts for misalignment between the IMU and the vehicle marker model.  $\mathbf{q}_{f_B}$  estimates the misalignment between the follower's IMU and the IR camera. Simulation and flight testing have shown that misalignment of even a few degrees onboard either aircraft creates a steady state position error that is significant enough to prevent docking.

$$\mathbf{x}_{l|f} = [P_{l|f} \ V_{l|f} \ h_B \ \mathbf{q}_{l_B} \ \mathbf{q}_{f_B}]^T \quad (1)$$

The downside to a quaternion parametrisation within an UKF is that the unit norm cannot be guaranteed when calculating the mean. To overcome this, quaternions are replaced by generalised Rodrigues parameters that represent the local rotation error, when constructing the sigma points and computing the UKF prediction and update equations. The details are omitted for brevity but can be found in [5]. Even with the additional computation required to accommodate quaternions, execution is faster than a Euler angle parametrisation because trigonometric functions are avoided.

The inputs to the system  $\mathbf{u}_{l|f}$ , are the separately estimated leader and follower attitude quaternion  $\mathbf{q}_l$  and  $\mathbf{q}_f$ ; bias corrected leader and follower acceleration measurements  $\tilde{a}_l$  and  $\tilde{a}_f$ ; and the Gaussian noise process mean of the pressure altitude bias,  $\mathbf{w}_{h_B}$ , leader alignment error  $\mathbf{w}_{\mathbf{q}_{l_B}}$  and follower alignment error  $\mathbf{w}_{\mathbf{q}_{f_B}}$ .

$$\mathbf{u}_{l|f} = [\tilde{a}_l \quad \tilde{a}_f \quad \mathbf{w}_{h_B} \quad \mathbf{q}_l \quad \mathbf{q}_f \quad \mathbf{w}_{\mathbf{q}_{l_B}} \quad \mathbf{w}_{\mathbf{q}_{f_B}}]^T \quad (2)$$

The augmented form of the unscented transform is implemented by concatenating  $\hat{\mathbf{x}}_{l|f}$  and  $\mathbf{u}_{l|f}$  and their respective covariance matrices. The augmented form with concatenated process noise has been shown to yield improved performance compared to the non-augmented form [28], at the expense of additional computation. The matrix of sigma points,  $\chi_{k-1}$ , is formed by perturbing concatenated copies of the augmented state vector by the root of the augmented covariance, scaled by  $\kappa$ .  $\kappa$  is chosen such that  $\kappa = \lambda - D_{\hat{\mathbf{x}}^a}$  where  $D_{\hat{\mathbf{x}}^a}$  is the dimension of the augmented state. We set  $\lambda$  to 3 as per [13]. Each of the sigma points are propagated by

$$\chi_k(i) = \mathbf{f}[\chi_{k-1}(i), k], \quad i = 0, 1, \dots, 2D_{\hat{\mathbf{x}}^a} \quad (3)$$

$$\mathbf{f}[\hat{\mathbf{x}}_{l|f,k}, k] = \begin{bmatrix} \hat{P}_{l|f,k} \\ \hat{V}_{l|f,k} \\ \hat{h}_{B_k} \\ \hat{\mathbf{q}}_{l_{B_k}} \\ \hat{\mathbf{q}}_{f_{B_k}} \end{bmatrix} = \begin{bmatrix} \hat{P}_{l|f,k-1} + V_{l|f,k} \Delta t \\ \hat{V}_{l|f,k-1} + \left( C_{\hat{\mathbf{q}}_{l_k}}^T \tilde{a}_l - C_{\hat{\mathbf{q}}_{f_k}}^T \tilde{a}_f \right) \Delta t \\ \hat{h}_{B_{k-1}} + \mathbf{w}_{h_B} \\ \hat{\mathbf{q}}_{l_{B_{k-1}}} \otimes \mathbf{w}_{\mathbf{q}_{l_B}} \\ \hat{\mathbf{q}}_{f_{B_{k-1}}} \otimes \mathbf{w}_{\mathbf{q}_{f_B}} \end{bmatrix} \quad (4)$$

where  $\otimes$  denotes quaternion multiplication,  $\hat{\mathbf{q}}_{l_k} = \mathbf{q}_l \otimes \hat{\mathbf{q}}_{l_{B_k}}$  and  $\hat{\mathbf{q}}_{f_k} = \mathbf{q}_f \otimes \hat{\mathbf{q}}_{f_{B_k}}$ . The notation  $C_{\mathbf{q}}$  is the direction cosine matrix (DCM) for the quaternion  $\mathbf{q}$ . The equivalent DCM for Euler angles is  $C_{(x,y,z)}$  and a rotation about the z-axis only is denoted  $C_z = C_{(0,0,z)}$ . The relative position and velocity are propagated by a first-order Euler integration while the pressure altitude bias and attitude error quaternions are modelled as random walks. The predicted mean and covariance are computed using the UKF equations in [5, 13].

#### A. Update

GPS, barometric and vision data are used, when available, to sequentially update the relative state. For each measurement update, each sigma point is transformed through the observation model to obtain the expected measurements  $\gamma_k$  as per

Eq. (5). For this step, we may use the existing propagated sigma point matrix  $\chi_k$  or resample the sigma points using the predicted mean. Our own empirical analysis did not show a significant performance difference so we shall proceed with  $\chi_k$ .

$$\gamma_k(i) = \mathbf{h}[\chi_k(i), k], \quad i = 0, 1, \dots, 2D_{\mathbf{x}} \quad (5)$$

When new GPS data is available, the observation model in Eq. (6) is used which is simply the current estimated relative position and velocity. The corresponding GPS measurements are calculated by transforming each of the position measurements to the navigation frame, then subtracting the follower position and velocity from the leader. GPS time of week is used within a queue based algorithm to synchronise the measurements and account for communication latency.

$$\mathbf{h}^{\text{GPS}}[\hat{\mathbf{x}}_{l|f,k}, k] = [\hat{P}_{l|f,k} \quad \hat{V}_{l|f,k}]^T \quad (6)$$

A barometric pressure update occurs at each timestep using Eq. (7) which is the estimated vertical displacement, corrected for an estimated altitude sensor bias. The corresponding sensor measurement is the follower's barometric altitude, subtracted from the leader's.

$$\mathbf{h}^{\text{press}}[\hat{\mathbf{x}}_{l|f,k}, k] = -\hat{Z}_{l|f,k} - \hat{h}_{B_k} \quad (7)$$

A vision update occurs when IR marker centroid observations are available. Each marker in the model  $\zeta_j^l, j = 1, 2, \dots, n$ , is de-rotated by the leader's alignment corrected attitude, offset by the estimated relative position then rotated to the follower's body frame as shown in Eq. (8a). Camera extrinsic parameters transform  $\zeta_j^f$  to the camera frame using Eq. (8b), where  $P_{f|c}$  and  $C_f^c$  are the translation and rotation from the followers body frame to the camera frame.  $C_f^c$  includes camera mounting orientation and the axes transformation.

$$\zeta_j^f = C_{\hat{\mathbf{q}}_{f_k}} (C_{\hat{\mathbf{q}}_{l_k}}^T \zeta_j^l + \hat{P}_{l|f,k}) \quad (8a)$$

$$\zeta_j^c = [C_f^c \quad P_{f|c}] \begin{bmatrix} \zeta_j^f \\ 1 \end{bmatrix} \quad (8b)$$

Finally, the expected centroid locations are calculated in Eq. (9), using the camera intrinsic matrix  $K$ .

$$\mathbf{h}^{\text{vis}}[\hat{\mathbf{x}}_{l|f,k}, k] = \begin{bmatrix} \bar{\delta}_j \\ 1 \end{bmatrix} = K \begin{bmatrix} \zeta_{x_j}^c / \zeta_{z_j}^c \\ \zeta_{y_j}^c / \zeta_{z_j}^c \\ 1 \end{bmatrix}, \quad j = 1, 2, \dots, n \quad (9)$$

Before an update can occur, correct correspondence between the projected set of markers  $\bar{\delta}_j$ , and the observed set must be determined. Our approach uses mutual-nearest-neighbour to determine the number of unique point pairs within a threshold, for every model-point hypotheses. In this way, outliers are rejected and correct correspondence is the hypothesis with the highest number of unique point pairs. Update uses the standard UKF equations [5, 13].

### III. APPENDAGE ESTIMATION

An accurate estimate of the leader-follower relative state is now known, but this alone is insufficient for docking with a drogue. To address this, another UKF is formulated with the state vector  $\mathbf{x}_a$  that describes the state of the drogue with respect to the leader. It comprises the orientation of an assumed rigid cable,  $\theta$  and  $\beta$ ; their time derivatives  $\dot{\theta}$  and  $\dot{\beta}$ ; the product of the drogue cross-sectional area and the aerodynamic drag coefficient in the longitudinal direction,  $\eta_x$ ; the same quantity in the lateral-vertical plane  $\eta_{yz}$ ; an air-relative velocity heading bias  $\psi_B$  and the length of the cable,  $L$ , as shown in Fig. 2. Estimating the drogue aerodynamic characteristics when vision is available improves the open-loop accuracy when vision is unavailable. More in-depth experimental validation of the drogue estimation algorithm can be found in author's previous work in [27], which is an extension of [26].

$$\mathbf{x}_a = [\theta \quad \dot{\theta} \quad \beta \quad \dot{\beta} \quad \eta_x \quad \eta_{yz} \quad \psi_B \quad L]^T \quad (10)$$

The system inputs  $\mathbf{u}_a$  are defined in Eq. (11) and comprise the leader airspeed  $V_{a_l}$ , leader vertical velocity  $V_{Z_l}$ , leader vertical acceleration  $\tilde{a}_l$  and the mean of each Gaussian noise processes that govern  $\eta_x$ ,  $\eta_{yz}$ ,  $\psi_B$  and  $L$ .

$$\mathbf{u}_a = [V_{a_l} \quad V_{Z_l} \quad \tilde{a}_l \quad \mathbf{w}_{\eta_x} \quad \mathbf{w}_{\eta_{yz}} \quad \mathbf{w}_{\psi_B} \quad \mathbf{w}_L]^T \quad (11)$$

#### A. Prediction

The system propagation begins by rotating the current estimate of the cable length through the current estimate of the cable orientation to obtain the the position of the drogue  $P_a = [X_a \ Y_a \ Z_a]^T$  in the leader's horizontal body frame.

$$P_a = C_{(0, \theta_{k-1}, \beta_{k-1})}^T [-L_{k-1} \quad 0 \quad 0]^T \quad (12)$$

The air-relative velocity  $v_a$ , is found by adding the motion due to a change in cable orientation, and the air-relative velocity of the frame due to the leader's motion.

$$v_a = [0 \quad \dot{\theta}_{k-1} \quad \dot{\beta}_{k-1}]^T \times P_a + [V_{a_l} \quad 0 \quad V_{Z_l}]^T \quad (13)$$

The aerodynamic drag force  $F_D$ , is calculated using the air density  $\rho$ , the current estimate of the aerodynamic parameters and the air-relative velocity vector  $v_a$ . The notation  $x^{\circ 2}$  represents the element-wise square of  $x$  and  $\vec{x}$  is the unit vector of  $x$ .

$$F_D = -\frac{1}{2} \rho [\eta_{x_{k-1}} \quad \eta_{yz_{k-1}} \quad \eta_{yz_{k-1}}]^T v_a^{\circ 2} \vec{v}_a \quad (14)$$

The moment  $M$  is the cross product of  $P_a$  and the sum of the forces due to drag, gravity and leader vertical acceleration. Angular acceleration of the cable is calculated in Eq. (16) using the inertia and mass of the drogue  $m$ .

$$M = [M_x \quad M_y \quad M_z]^T = P_a \times [F_D - mC_{\tilde{a}_k}^T \tilde{a}_l] \quad (15)$$

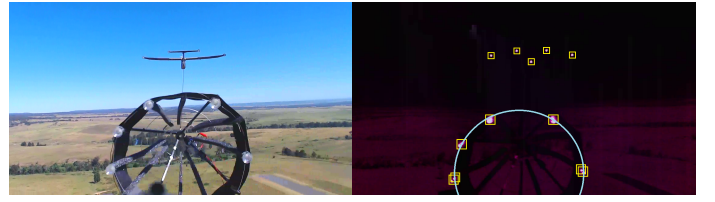


Fig. 4: An approach to docking in RGB on the left and IR on the right. IR marker detections are indicated by yellow squares and the extracted ellipse in cyan.

$$\ddot{\theta} = \frac{M_y}{m(X_a^2 + Z_a^2)} \quad (16a) \quad \ddot{\beta} = \frac{M_z}{m(X_a^2 + Y_a^2)} \quad (16b)$$

A first-order integration over  $\Delta t$  propagates the system.

$$\begin{aligned} \mathbf{f}[\hat{\mathbf{x}}_{a_k}, k] &= [\theta_k \quad \beta_k \quad \dot{\theta}_k \quad \dot{\beta}_k \quad \eta_{x_k} \quad \eta_{yz_k} \quad \psi_{B_k} \quad L_k]^T \\ &= \hat{\mathbf{x}}_{a_{k-1}} + \Delta t [\dot{\theta}_k \quad \dot{\beta}_k \quad \ddot{\theta} \quad \ddot{\beta} \quad \mathbf{w}_{\eta_x} \quad \mathbf{w}_{\eta_{yz}} \quad \mathbf{w}_{\psi_B} \quad \mathbf{w}_L]^T \end{aligned} \quad (17)$$

#### B. Update

Visual measurements of the drogue-mounted IR markers update the state. Rather than using the raw marker centroid coordinates as in Section II-A, an ellipse is fit to the data using the algorithm in [26]. The ellipse centre, semi-major axis, semi-minor axis and rotation form a pseudo sensor measurement. This technique avoids marker model correspondence determination, it is robust to partial occlusion and tolerant to symmetry. These characteristics are particularly important during the approach to docking as shown in Fig. 4, where the drogue is partially occluded and two of the markers return multiple centroids due to the close proximity. The ellipse fitting strategy simply requires at least five markers, located at the circumference of the drogue entry.

The expected ellipse parameters are determined by calculating five evenly spaced points in a circle of radius  $r_a$  about the origin, translating by the estimated cable length and rotating by the estimated cable orientation to generate a leader relative marker model  $\zeta_j^a$  in Eq. (18).

$$t_j = \frac{2\pi}{m}(j-1), \quad j = 1, 2, \dots, 5 \quad (18a)$$

$$\zeta_j^a = C_{(0, \theta_k, \beta_k - \psi_{B_k})}^T \begin{bmatrix} -L_k - d \\ r_a \sin t_j \\ r_a \cos t_j \end{bmatrix} + P_{\text{cable}} \quad (18b)$$

The marker model can now be transformed to the camera's pixel frame using Eqs. (8) and (9), taking care to only transform the leader by  $\psi_l$  only, rather than the full attitude. Finally, an ellipse is fit to the marker pixel coordinates using the algorithm in [9]. The result is the vector of expected ellipse parameters in Eq. (19).

$$\mathbf{h}^{\text{ellipse}}[\hat{\mathbf{x}}_{a_k}, k] = [\mu \quad a \quad b \quad \varphi]^T \quad (19)$$

where  $\mu$  is the centre of the ellipse,  $a$  and  $b$  are the semi-major and semi-minor axes and  $\varphi$  is the rotation. Care should

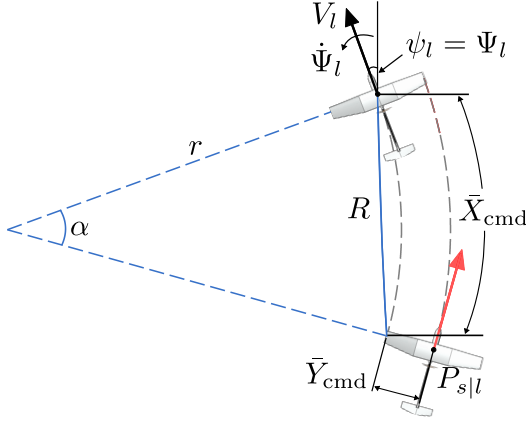


Fig. 5: The setpoint state to compensate for a leader turn.

be taken to ensure  $\zeta_z^c > 0$  in Eq. (9). This problem occurred during the first docking where the marker and focal planes became collinear, yielding  $\zeta_z^c \approx 0$  and diverging the filter.

#### IV. RELATIVE GUIDANCE

Once an estimate of the leader relative state and appendage motion is known, actions must be undertaken to transition from the initial position, to close formation flight and then to docking. A deterministic relative guidance algorithm has been developed for this purpose, where the inputs are individual vehicle, relative and appendage state estimates and the output are low level controller commands. Within this algorithm is a formation setpoint augmentation algorithm that compensates for leader dynamics; a terminal setpoint algorithm that guides the follower into the drogue; and a guidance algorithm that minimises the error between the setpoint state and the estimated state.

##### A. Formation Setpoint

The commanded formation configuration  $P_{cmd} = [X_{cmd} \ Y_{cmd} \ Z_{cmd}]^T$ , is a position in the leader's horizontal body frame and may represent a loose formation or close formation configuration. This section augments  $P_{cmd}$  to compensate for leader turn dynamics. The output is a desired or setpoint position, relative to the leader  $P_{s|l}$  and a setpoint velocity  $V_s$ , that shall be maintained at  $P_{s|l}$ . First,  $P_{cmd}$  is rotated to be in the direction of  $V_l$ , i.e.  $\bar{P}_{cmd} = C_{\psi_l - \Psi_l}^T P_{cmd}$ . During straight and level flight, the rotation rate of the leader's inertial velocity  $\dot{\Psi}_l$ , is zero so  $P_{s|l} = \bar{P}_{cmd}$  and  $V_s = V_l$ . However, when  $\dot{\Psi}_l \neq 0$ ,  $P_{s|l}$  lies on an arc as shown in Fig. 5. The direction of  $V_s$  is the tangent to the arc at  $P_{s|l}$  and the magnitude of  $V_s$  changes when  $\dot{Y}_{cmd} \neq 0$ . Rather than differentiating  $\Psi_l$  to obtain  $\dot{\Psi}_l$ , an analytical solution is provided in Eq. (20). Here,  $\dot{\psi}_l$  is calculated using the leader gyro measurements and a transformation in Section 3.2 of [1].

$$\dot{\Psi}_l = \dot{\psi}_l \frac{V_{a_l} \bullet V_l}{V_l \bullet V_l} \quad (20)$$

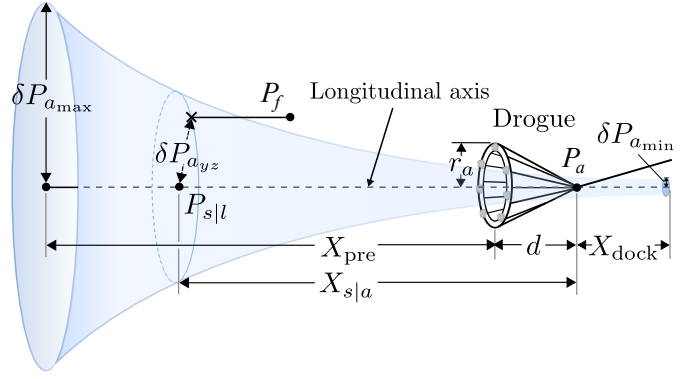


Fig. 6: The 4th-order terminal setpoint algorithm.

The radius  $r$ , central angle  $\alpha$  and chord length  $R$  that form the arc are then calculated.

$$r = \frac{\|V_l\|}{\dot{\Psi}_l} \quad (21a) \quad \alpha = \frac{\bar{X}_{cmd}}{r} \quad (21b) \quad R = 2r \sin \frac{\alpha}{2} \quad (21c)$$

The setpoint position  $P_{s|l}$  can now be calculated by rotating the chord  $R$  by  $\alpha/2$ , adding the lateral component of  $P_{s|l}$  rotated by  $\alpha$ , then rotating the result by  $\Psi_l$ .

$$P_{s|l} = C_{\Psi_l} \left( C_{\alpha/2} [R \ 0 \ \bar{Z}_{cmd}]^T + C_\alpha [0 \ \bar{Y}_{cmd} \ 0]^T \right) \quad (22)$$

Exploiting the fact that the rotation rate of  $V_s$  and  $V_l$  are equal, and combining with Eq. (21a) we obtain an expression for the magnitude of  $V_s$ .

$$\dot{\Psi}_l = \frac{\|V_s\|}{r - \bar{Y}_{cmd}} = \frac{\|V_l\|}{r} \quad (23a)$$

$$\|V_s\| = \|V_l\| - \bar{Y}_{cmd} \dot{\Psi}_l \quad (23b)$$

The setpoint velocity is then calculated by rotating the unit vector of  $V_l$  by  $\alpha$  and scaling by the magnitude of  $V_s$ .

$$V_s = \|V_s\| C_\alpha \bar{V}_l \quad (24)$$

##### B. Terminal Setpoint

The terminal approach from the pre-contact formation position to docking occurs by commanding a setpoint position that pulls the nose of the follower through the entrance to the drogue. The algorithm is depicted in Fig. 6 where the setpoint position  $P_{s|l}$  is situated along the longitudinal axis of the drogue, passing through  $P_a$ . The position along the longitudinal axis is a function of the scalar lateral-vertical position error  $\delta P_{a_{yz}}$ , between the nose of the follower  $P_N$ , and  $P_a$  as per Eq. (25). The error is saturated to the range  $[\delta P_{a_{min}} \ \delta P_{a_{max}}]$  which corresponds to the radii of the smaller and large funnel entries.

$$\delta P_a = C_{\psi_l} \left( P_{l|f} - C_{\bar{q}_f}^T P_N \right) + P_a + P_{cable} \quad (25a)$$

$$\delta P_{a_{yz}} = \sqrt{\delta Y_a^2 + \delta Z_a^2}, \{ \delta P_{a_{min}} < \delta P_{a_{yz}} < \delta P_{a_{max}} \} \quad (25b)$$



When the error is small, the nose is lined up with the entrance to the drogue so  $P_{s|l}$  is situated further through the drogue entrance. The fourth-order function to calculate the longitudinal position  $X_{s|a}$ , relative to  $P_a$  is provided in Eq.(26). The final setpoint position is calculated in Eq. (27).

$$X_{s|a} = X_{\text{dock}} - \left[ \frac{(\delta P_{a_{yz}} - \delta P_{a_{\min}})(X_{\text{pre}} + d + X_{\text{dock}})^4}{\delta P_{a_{\max}}} \right]^{\frac{1}{4}} \quad (26)$$

$$P_{s|l} = P_a + P_{\text{cable}} + \begin{bmatrix} 0 & 0 & X_{s|a} \end{bmatrix}^T \quad (27)$$

Pictorially,  $P_{s|l}$  lies on the drogue's longitudinal axis, in the same lateral-vertical plane as the intersection of the funnel surface with a line that passes through  $P_f$ , parallel to the longitudinal axis. The position along the axis is saturated by  $X_{\text{pre}}$  and  $X_{\text{dock}}$ , which are the pre-contact standoff distance and a docked offset, that assures connection is made.

To compensate for the high dynamics,  $\delta P_a$  can also be predicted, based on the estimated time to contact, which is computed using  $\delta P_a$  and  $\delta \dot{P}_a$ . The predicted  $\delta P_a$  is then calculated by propagating  $P_a$  by the time to contact using Eqs. (12)-(17); adding  $\delta P_a$ ; and adding a first-order integration of the relative velocity.

$$\delta \dot{P}_a = C_{\psi_l} V_{l|f} + \begin{bmatrix} 0 \\ \dot{\theta} \\ \dot{\beta} \end{bmatrix}^T \times P_a + \begin{bmatrix} 0 \\ 0 \\ \dot{\psi}_l \end{bmatrix}^T \times (P_a + P_{\text{cable}}) \quad (28)$$

### C. Guidance

Deterministic formation guidance was developed to minimise the error between the setpoint state  $P_{s|l}$ , and the estimated relative state  $P_{l|f}$  that was determined in Section II. As depicted in Fig. 7, the fundamental strategy is to define a commanded velocity vector  $V_{a_{\text{vec}}}$ , that incorporates a closing velocity to minimise the position error and the setpoint velocity that we wish to maintain when we reach the desired formation configuration. The commanded rotation rate of the commanded velocity vector is also defined, to maintain the formation configuration during turns. Low level commands are formulated to align the magnitude and direction of the vehicle's air-relative velocity vector with the commanded air-relative velocity vector. The algorithm could also be formulated in terms of inertial velocity, however this would introduce latency in the feedback. In the context of fixed-wing UAVs, the low level commands are bank angle,  $\phi_{\text{cmd}}$ , airspeed  $V_{a_{\text{cmd}}}$  and vertical velocity  $V_{Z_{\text{cmd}}}$ . First, the position error  $P_{s|f}$  is calculated and used to define a closing velocity  $V_{\text{close}}$ , using separate horizontal and vertical gains  $K_{XY}$  and  $K_Z$  respectively. The operator  $\circ$  denotes element-wise multiplication.

$$P_{s|f} = P_{s|l} + P_{l|f} \quad (29)$$

$$V_{\text{close}} = P_{s|f} \circ [K_{XY} \quad K_{XY} \quad K_Z]^T \quad (30)$$

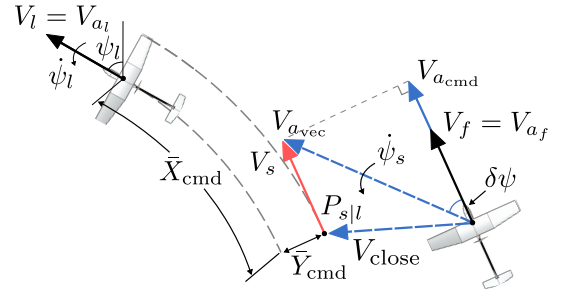


Fig. 7: The formation guidance strategy used to transition to and maintain a commanded formation configuration.

The commanded air-relative velocity vector  $V_{a_{\text{vec}}}$  is calculated in Eq. (31) by combining the closing velocity, the setpoint velocity and the estimated relative velocity, then correcting for wind  $V_w$  where  $V_w = V_f - V_{a_f}$ . It is worth noting that wind is estimated onboard each UAV and used to infer the vehicles airspeed. This is important because the squared relationship between airspeed and dynamic pressure lowers the resolution of the directly measured airspeed. The commanded vertical velocity  $V_{Z_{\text{cmd}}}$ , is the vertical component of  $V_{a_{\text{vec}}}$ .

$$V_{a_{\text{vec}}} = V_{\text{close}} + V_s + V_{l|f} - V_w \quad (31)$$

The commanded airspeed,  $V_{a_{\text{cmd}}}$  is then the scalar projection of  $V_{a_{\text{vec}}}$  in the direction of  $V_{a_f}$ .

$$V_{a_{\text{cmd}}} = V_{a_{\text{vec}}} \bullet \vec{V}_{a_f} \quad (32)$$

The lateral guidance aims to minimise the angle  $|\delta\psi|$ , between the X-Y component of the desired air-relative velocity vector  $V_{a_{\text{vec}_{XY}}}$  and the vehicle's current air-relative velocity vector  $V_{a_{f_{XY}}}$ .

$$\delta\psi = \tan^{-1} \left( \frac{V_{a_{\text{vec}_{XY}}} \times V_{a_{f_{XY}}}}{V_{a_{\text{vec}_{XY}}} \bullet V_{a_{f_{XY}}}} \right) \quad (33)$$

First, the rotation rate  $\dot{\psi}_s$ , of  $V_{a_{\text{vec}}}$  is calculated in Eq. (34), in a similar way to Eq. (20).

$$\dot{\psi}_s = \dot{\psi}_l \frac{(V_s - V_w) \bullet V_{a_{\text{vec}}}}{V_{a_{\text{vec}}} \bullet V_{a_{\text{vec}}}} \quad (34)$$

Then a commanded heading rate  $\dot{\psi}_{\text{cmd}}$ , is computed using a lateral gain  $K_\psi$ .

$$\dot{\psi}_{\text{cmd}} = \delta\psi K_\psi + \dot{\psi}_s \quad (35)$$

The bank angle command  $\phi_{\text{cmd}}$  is then calculated using the commanded heading rate, follower airspeed and gravity  $g$ .

$$\phi_{\text{cmd}} = \tan^{-1} \left( \frac{\|V_{a_f}\| \dot{\psi}_{\text{cmd}}}{g} \right) \quad (36)$$

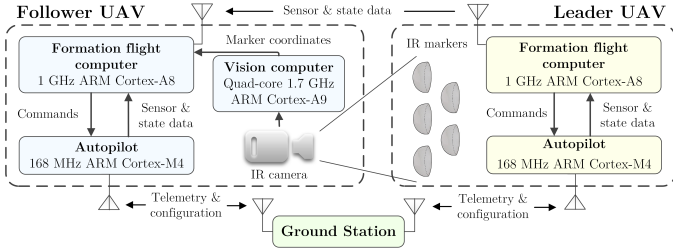


Fig. 8: The dual-UAV system architecture.

## V. AIRBORNE DOCKING EXPERIMENTS

Sustained close formation flight and repeatable docking was demonstrated in a series of flight tests. This section describes the dual-UAV system that was used in the experiments, then presents results from the tests. To the best of the author's knowledge, these are the closest autonomous formation flights to date and the first demonstration of UAV airborne docking.

### A. Dual-UAV Testbed

The dual-UAV system in Fig. 8 consisted of two fixed-wing aircraft, each equipped with a purpose built autopilot to handle low level guidance, navigation and control. A separate formation flight computer (FFC) implemented the relative navigation, appendage estimation and relative guidance algorithms. Sensor and state data was exchanged wirelessly between the FFCs and autopilot-to-ground communications allowed telemetry downlink and configuration upload.

The vision system consisted of a follower mounted camera with a narrow bandpass filter; a quad-core ARM computer to process the images; five IR LED markers positioned at the extremities of the leader, and seven IR markers mounted around the circumference of the drogue, as shown in Fig. 2. Camera exposure, brightness and contrast were tuned in worst-case environmental IR conditions to optimise interference rejection and marker sensitivity.

Marker centroid extraction was implemented by converting the image to grayscale; thresholding; calculating image contours; approximating these with a polynomial; then returning the centre of the minimum enclosing rectangles. The vision system is able to process 720p frames at 20 FPS with computational capacity to spare.

### B. Flight Test Results

During the flights, the leader was commanded to track a  $760 \times 160$  m racetrack pattern, at an altitude of 80 m AGL and a constant airspeed of 16 m/s, while transmitting its current state and sensor data to the follower. Leader autonomy was not necessary but ensured a consistent and predictable path. When initiated, the follower was commanded to a) rendezvous to  $P_{s|l} = [-10 \ 0 \ 1]^T$  and obtain vision measurements b) transition to a pre-contact position  $X_{pre} = 3$  m to observe the drogue, then c) dock with the 0.17 m radius drogue.

Figure 9a shows relative position data from a close formation flight and compares the vision-aided estimate with an unaided estimate, through three circuits of a racetrack pattern.

Assuming the vision-aided estimate is close to the truth, errors in the unaided estimate are often in excess of 5 m which motivates the use of vision. During the two visual dropouts at 140 s and 220 s, the horizontal component of the aided estimate gracefully degraded to the accuracy of the unaided estimate. The vertical component maintained a constant offset which was equal to  $h_B$ . The dropouts also show a retreat from close formation to avoid a collision and to regain vision.

Docking was demonstrated repeatedly in multiple flights, including five successful docking manoeuvres in less than seven minutes of a single flight. This seven minute segment is shown in Fig. 9b, where D1-5 indicate each docking and C1-2 indicate instances where the drogue made brief contact with follower, but docking did not occur. Due to the scale of the aircraft, the drag created by the drogue represents a large portion of the leader's total drag. When docking occurs, the extra drag is instantaneously eliminated, so the leader experiences a longitudinal and vertical acceleration, until the controllers adjust. This is shown by sharp spikes in the leader's altitude and airspeed in Fig. 9b. The docking indicators are further magnified by passing the 3D airspeed through a high-pass filter and squaring the result to obtain the trace at the bottom of Fig. 9b.

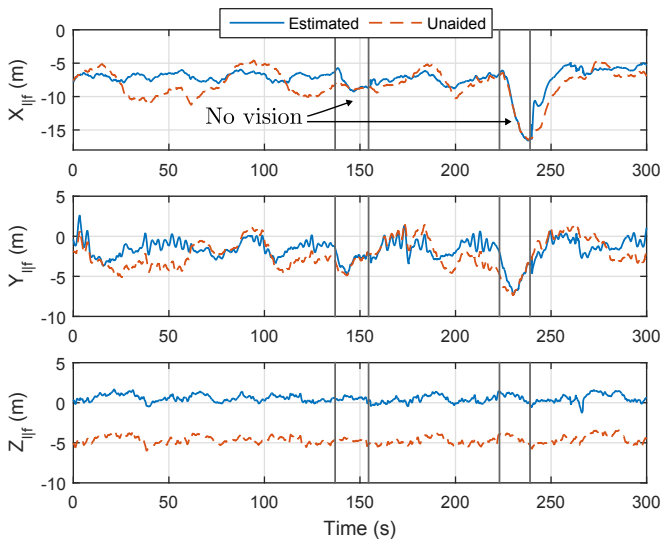
The terminal position error  $\delta P_a$  between the drogue and the follower's nose is shown for each approach to docking in Fig. 9c. Docking occurs when  $\delta X_a < 0$ ,  $|\delta Y_a| < r_a$  and  $|\delta Z_a| < r_a$ . The first two seconds of connection are indicated by the shaded area. Once docked,  $\delta Z_a$  has a characteristic decrease due to the instantaneous reduction of leader drag. Time on the x-axis is simply shown for scale. An image sequence of the approach to docking, then station keeping, taken from the left wing of the follower, is shown in Fig. 10.

Throughout the flight, in Fig. 9b,  $X_{dock}$  was increased so that the terminal guidance was more aggressive. This resulted in more frequent docking attempts but with a lower likelihood of success, where each attempt was a transition from pre-contact to a dock attempt. For comparison, D1-2 were successful on the first attempt, but each took 29 s and 48 s respectively. D3-5 were successful on the 5th, 8th and 4th attempt, respectively, with an average time of 17 s.

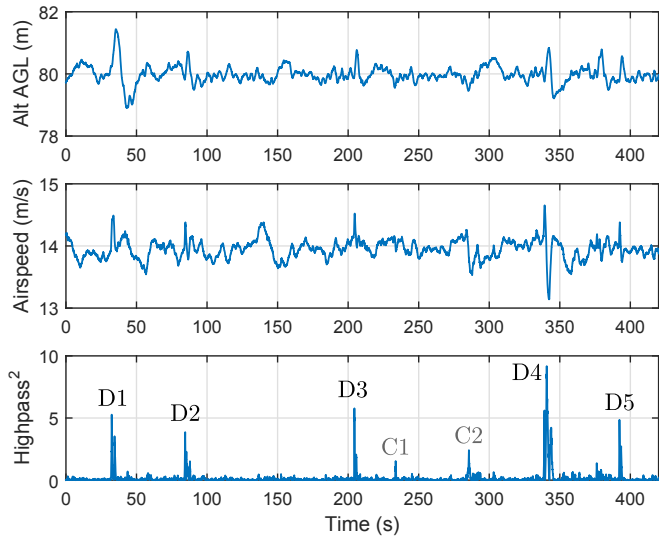
## VI. CONCLUSION AND FUTURE WORK

This paper has presented a complete solution to autonomous airborne docking, from theoretical inception, to practical demonstration. The vision-aided guidance and navigation algorithms contain principled solutions to problems that are only encountered during actual implementation. An illuminating example is the minor IMU alignment, which was found to cause a consistent lateral position bias between the follower and the drogue. Although the bias was only  $\approx 0.5$  m, it would have continued to prevent docking, if it were not estimated within the filter.

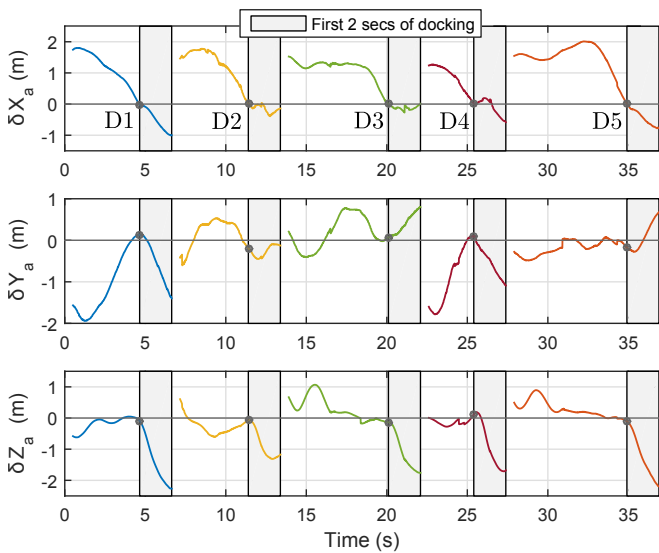
As the first results of their kind, this work contributes a verified starting point for future research. This may include the development of a docking appendage for small UAV refueling and recharging; and application to UAV aerial capture.



(a) The estimated relative position during close formation flight when aided by vision in blue, and unaided in dashed red.



(b) The leader's altitude, airspeed and the square of the high-pass filtered airspeed, which is used to detect docking.



(c) The terminal position error  $\delta P_a$ , for five approaches to docking.

Fig. 9: Results from close formation flight and docking.



Fig. 10: The approach to docking, then station keeping, taken from the left wing of the follower.



## REFERENCES

- [1] R. W. Beard and T. W. McLain. *Small Unmanned Aircraft: Theory and Practice*. Princeton University Press, 2012.
- [2] A. Betser, P. Vela, and A. Tannenbaum. Automatic tracking of flying vehicles using geodesic snakes and Kalman filtering. In *Decision and Control, 43rd IEEE Conference on*, December 2004.
- [3] J. D. Boskovic, S. Li, and R. K. Mehra. Semi-globally stable formation flight control design in three dimensions. In *Decision and Control, Proceedings of the 40th IEEE Conference on*, 2001.
- [4] G. Campa, M. L. Fravolini, A. Ficola, M. Napolitano, B. Seanor, and M. Perhinschi. Autonomous Aerial Refueling for UAVs Using a Combined GPS-Machine Vision Guidance. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*. American Institute of Aeronautics and Astronautics, 2004.
- [5] J. L. Crassidis and F. L. Markley. Unscented Filtering for Spacecraft Attitude Estimation. *Journal of Guidance, Control, and Dynamics*, 26(4):536–542, 2003.
- [6] M. Darling. Autonomous Close Formation Flight of Small UAVs Using Vision-Based Localization. Master’s thesis, California Polytechnic State University, 2014.
- [7] DARPA. Making connections at 45,000 feet: future UAVs may fuel up in flight, October 2012. [Online; accessed 21-July-2014].
- [8] R. P. Dibley and M. J. Allen. Autonomous Airborne Refueling Demonstration Phase I Flight-Test Results. In *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, 2007.
- [9] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least-squares fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.
- [10] A. Fosbury and J. Crassidis. Relative Navigation of Air Vehicles. *Journal of Guidance, Control, and Dynamics*, 31(4):824–834, 2008.
- [11] F. Giulietti, L. Pollini, and M. Innocenti. Autonomous formation flight. *Control Systems Magazine, IEEE*, 20(6):34–44, 2000.
- [12] E. N. Johnson, A. J. Calise, Y. Watanabe, J. Ha, and J. C. Neidhoefer. Real-Time Vision-Based Relative Aircraft Navigation. *Journal of Aerospace Computing, Information, and Communication*, 4(4):707–738, 2007.
- [13] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte. A new approach for filtering nonlinear systems. In *American Control Conference, Proceedings of the*, 1995.
- [14] M. Mammarella, G. Campa, M. R. Napolitano, M. L. Fravolini, Y. Gu, and M. G. Perhinschi. Machine Vision/GPS Integration Using EKF for the UAV Aerial Refueling Problem. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(6):791–801, 2008.
- [15] C. Martinez, T. Richardson, P. Thomas, L. B. Jonathan, and P. Campoy. A vision-based strategy for autonomous aerial refueling tasks. *Robotics and Autonomous Systems*, 61(8):876–895, 2013.
- [16] R. Mati, L. Pollini, A. Lunghi, M. Innocenti, and G. Campa. Vision-based autonomous probe and drogue aerial refueling. In *Control and Automation, 14th Mediterranean Conference on*. IEEE, 2006.
- [17] R. K. Nangia. Operations and aircraft design towards greener civil aviation using air-to-air refuelling. *The Aeronautical Journal*, 110(1113):705–721, 2006.
- [18] J. W. Nichols, L. Sun, R. W. Beard, and T. McLain. Aerial Rendezvous of Small Unmanned Aircraft Using a Passive Towed Cable System. *Journal of Guidance, Control, and Dynamics*, 37(4):1131–1142, 2014.
- [19] D. Scott, M. Toal, and J. Dale. Vision-based sensing for autonomous in-flight refueling. In *Unmanned Systems Technology IX*, May 2007.
- [20] M. J. Tahk, C. S. Park, and C. K. Ryoo. Line-of-Sight Guidance Laws for Formation Flight. *Journal of Guidance, Control, and Dynamics*, 28(4):708–716, 2005.
- [21] P. R. Thomas, U. Bhandari, S. Bullock, T. S. Richardson, and J. L. du Bois. Advances in air to air refuelling. *Progress in Aerospace Sciences*, 71(0):14 – 35, 2014.
- [22] J. Valasek, K. Gunnam, J. Kimmet, M. D. Tandale, J. L. Junkins, and D. Hughes. Vision-based sensor and navigation system for autonomous air refueling. *Journal of Guidance, Control, and Dynamics*, 28(5):979–989, 2005.
- [23] Y. Watanabe. *Stochastically optimized monocular vision-based navigation and guidance*. PhD thesis, Georgia Institute of Technology, 2008.
- [24] W. R. Williamson, G. J. Glenn, V. T. Dang, J. L. Speyer, S. M. Stecko, and J. M. Takacs. Sensor Fusion Applied to Autonomous Aerial Refueling. *Journal of Guidance, Control, and Dynamics*, 32(1):262–275, 2009.
- [25] D. B. Wilson, M. A. T. Soto, A. H. Göktoğan, and S. Sukkarieh. Real-time rendezvous point selection for a nonholonomic vehicle. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 3941–3946, May 2013.
- [26] D. B. Wilson, A. H. Göktoğan, and S. Sukkarieh. Drogue Motion Estimation Using Air-to-Air Observations. In *Australasian Conference on Robotics and Automation (ACRA)*, 2014.
- [27] D. B. Wilson, A. H. Göktoğan, and S. Sukkarieh. Experimental Validation of a Drogue Estimation Algorithm for Autonomous Aerial Refueling. In *Robotics and Automation (ICRA), IEEE International Conference on*, pages 4988–4995, May 2015.
- [28] Y. Wu, D. Hu, M. Wu, and X. Hu. Unscented Kalman Filtering for Additive Noise Case: Augmented vs. Non-augmented. In *American Control Conference, proceedings of*, 2005.
- [29] Y. Yin, D. Xu, X. Wang, and M. Bai. Detection and Tracking Strategies for Autonomous Aerial Refuelling Tasks Based on Monocular Vision. *International Journal of Advanced Robot Systems*, 97(11), 2014.