

N. R. J. Lawrance

(Australian Centre for Field Robotics,
University of Sydney)

J. J. Acevedo

(University of Sevilla)

J. J. Chung, J. L. Nguyen,

D. Wilson, S. Sukkarieh

(Australian Centre for Field Robotics,
University of Sydney)

E-mail : n.lawrance@acfr.usyd.edu.au

DOI : 10.12762/2014.AL08-05

Long Endurance Autonomous Flight for Unmanned Aerial Vehicles

This paper presents a summary of research performed at the University of Sydney towards extending the flight duration of fixed-wing unmanned aerial vehicles. A historical context to extended flight is provided and particular attention is paid to research in autonomous soaring and aerial refueling. Autonomous soaring presents a unique set of challenges whereby an aircraft must autonomously identify sources of energy in the wind field and generate trajectories to exploit these conditions to collect energy. The basic mechanisms of soaring flight are examined and methods for generating energy gaining trajectories for exploration, information gathering and patrolling missions with multiple aircraft are detailed. Aerial refueling represents a complementary approach for extending flight duration, and the challenges and current efforts towards autonomous refueling between small aircraft are also detailed.

Introduction

Whilst Unmanned Aerial Vehicles (UAVs) have become increasingly capable platforms used in a wide variety of applications, most are still limited in their endurance by the necessity for on-board energy storage for propulsion. However, recent research has aimed to address this issue by examining methods for extending flight duration by collecting energy during flight. There are two distinct approaches to this problem. The first is the direct and ongoing capture of energy from the vehicle's surrounding environment, such as soaring in wind or collecting solar energy using solar panels. The second is the deliberate resupply of energy from other vehicles using aerial refueling or recharging.

Soaring

Soaring is the process of exploiting wind to collect energy. Soaring was discovered when some birds were noted to be capable of flying for extended periods of time without flapping their wings and seemingly without losing airspeed or altitude. Early aerodynamic research had shown that energy must be lost to drag by any object moving through a fluid. Thus, it was determined that since the birds were not directly expending energy they must be capturing energy from the wind [49, 52, 2]. This process is known as soaring and there are two primary methods for energy capture in wind, static soaring and

dynamic soaring. Static soaring is the process of flying through air that is rising relative to the ground. This method is utilized by both birds and manned gliders where there are naturally occurring sources of rising air (thermals) [60, 50]. Thermals occur when an area of the ground is heated (usually by the sun) to a warmer temperature than surrounding areas. The warm air is less dense and rises with respect to the cooler surrounding air, and an aircraft that flies in the rising air will collect energy (illustrated in figure 1). Thermals are favoured by both birds and human glider pilots because they are relatively common and easy to utilise for energy gain [51, 57].

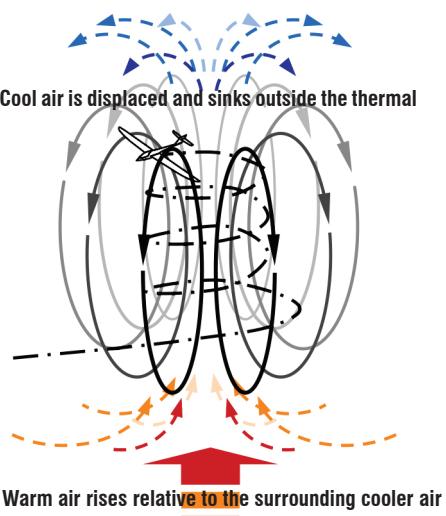


Figure 1 - Static soaring in a thermal

Dynamic soaring utilises trajectories through distributions of wind speed (wind shear) to obtain increased kinetic energy (see figure 2). This is often performed cyclically, with the energy gained in each cycle being used to travel before starting the next cycle. This method was originally discovered being used by birds such as albatrosses over the ocean [24, 62, 69]. Dynamic soaring generally requires good knowledge of the wind field to calculate trajectories which result in energy gain. Naturally occurring sources of wind shear are boundary layers which occur over surfaces (such as the ground or ocean), shear generated by flow around geographic obstacles, and meteorological shear.

Early research in soaring focused on how birds identified and used sources of energy, the amount of energy that could be obtained, and how they integrated soaring behaviours with their needs for travelling and foraging flight [63, 55]. Parallel research for manned aircraft has focused mainly on static soaring which is the primary energy capture method for manned gliders. This led to the development of relatively simple algorithms such as the speed-to-fly rules [39, 40] for cross country gliding. These rules are used to determine when a vehicle should utilise a thermal and when to travel to maximise overall average speed based on an estimate of thermal strengths.

Whilst the bird and manned glider problems have received significant attention, it is only in recent years that soaring with UAVs has been addressed as a research problem. The earliest research attempted to imitate the behavior of manned glider pilots by utilising simple gliding rules (such as estimating the best orbit radius for a thermal) for cross-country flights in autonomous UAVs [4, 5, 3]. Extensions to this work culminated in the demonstration of a fully autonomous glider which flew for over 4 hours on a 97 km round trip flight unofficially setting a new soaring record [18]. Dynamic soaring has also been considered but due to the difficulty of testing has mainly been limited to simulation. Previous work used off-line numerical optimization techniques to calculate the wind strengths required for feasible dynamic soaring and showed that shear layers over the ocean should contain sufficient energy to provide continuous or assisted flight for small (< 10 kg) UAVs [70, 71, 37, 17]. Further work examined on-line reactive strategies for soaring in shear [8, 7, 32] and turbulent fields [29, 15]. Of further interest is the application of machine learning to the soaring problem. Reinforcement learning seems like a natural choice for this type of problem due to the inherent goal of balancing exploration and exploitation, but previous implementations suffered from issues of state space complexity and slow learning rates [64, 65, 28].

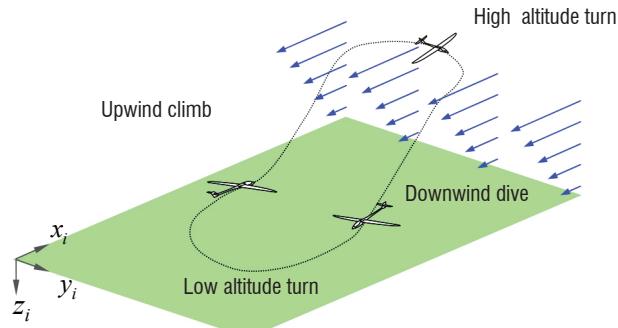


Figure 2 - Dynamic soaring in wind shear

Autonomous aerial refueling

An alternative to capturing energy from the environment is the deliberate resupply of an aerial vehicle during flight. Aerial refueling has been used to extend the mission duration of manned aircraft for decades. This has been almost exclusively within the military space where two strategies are employed. In probe and drogue refueling, a drogue is unwound behind a tanker and a pilot navigates a receiver's probe to contact. Boom and receptacle refueling requires a pilot to station-keep relative to a tanker aircraft while a human operator manually navigates an aerodynamically controlled rigid boom to contact with the receiver's receptacle. Due to the close proximity operations required by these methods, the procedure is inherently dangerous and requires significant pilot training, practice and skill.

With the recent surge in the number of operational UAVs, automating this procedure is the natural next step and is not limited to the military space, or even large UAVs. To date, completely autonomous aerial refueling (AAR) between two UAVs has not been demonstrated. Although an AAR procedure comprises many phases, the barrier to success is generally accepted as being the sensing and navigation challenges surrounding tight formation flight. Here, separation is defined as being less than one wingspan [56] and can translate to only a few meters in some cases.

Work on tight formation relative navigation has focused on utilising air-to-air relative observations to obtain an accurate and timely relative state estimate. Vision-only techniques have included active visual contours [16], silhouette based techniques [27], template matching [44] and feature extraction [41, 61]. Other approaches have fused the vision measurements with inertial and GPS sensors from one or more aircraft in loosely-coupled [43, 21] and tightly coupled arrangements [47, 20, 66]. Experimental validation has mostly occurred in simulations of varying fidelity with notable exceptions being rendezvous experiments [48], GPS-based loose formation flights [22] and the first closed-loop, vision-only loose formation flight [25].

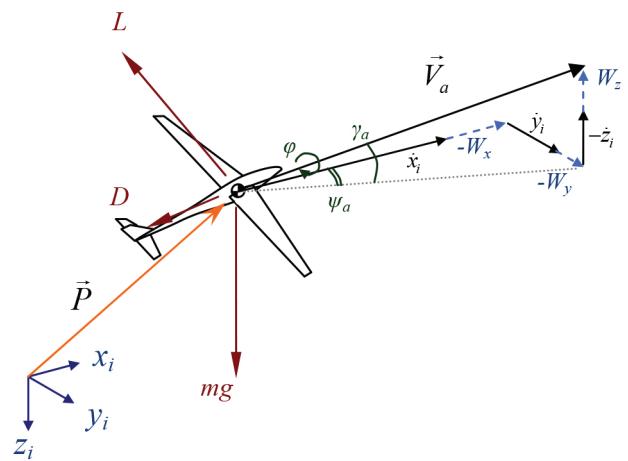


Figure 3 - Air-relative velocity and applied forces for a gliding aircraft. The air relative quantities represent the motion of the vehicle with respect to a stationary air frame, which is actually moving through inertial space due to the wind.

Soaring

Dynamics of soaring flight

The conditions required for energy-gain flight can be identified by analysing the equations of motion of an aircraft in wind. This section briefly describes a dynamic model for a soaring aircraft and examines the mechanisms of soaring through analysis of the resulting energy equations.

Dynamic model

The dynamic model used in the analysis and simulation for this work is an aerodynamic point mass model. The applied forces are the aerodynamic force (decomposed into lift, L, and drag, D) and the weight force (mg). The aerodynamic force is a function of the motion of the vehicle relative to the surrounding air and the physical properties of the aircraft (shape, size, surface properties). Body force due to sideslip is not considered. Weight is the force due to gravity and is directed down (a flat Earth model is assumed due to the relatively small scale of the aircraft and flight paths). Thus there are two important frames of reference: an inertial frame which is fixed with respect to the ground, and the air-relative frame which is the aircraft motion relative to the surrounding air. Figure 3 illustrates the forces acting on a gliding aircraft in wind.

Wind is defined in inertial space and represents the motion of the air relative to the ground-fixed inertial frame. The air-relative velocity vector represents the motion of the vehicle with respect to the surrounding air by treating the wind field as a stationary frame. Thus the airspeed is the magnitude, the air-relative climb γ_a is the vertical angle and ψ_a is the heading angle of this air-relative motion vector. The bank angle ϕ is the rotation of the lift vector around the velocity vector. The air-relative to inertial transformation matrix is denoted C_a^i and is made up of the standard rotation transformation matrices such that $C_a^i = L_z(\psi_a)L_y(\gamma_a)L_x(\phi)$.

The air-relative velocity can be described in terms of the airspeed V_a , heading ψ_a and air-relative climb angle γ_a :

$$\vec{V}_a = C_a \begin{bmatrix} V_a \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} V_a \cos \gamma_a \cos \psi_a \\ V_a \cos \gamma_a \sin \psi_a \\ 0 \end{bmatrix} \quad (1)$$

The drag coefficient is estimated using the common approximation where the effective drag coefficient (C_D) is the sum of parasitic ($C_{D,0}$) and lift-induced ($C_{D,i}$) drag components [6]. Induced drag is a function of the lift coefficient C_L , aspect ratio \mathcal{A} and efficiency factor ε .

$$C_D = C_{D,0} + \frac{C_L^2}{\pi \mathcal{A} \varepsilon} \quad (2)$$

Consider the case of locally spatially-fixed linear wind gradients. Let J_w be the spatial wind gradients at a particular location,

$$J_w = \begin{bmatrix} \frac{\partial W_x}{\partial x} & \frac{\partial W_x}{\partial y} & \frac{\partial W_x}{\partial z} \\ \frac{\partial W_y}{\partial x} & \frac{\partial W_y}{\partial y} & \frac{\partial W_y}{\partial z} \\ \frac{\partial W_z}{\partial x} & \frac{\partial W_z}{\partial y} & \frac{\partial W_z}{\partial z} \end{bmatrix} \quad (3)$$

By summing the applied forces (lift, drag, and weight) in the inertial frame and differentiating the velocities for acceleration, a set of dynamic equations can be obtained which describe the motion of a gliding aircraft. Assuming that roll rate ($d\phi/dt$) is directly controlled, the system can be solved by specifying climb angle rate ($d\gamma_a/dt$) or specific lift (L/m) as a control input. Physical limitations mean that the maximum specific lift is limited by maximum lift coefficient ($C_{L,max}$) and load factor constraints (n_{min}, n_{max}). In such cases, the lift is specified and (6) returns the climb angle rate. The resulting equations for the system are shown in (5)–(9), where \dot{P} is the velocity of the vehicle in the inertial frame.

Further information and full derivation of the equations can be found in [33].

$$\vec{V}_w = \begin{bmatrix} \cos \psi_a \sin \gamma_a \\ \sin \psi_a \sin \gamma_a \\ \cos \gamma_a \end{bmatrix}^T J_w \dot{P} \quad (4)$$

$$\frac{L}{m} = \frac{1}{\cos \phi} \left(V_a \frac{d\gamma_a}{dt} + g \cos \gamma_a - \vec{V}_w \right) \quad (5)$$

$$\frac{d\gamma_a}{dt} = \frac{1}{V_a} \left(\frac{L}{m} \cos \phi - g \cos \gamma_a + \vec{V}_w \right) \quad (6)$$

$$\frac{D}{m} = \frac{1}{2} \rho V_a^2 S C_{D,0} + \frac{1}{2} \rho V_a^2 S \pi \varepsilon \quad (7)$$

$$\frac{dV_a}{dt} = \frac{-D}{m} - g \sin \gamma_a - \hat{V}_a J_w \dot{P} \quad (8)$$

$$\frac{d\psi_a}{dt} = \frac{1}{V_a \cos \gamma_a} \left(\frac{L}{m} \sin \phi + [\sin \psi_a, -\cos \psi_a, 0] J_w \dot{P} \right) \quad (9)$$

These equations can be integrated numerically to simulate gliding flight in wind. This model assumes that a controller is able to track a specified roll rate and either pitch rate or lift coefficient. The simulations used in the following sections are based on a remote control RnR SBXC scale glider. Relevant parameters are listed in table 2.

Soaring energy

The dynamic model can be examined in terms of the energy gained or lost to give an understanding of how wind contributes to the overall energy of the platform. The energy of a point mass can be described as the sum of gravitational potential and kinetic energy. We define the air-relative energy E_a as the aerodynamic energy of the vehicle with respect to the surrounding air (treating the air as a local inertial frame with respect to the aircraft). Taking the time derivative of the air-relative energy and substituting the airspeed acceleration (equation (8)) yields the overall specific power.

$$\frac{\dot{E}_a}{m} = -V_a \frac{D}{m} - g W_z - \vec{V}_a^T J_w \vec{V}_w \quad (10)$$

This equation illustrates how a gliding aircraft can gain or lose air-relative energy from a wind field. The first term is the power loss due to drag. This is always an energy loss term since the airspeed must be greater than zero. The second term is the static soaring term representing energy gained or lost from vertical wind. The third term is the dynamic soaring term and represents energy gained or lost due

to wind gradients and is affected by airspeed, climb angle and wind gradients. Equation (10) shows that energy gained or lost from wind shear is proportional to the airspeed. In general terms, increased airspeed increases energy capture or loss rate and allows energy capture from lower magnitude wind gradients.

These equations can be used to identify important conditions for gliding flight, such as the optimal airspeed and climb angle for minimum sink and maximum range. Further, the optimal conditions for energy gain in shear can be calculated when the magnitude of the wind gradient is known. These identities are beyond the scope of this article, but can be found in [31].

Exploration of wind fields

An interesting problem in soaring is that of a vehicle attempting to simultaneously explore and exploit an unknown wind field using soaring. This introduces a requirement for on-line mapping as the vehicle must now also attempt to create a map of the field whilst using that map to generate feasible and energy-gaining paths. This task is complicated by the fact that generally the wind cannot be remotely observed by a small UAV with common on-board sensors. Further, wind fields vary spatially and temporally so observations are only locally valid for a limited period of time.

We consider the case of a UAV with the ability to estimate the local wind vector using an air data sensor and an inertial sensor. We assume that the air data sensor measures the speed and direction of airflow relative to the centre of mass of the aircraft, and the inertial sensor measures the acceleration and speed of the vehicle in a fixed inertial frame. Whilst airspeed and inertial measurement units are common on UAVs, air angle measurements are not. One method for estimating air angles is an alpha-beta vane system which records the wind direction around two perpendicular axes each using a lightweight wind vane and potentiometer [23]. An alternative system is a multi-hole pressure sensor which estimates the air angles by calculating the pressure at different orientations to the wind and solving for the wind direction [59]. It is difficult to determine the accuracy required from these types of sensor for accurate wind reconstruction, and this is part of ongoing work. The collection of truth data is also difficult, though some work had been performed towards quantifying the accuracy available from typical sensors available on small UAVs [30].

Gaussian process mapping and control-sampled planning

Our earliest attempt at solving this problem was to use Gaussian Process (GP) regression, a non-parametric regression technique, for wind mapping to avoid the need for explicit wind feature models. The advantage of using a GP to generate the wind map is that the GP provides both the mean estimate and a variance estimate which can be used to identify the uncertain regions of the map. The exploration/exploitation trade-off is managed by weighting a utility function such that when platform energy is low the utility weights towards known areas (low variance in the GP) with high energy gain (exploitation action) and when platform energy is high the utility weights towards unknown areas (exploration action). An overview of the system is provided in figure 4.

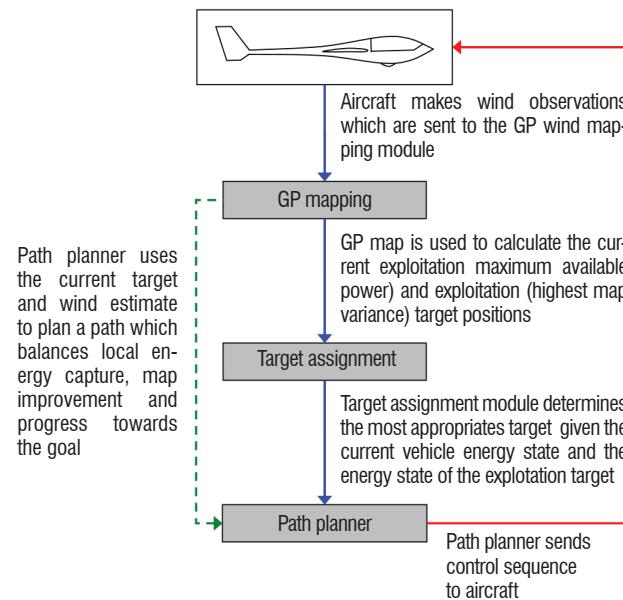


Figure 4 - System overview of simultaneous exploration and exploitation path planning architecture for a gliding UAV [34].

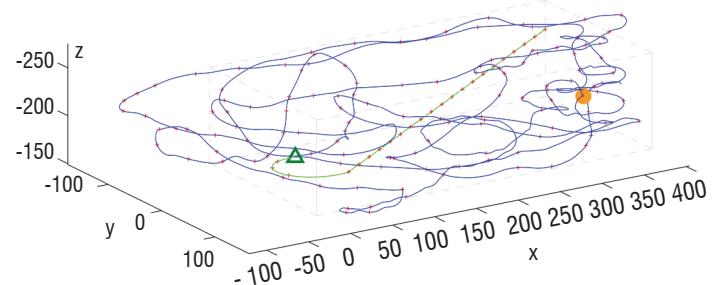


Figure 5 - Single thermal exploration at $t = 500$ s. Autonomous soaring flight starts at the green triangle. There is a single thermal bubble with core vertical wind speed of 3 m/s illustrated by a filled circle [34].

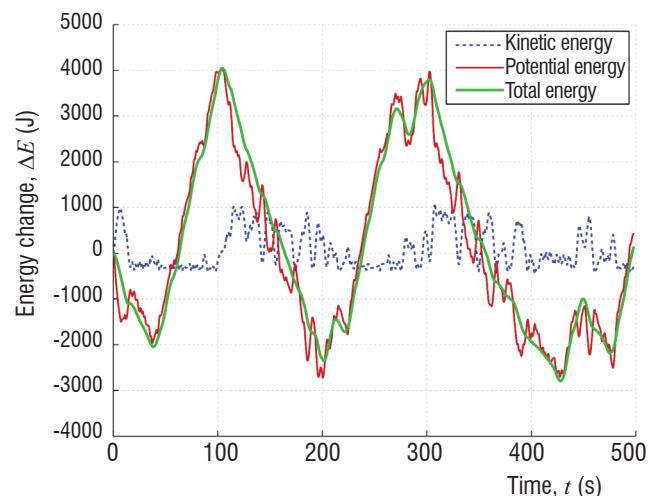


Figure 6 - Energy change during single thermal exploration [34]

Plans are generated by a control-space sampling planner, that selects from a number of control actions (in this case three pitch and three roll actions), forward simulates the resulting trajectories using the mean estimate from the GP wind map for a given search horizon, and selects the resulting path which yields the highest estimated utility. The utility is a weighted sum of the total estimated energy change over the segment, a navigation reward which estimates progress towards a goal, and an exploration reward which ‘optimistically’ estimates the amount of energy available at the projected sample locations, where the ‘optimism’ is a function of the estimated map variance. The current target is selected during each planning cycle based on the energy of the aircraft and is either an exploration target (the map location with maximum estimated variance) or an exploitation target (the position at which maximum power gain was recorded).

Simulation results

The planner showed that it was capable of generating paths that continually improved the map for a specified search region whilst utilising energy sources found during the flight. Figure 5 shows the results of a single simulation where the aircraft is tasked with exploring a three-dimensional rectangular region which contains a single thermal. After approximately 40 s the aircraft finds the thermal and then repeats a process of alternately travelling through the field to reduce variance and using the known thermal to collect energy. The time history of the vehicle energy is shown in figure 6. Further details and additional results are available in [34]. This method showed that a relatively simple control strategy with good knowledge of the vehicle motion model can map and explore a wind field while using energy found during exploration. This was further extended in [35] to account for temporally varying wind fields.

Reinforcement learning for exploration and exploitation management

An alternative or complementary approach to the soaring problem is to pose it as a reinforcement learning (RL) problem where the glider agent must learn the best control action to take given its current state in the wind field. The main advantage of this formulation is that one can elegantly combine the goals of exploring and exploiting the wind field in the RL framework. Further, this does not require explicit identification of soaring behaviours; allowing the system to learn behaviours based on the wind field experienced and resulting energy changes should allow the system to find new soaring behaviours in unknown wind fields.

Reinforcement learning

Standard value-function-based RL control algorithms such as SARSA (λ) [54] learn the value, $Q(s,a)$, of taking a particular action, a , whilst in a particular state, s , by observing the immediate state transition reward, r , and updating the estimated value function according to the following backup equation :

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha \delta_t e_t(s,a) \quad (11)$$

where α is the learning step size, the temporal difference is computed as,

$$\delta_t = r_{t+1} + \Gamma Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t) \quad (12)$$

and the eligibility trace is given by,

$$e_t(s,a) = \begin{cases} 1 & \text{if } (s,a) = (s_t, a_t) \\ \Gamma \lambda e_{t-1}(s,a) & \text{otherwise} \end{cases} \quad (13)$$

for all (s,a) . The two discount factors, Γ and λ , control the contribution of the current reward to the expected return and the value of state-actions previously visited in the trajectory history, respectively. The interested reader is directed to [58] for a full description of SARSA (λ) and other RL algorithms.

The autonomous soaring problem lends itself naturally to this learning framework since it can be considered as a policy learning problem with a well-defined reward (platform energy). Furthermore, the eligibility trace allows credit assignment along the state-action history, promoting the learning of long and potentially complex trajectories, which we expect as a feature of successful soaring policies. However, the state-action space in the soaring problem is continuous, whereas traditional SARSA (λ) deals only with discrete spaces. Thus, value function approximation must be applied to extend SARSA(λ) to handle this. We propose using a GP model to approximate the state-action value function since the GP not only provides an estimate of the mean, it also computes a measure of the uncertainty in the form of a variance. Specifically, the potential reduction in uncertainty (or information gain) from future actions can be used to quantify their exploration utility.

Information measure

The GP variance represents a bounding volume around the estimated function surface and the change in this volume over successive observations can be defined as the information gain of taking those observations [10]. The variance volume can be computed as the integral of the GP covariance function over the state-action space, which has the dimensions $[x_1, \dots, x_n] = [s_1, \dots, s_u, a_1, \dots, a_v]$.

$$V_{\text{bound}_t} = \int_{x_{n_a}}^{x_{n_b}} \dots \int_{x_{1_a}}^{x_{1_b}} \text{cov}([x_1, \dots, x_n] | X_t) dx_1 \dots dx_n \quad (14)$$

$$I_{\text{gain}} = V_{\text{bound}_t} - V_{\text{bound}_{t+1}} \quad (15)$$

The GP training set X_t consists of the observed n -dimensional state-action pairs, furthermore, $X_{t+1} = X_t \cup x_{t+1}$. Given an integrable covariance function, an analytical solution to (14) and subsequently (15) can be found, see [10] for the full solution for the squared exponential covariance function. Drawing inspiration from the eligibility trace for the value function, we define an information value that encapsulates the discounted sum of the information gain from all future state-action observations simulated forward from the next proposed action. Figure 7 gives a graphical representation of the rollout information value calculation :

$$I_{a_{\text{total}}} = I_{\text{gain}_0} + \Gamma_r I_{\text{gain}_1} + \Gamma_r^2 I_{\text{gain}_2} + \dots + \Gamma_r^p I_{\text{gain}_p} \quad (16)$$

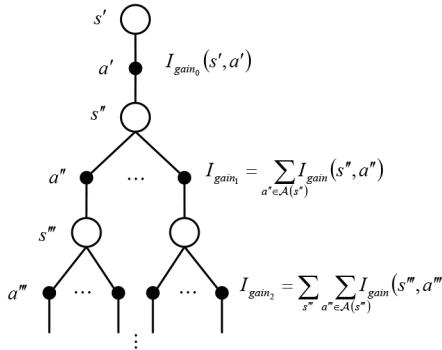


Figure 7 - The rollout method introduced in [10]. The information gain of each rollout level considers all the reachable state-actions at that level. The total information gain is the discounted sum of the information gain of each level.

Action selection

Given the state-action (exploitation) value estimated from the GP model and the information (exploration) value calculated from (16), it is necessary to consolidate the two competing objectives into a single control policy. Prior work [12] has investigated the potential for applying a dynamic scaling of exploration and exploitation metrics according to available platform energy, and this has been adapted in [11] to a 2D simulation of a glider learning to soar in a wind field containing a thermal updraft and a wind shear region. The objective function used normalises both the state-action value and information value and combined the two using a dynamic weighting factor,

$$\begin{aligned} \hat{Q}_t &= \frac{\bar{Q}_t}{\max |\bar{Q}_t|} & \hat{I}_t &= \frac{I_{t_{\text{total}}}}{\max |I_{t_{\text{total}}}|} \\ J_t &= \hat{Q}_t + \max \left(0, \min \left(1, \frac{E_t}{E_{\max}} \right) \right) \hat{I}_t \end{aligned} \quad (17)$$

where E_t is the current platform energy and E_{\max} is the maximum achievable platform energy (corresponding to a maximum speed at a maximum altitude).

Simulation results

The objective function (17) was tested as the sampling policy in the SARSA(λ) learning algorithm with GP value function approximation

for a 3D 6DOF glider simulation with a single thermal updraft in the wind field. It was assumed that the thermal centre location was known and so the learning state dimension could be reduced to $\mathbf{s} = [r_{\text{therm}}, \psi_{\text{therm}}, v_a]$, the relative distance to the thermal centre, the bearing to the thermal centre, and the glider airspeed. The action set dimensions were chosen as $\mathbf{a} = [a_\phi, a_\psi]$, the roll rate and pitch rate. Finally, the rewards were based on the specific energy gain of the platform over each state-action transition, as well as two discrete event costs to cover the stall and crash conditions. The reward function is computed as :

$$\begin{aligned} r_{\text{stall}} &= -25\% \times \frac{1}{2} v_{\text{stall}}^2 & r_{\text{crash}} &= -\frac{2E_{\max}}{m} \\ r_t &= \frac{E_t}{m} + \text{stall} \times r_{\text{stall}} + \text{crash} \times r_{\text{crash}} \end{aligned} \quad (18)$$

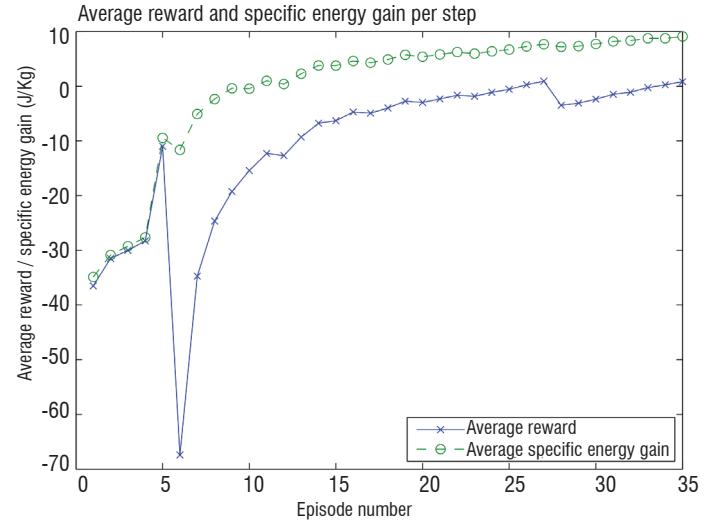


Figure 9 - Progression of the average reward per step and average specific energy gain per step across the learning episodes.

Learning occurred over a period of 35 episodes and a nominal set of flight paths are shown in figure 8. As more observations are taken and the algorithm updates the value function, the flight trajectories rapidly evolve to successfully gain enough energy to exit the field via the upper boundary by episode 14. The energy gaining efficiency of the learnt policy is also seen to increase as shown by the progression of the average reward and average specific energy gain in figure 9.

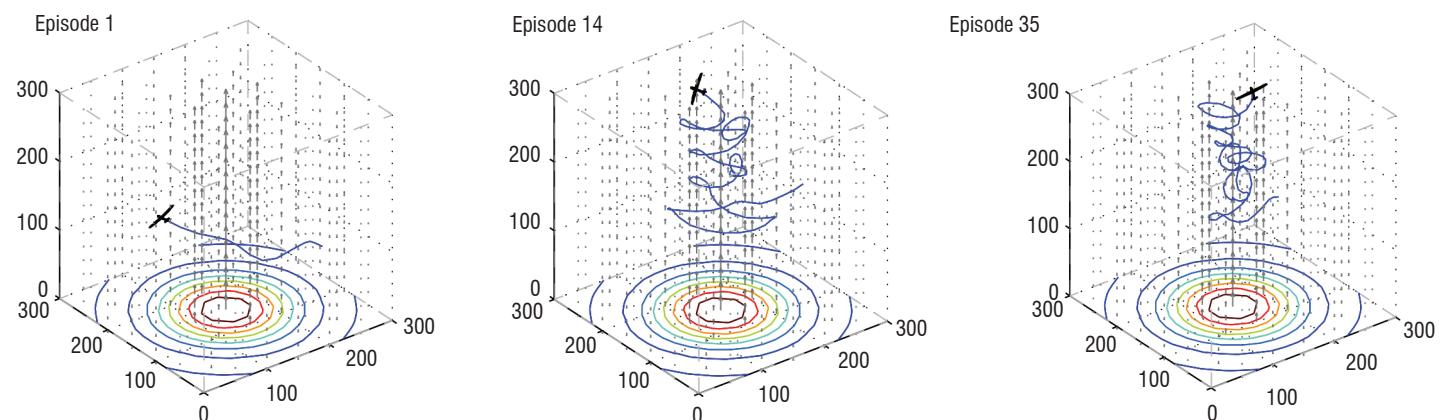


Figure 8 - Evolution of the sampled flight trajectories as learning progresses from episode 1, which terminates when the glider exits the lateral field boundary, to the first instance of the glider exiting via the upper boundary in episode 14, and the final flight trajectory in episode 35.

Soaring for UAV missions

The two previous methods looked at the soaring problem purely as a problem of exploring a wind field and finding the optimal actions to gain energy. However, in most cases there are additional goals for a UAV. The following sections look at methods for how a UAV should manage external goals with the need to capture energy in-flight. Two scenarios are considered : a single glider searching for a lost ground target, and multiple gliders tasked with a long-term surveillance mission over an area of interest.

Long-term information gathering

The first scenario uses thermal static soaring to aid long-term information gathering, as illustrated in figure 10. A gliding UAV equipped with sensors is tasked with searching for a stationary lost ground target. Information of the target state ξ is represented with a probabilistic belief function b^ξ on which UAV control actions are planned to increase information. Given a long-duration mission, the UAV is energy-constrained and must also periodically replenish energy at thermals. It is assumed that all thermals are known, stationary, and cover the search area densely enough for it to be explored. The objective is to find an optimal path plan P that maximises information gain over the entire mission time.

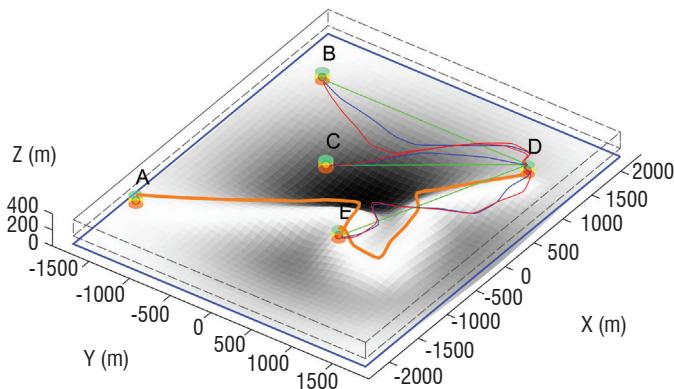


Figure 10 - A long-duration search scenario enabled by thermal soaring. The underlying surface is a probabilistic belief function of a ground target's location; darker regions represent more information. The UAV started at the top of thermal A, and has planned definite path segments to thermal D (orange lines). Green, blue and red lines represent possible next-step trajectories of varying information gain and energy expenditure.

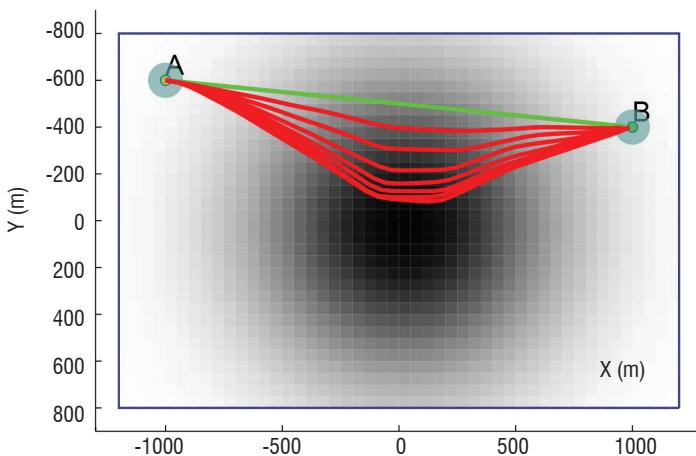


Figure 11 - The set R of inter-thermal paths between thermals A and B, including the initial path r_0 (green). Darker regions represent more information.

Path planning formulation

The target-search problem involves forward propagating UAV actions and applying a sensor model along the resulting paths to capture information. This results in an enormous action search space; to make the problem tractable, we firstly observe that any feasible plan can only consist of (1) information gathering path segments in between thermals, and (2) climbing within thermals to increase energy.

We introduced this as the informative soaring problem in [45] and formulated it as a discrete tree search problem by constructing a tree T of nodes $v \in V$. Starting with an empty root v_0 , a child node v_d at depth d is constructed by appending one inter-thermal path segment to its parent v_{d-1} , so that any v represents a sequence of path segments. Each node has an associated utility $J(v)$ and cost $C(v)$. Here, $J(v)$ is the cumulative probability of target detection defined in [68], and $C(v)$ is the traversal time. A feasible plan is a leaf or terminal node with cost $C(v_t)$ within some budget $B > 0$. An optimal plan is one with maximum $J(v_t) \leq 1$.

Between thermals, there exists a continuous spectrum of inter-thermal path segments varying in $J(\cdot)$ and $C(\cdot)$. However, we only consider a subset Q of up to three options : 1) the maximum utility, 2) the minimum cost, and 3) the median utility/cost path segments. $Q \subseteq R$ is selected from the path segments set R generated by deforming an initial path r_0 using gradient descent on the underlying belief function b^ξ shown in figure 11. Refer to [45] for equations describing the sensor model and utility function.

Depth-Limited tree Search (DLS)

The optimal plan is ascertained with an exhaustive tree search ; however, this is computationally intractable for large time budgets B . A solution method that trades optimality for reduced computation is depth-limited (or finite-horizon) tree search (DLS). The process begins by building a complete subtree up to a fixed depth. The highest utility leaf node is identified, and a transition to the child node v_t of the root v_0 along this branch is executed. The child becomes the new root and these steps are repeated until $C(v_t) \geq B$. DLS ideally allows future information gain to influence local decisions.

Monte Carlo Tree Search (MCTS)

While DLS offers computational practicality, a good search depth is unknown for arbitrary problem instances. DLS complexity also remains exponential in the search depth $\mathcal{O}(|Q|^d)$ such that a better solution at depth $d+1$ may just take too long to compute. Monte Carlo Tree Search (MCTS) is an anytime algorithm that can achieve further computational reductions for similarly high-quality plans to DLS by selectively expanding relevant tree nodes. MCTS achieves this with random rollouts at each node expansion to bias tree growth towards high-yield end states. The algorithm is simple yet powerful, and is well described in [9].

Cluster tree search

Interesting scenarios arise when the a-priori probabilistic belief function b_{init}^ξ is partitioned. This could be due to a series of prior uninformed local-area searches, and now the search region has been broadened. For long-term planning, this is problematic for state-of-the-art search schemes, which can be too myopic for a given computational-

limit. Results from DLS illustrate that good solutions concentrate tree search effort at clusters of information. We draw on this observation to explicitly perform local cluster searches and optimally combine the resulting local plans with dynamic programming (DP) based on $J(v)$ and $C(v)$ of every node of each cluster tree [46]. Alg. 1 outlines the approach.

First, information clusters are identified using a Gaussian mixture model G_{mm} . They are ordered in sequence of distance from the UAV start position. Trees for each cluster are stored in the set T . Cluster assigns time budget to the terminal state s_t^i proportionally to a fraction ε greater than the belief proportion b_{init}^ξ in b_{init}^ξ . The DP optimal plan may involve taking time from one cluster and allocating it to another ; ε allows for this. Search can be any tree search algorithm such as DLS or MCTS. For DLS, it suffices to use near-greedy search depths (e.g. 2 or 3) because the search effort is already concentrated at clusters. Minimum-cost path segments are used to link cluster plans together. Finally, Dp returns the optimal plan P .

Algorithm 1 ClusterSearchDp

```

1:  $[b_1^\xi, b_2^\xi, \dots, b_n^\xi] \leftarrow G_{mm}(b_{init}^\xi)$ 
2:  $\mathbb{T} \leftarrow \emptyset$ 
3: for  $i = 1 : n$  do
4:    $s_t^i \leftarrow \text{Cluster} + (b_1^\xi, b_2^\xi, B, \varepsilon)$ 
5:    $T^i \leftarrow \text{Search} + (s_t^i)$ 
6:    $\mathbb{T} \leftarrow \mathbb{T} \cup T^i$ 
7: end for
8: return  $P \leftarrow +Dp + (\mathbb{T})$ 
```

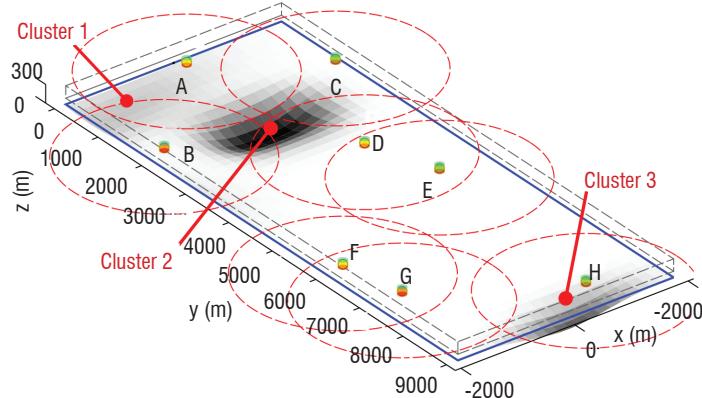


Figure 12 - A complex map scenario with three clusters of belief uncertainty (or information) represented by the dark patches. Thermals are labelled A to H.

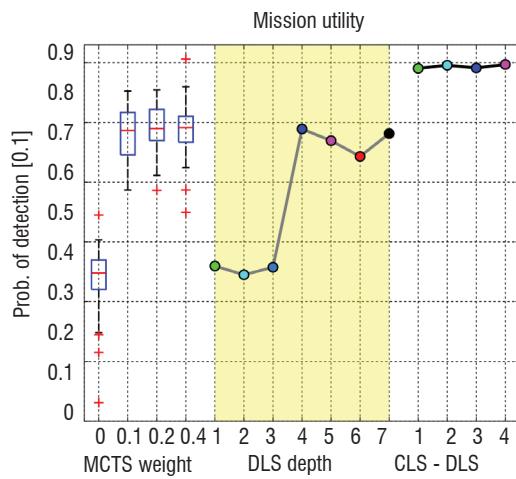


Figure 13 - Comparative results for 1) MCTS, 2) DLS and 3) cluster tree search using DLS on the scenario in figure 12. The trade-off between optimality and complexity is controlled by an exploration weight in MCTS, and the search depth in DLS.

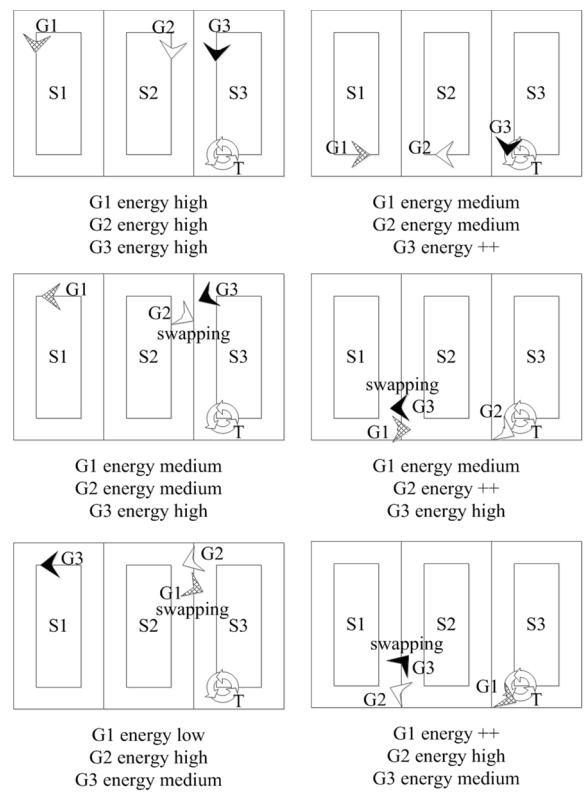


Figure 14 - A team of 3 gliders (G1, G2 and G3) uses the swapping approach to manage the thermal T, extending the mission but keeping the partitioning patrolling strategy.

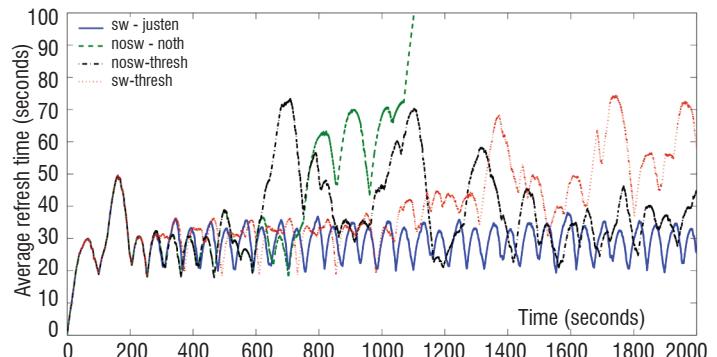
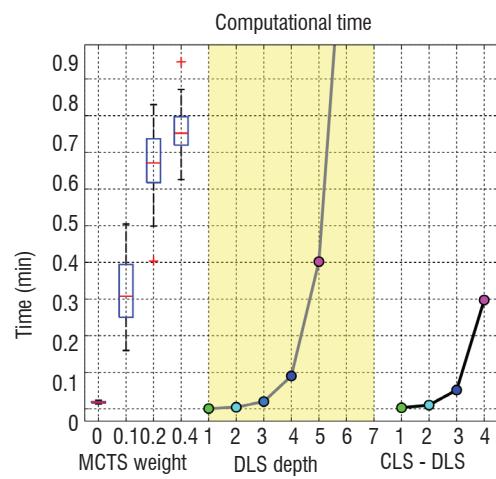


Figure 15 - The average maximum refresh time for a scenario with two thermals and four gliders, tested with different approaches : swapping or non-swapping and just-enough, threshold or no-thermalling approaches.



Simulation results

In the interest of space, results are only shown for one map scenario although these methods have been extensively tested on complex target-search scenarios and have been shown to outperform current greedy planners [46]. This map (figure 12) has three information clusters and eight thermals. The UAV starts at thermal A and has a budget of 60 minutes to maximise information gain.

Performance results are illustrated in figure 13. As MCTS is a randomised solver, we conducted 50 trials for each exploration weight $\{0, 0.1, 0.2, 0.4\}$. A higher weight corresponds to denser tree search which expands more nodes at the expense of increased computation. Figure 13 suggests that 0.1 works the best here. Its utility is close to that of the best DLS depth of 4. The large step in DLS utility between depths 3 and 4 is a result of a longer horizon directing the plan towards cluster 3 in figure 12. In contrast, cluster tree search with DLS can achieve very high utility for greedy and near-greedy search depths.

Persistent monitoring with multiple gliders

A natural extension of the previous work is to examine how soaring missions could be performed with multiple gliding vehicles. This is particularly applicable to large area monitoring problems where the advantages of long endurance and multiple vehicles are readily apparent. However, multi-UAV systems present some difficult challenges to overcome related to coordination and distributed decision-making. Consider a problem where the objective is for multiple UAVs to cooperate to distribute the surveillance of a target in a persistent manner while minimising the refresh time T_r (period between consecutive visits) of all positions in the area. Previous work [1] considered this problem with powered UAVs, and the solution presented was shown to be robust to the failure of vehicles or the introduction of new vehicles whilst maintaining coverage of the target region and minimising refresh time.

However, extending the problem to consider gliding UAVs that require regular energy resupply is non-trivial. We extend the previous problem to consist of a team of N autonomous gliders tasked with monitoring an area S where there are M known thermal sources [72]. A cooperative patrolling strategy is required to coordinate the gliders such that the objectives can be optimised. A partitioning strategy is proposed to solve this problem because it allows consideration of communication constraints and gliders with different capabilities. The problem is partitioned such that each glider is in charge of patrolling a different non-overlapped region and thermals are associated with the regions in which they are located. In this case, the location and state of thermal sources can be considered as resources to be shared between the UAVs depending on their capabilities. This can be approached as a dynamic resource allocation problem to assign the thermals to the most suitable glider, such that all gliders maintain sufficient energy for continuous flight.

A *swapping* method based on one-to-one coordination can be used to manage the thermals in a distributed manner ensuring that all the regions continue being patrolled by at least one of the gliders. The swapping method implies that neighbouring UAVs share information about the thermal sources (location, state) in their own sub-areas. According to this information and the glider states each pair of neighbouring gliders can decide if they need to swap regions, as shown in

figure 14. They decide according to the gliders' current energy and the distance to the nearest thermals. As the gliders converge in their knowledge of the world, they can decide independently and obtain consistent solutions. In this way, each UAV can reach the nearest thermal resource in a known finite time if it needs to gain energy. This approach ensures that the information about the thermals will be shared between all the gliders. Then, the thermals can be dynamically allocated between the gliders whilst maintaining the patrolling strategy.

Finally, another relevant issue is to decide how long a glider has to remain in the thermal. Two general approaches are defined. A threshold approach implies that the gliders go to the nearest thermal when they detect an energy level less than a *threshold* and remain in it until gaining the maximum possible altitude. Alternatively, a just-enough approach implies that when a glider reaches a thermal, it estimates the amount of energy that it will require to reach the next thermal, and remains in the thermal until gaining that estimated energy. This value can be estimated based on the model described in (1) and assuming an *a priori* known path to patrol the whole area.

Early research has shown that a combined just-enough swapping method obtains promising results for cooperative large area monitoring missions with a team of autonomous gliders that exploit the thermal sources. Figure 15 shows the maximum average refresh time computed along the whole path during a mission assuming four gliders and two thermals. A cooperative path partitioning strategy and four different thermal access approaches are considered in the test.

Aerial resupply

In the preceding sections, this paper has described recent work on persistent autonomous flight through the opportunistic exploitation of readily available atmospheric energy. A different, yet complementary approach to persistent flight is the deliberate in-flight resupply of energy. This is a complementary strategy when atmospheric energy is insufficient or unavailable, and an alternative when the platform is not optimised for atmospheric energy collection. In the past this has occurred manually with large, manned aircraft at high altitude, however automating this process allows the pilot to be removed and much smaller aircraft, such as autonomous gliders to utilise the procedure.

It is generally accepted in the literature that the primary barrier to autonomous in-flight resupply is the sensing and navigation challenges surrounding tight formation flight. This challenge is amplified in our work where we consider small, dynamic vehicles that are operating in a turbulent, low altitude environment. These conditions require an accurate and timely relative state estimate that is robust to a dynamic environment. One method to obtain this estimate is to subtract one vehicle's INS/GPS-based state estimate from the other, where the result is known as the raw relative estimate. This approach has merit during high separation formation but the accuracy, particularly in relative position, is in the order of metres and is not sufficiently accurate for close proximity formation flight. The low accuracy is not only attributed to the individual sensor accuracy, but also errors in measurement time synchronisation since absolute measurements are being differenced. This problem also applies to more accurate DGPS systems. Highly dynamic vehicles and sporadic communication dropouts further amplify this effect.

To achieve the necessary accuracy, directly observed relative measurements must be used. Vision is a popular sensor selection in the aerial domain due to its availability, compact size and low weight. The downside to vision is the susceptibility to observation dropouts as a result of occlusion, a constrained field of view (FOV) and uncertain lighting conditions. Further, incorrect observations resulting from clutter and false feature correspondences must be detected and excluded. To negate these shortcomings and create a resilient yet accurate relative state estimator, it is important to incorporate constantly available, albeit absolute, information from sensors such as inertial, magnetic, atmospheric and GPS.

In our scenario, a leader and follower UAV are flying in formation. Visual markers are mounted on the leader at each wingtip, at the top of the tail fin and on the right of the tail plane as shown in figure 16. A forward facing camera is mounted on the follower and provides relative measurements to the leader's visual markers. All onboard sensor data from both aircraft are available on the follower UAV in real-time via wireless communications. The following sections summarise the relative estimation framework, the vision integration and provide preliminary implementation results.

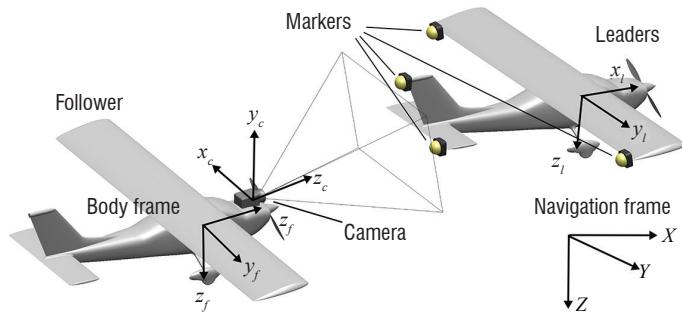


Figure 16 - Leader-follower coordinate frames and the marker based vision system

Multi-vehicle relative navigation

The proposed relative estimator fuses vehicle-to-vehicle visual measurements with information from GPS, inertial, magnetic and atmospheric sensors, located on each UAV, in a tightly coupled fashion. An unscented Kalman filter (UKF) [26] provides the filter framework to estimate the position P , velocity V and attitude quaternion q for each aircraft. P and V are expressed in the local tangential frame, relative to a ground station.

$$x_{lf} = [P_l \ V_l \ q_l \ P_f \ V_f \ q_f]^T \quad (19)$$

$$\begin{bmatrix} P_{lf} \\ V_{lf} \\ q_{lf} \end{bmatrix} = \begin{bmatrix} P_l - P_f \\ V_l - V_f \\ q_f^{-1} \otimes q_l \end{bmatrix} \quad (20)$$

A UKF has several advantages over the traditional extended Kalman filter (EKF). It provides at least second-order nonlinear approximation as opposed to the first-order EKF; derivation of Jacobians is not necessary, the filter is more robust to initial errors and computation can occur in parallel. Resilience to initial error is particularly important because of the large difference in accuracy between the GPS and vision-based measurements. A downside of the UKF is that a quaternion parametrisation of the attitude results in a non-unit quaternion when the mean is computed. A brute force normalisation can be made to

work, but is undesirable. Instead, we use generalised Rodrigues parameters to represent the attitude error, as proposed in [13].

Each vehicle state is propagated using the bias corrected onboard inertial measurements and the mechanization equations found in [53]. Details pertaining to the vehicle atmospheric, magnetic and GPS sensor updates are also omitted but can be found in [19]. UKF prediction and update equations are well known and are provided in [26, 47, 13].

Vision integration

Relative pose estimation using vision sensors has been well researched and many valid approaches exist. Our work employs a feature based method where visual markers of a known configuration are mounted on a leader vehicle and observed by a follower. Using the set of n correspondences between the 3D marker positions, ζ_j^l and the 2D observations $\tilde{\delta}_j$, as well as the camera intrinsic parameters, the relative pose can be calculated directly. This requires $n \geq 3$ for a solution and $n \geq 4$ for a unique solution. A number of algorithms are available to solve this PnP problem, including POSIT [14] which is used as a benchmark in Section 3.3, the Lu-Hager-Mjolsness algorithm [38] and an efficient approach called EPnP [36].

The downside to this vision only approach is that it fails with incorrect point matching, occlusion or a target outside the FOV. These brief or prolonged measurement dropouts are highly undesirable, particularly during close proximity operations. Alternatively, one could fuse the pose estimate from one of the aforementioned algorithms with the onboard sensor data in a loosely-coupled arrangement, however preliminary results with a fixed measurement covariance displayed inferior performance to the tightly-coupled equivalent. One reason may be that the measurement covariance is dynamic and a function of the relative pose, in addition to the pixel noise. Deriving an expression for this is difficult.

Instead, we propose a tightly-coupled approach which uses n raw 2D marker observations, $\tilde{\delta}_j = [u_j \ v_j]^T$, $j=1, \dots, n$. In our case $n=4$ and $n \geq 3$ is required for observability within the UKF [66, 20]. The expected observations $\bar{\delta}_j$, $j=1, \dots, n$ are calculated by first transforming ζ_j^l from the leader's body frame to the world frame, ζ_j^f using (21). In this case the world frame is the follower's body frame.

$$\zeta_j^f = C_n^f (C_l^n \zeta_j^l + P_{lf}) \quad (21)$$

Next, the vision sensor extrinsic parameters transform ζ_j^f to the camera frame using (22). P_{fc} and C_f^c are the translation and rotation from the follower's body frame to the camera frame. C_f^c includes both the camera mounting orientation and the axes transformation.

$$\zeta_j^c = [C_f^c \ P_{fc}] \begin{bmatrix} \zeta_j^f \\ 1 \end{bmatrix} \quad (22)$$

$\bar{\delta}_j$ is calculated using K , the camera intrinsic matrix which encapsulates the camera focal length, aspect ratio, principal point and distortion. The final vision measurement model is provided in (24) and the correction occurs sequentially.

$$\begin{bmatrix} \bar{\delta}_j \\ 1 \end{bmatrix} = K \begin{bmatrix} \zeta_{yj}^c / \zeta_{xj}^c \\ \zeta_{yj}^c / \zeta_{zj}^c \\ 1 \end{bmatrix} \quad (23)$$

$$h^{vision}[x, k] = [\bar{\delta}_1 \bar{\delta}_2 \dots \bar{\delta}_n]^T \quad (24)$$

The vision-based observation model presented in this section updates both the relative position and orientation by assuming correct point matching. This is not always possible, particularly when the target is far and the points are difficult to distinguish from one another. Rather than neglecting such a measurement, a simpler observation model can be utilised to extract P_{lf} information. Here, the vision observation becomes the average or centroid of $\tilde{\delta}_j, j = 1, \dots, n$ as an approximation for the target's centre of gravity and (21) is replaced with (25) where $m = 1$. Although relative orientation and range become unobservable, P_{lf} and V_{lf} accuracy is improved.

$$\zeta_j^f = C_j^f P_{lf} \quad (25)$$

Before y_{lf}^{vision} can be used, correspondences between the observed points $\tilde{\delta}_i$ and the projected model points $\bar{\delta}_j$ must be determined. To do this, unique marker characteristics could be used, which may include colour, size, intensity and frequency. However, in our application we have chosen to use homogeneous visual markers to simplify the MV task and instead use the marker model to match the points. To do this, we use a computationally efficient, deterministic mutual nearest point procedure [42]. Before this is implemented, we eliminate the linear translation between the point sets by subtracting the vector $(\bar{\delta}_\mu - \tilde{\delta}_\mu)$ from $\bar{\delta}_j$. This eliminates errors in relative azimuth, elevation as well as follower attitude and simplifies the matching process. The matrix Θ is then populated with the pixel distances between $\tilde{\delta}_i$ and $\bar{\delta}_j$.

$$\Theta = \begin{bmatrix} d(\bar{\delta}_1, \tilde{\delta}_1) & \dots & d(\bar{\delta}_1, \tilde{\delta}_n) \\ \dots & \ddots & \dots \\ d(\bar{\delta}_m, \tilde{\delta}_1) & \dots & d(\bar{\delta}_m, \tilde{\delta}_n) \end{bmatrix} \quad (26)$$

Where $d(.,.)$ is the linear pixel distance between points. Θ_{col}^{\min} and Θ_{row}^{\min} are the minimum value of each column and row of Θ , respectively and Θ_{col}^{index} contains the index of the minimum value in each column.

$$\begin{aligned} \Theta_{col}^{\min} &= [min(d(\bar{\delta}_1, \tilde{\delta}_1)) \dots min(d(\bar{\delta}_1, \tilde{\delta}_m))] \\ \Theta_{row}^{\min} &= [min(d(\bar{\delta}_1, \tilde{\delta}_1)) \dots min(d(\bar{\delta}_n, \tilde{\delta}_1))] \\ \Theta_{col}^{index} &= [\Theta_{col_1}^{\min} \dots index(\Theta_{col_n}^{\min})] \end{aligned} \quad (27)$$

For a point to be valid, it must satisfy (28), that is to say a valid point in Θ must be the minimum of both its column and row. A threshold of validity can also be implemented to reject outliers and noise.

$$\Theta_{col}^{\min}[i] = \Theta_{col}^{\min}[\Theta_{col}^{index}[i]] \quad (28)$$

	Raw (1 σ)	POSIT (1 σ)	$x_{lf}(1\sigma)$	Raw impr.
P_{lf}				
N	2.069 (0.86)	0.353 (0.15)	0.205 (0.23)	90.1 %
E	1.901 (0.76)	1.017 (0.34)	0.427 (0.24)	77.5 %
D	0.152 (0.11)	0.123 (0.03)	0.129 (0.07)	15.1 %
V_{lf}				
N	0.203 (0.09)	-	0.216 (0.25)	-6.54 %
E	0.227 (0.18)	-	0.183 (0.18)	19.4 %
D	0.068 (0.09)	-	0.052 (0.05)	23.5 %
Q_{lf}				
φ	1.148 (0.81)	0.295 (0.01)	0.173 (0.24)	85.0 %
θ	1.095 (1.12)	0.660 (0.02)	0.208 (0.17)	81.0 %
ψ	14.743 (8.90)	0.577 (0.01)	0.443 (0.85)	97.0 %

Table 1 - Raw relative, POSIT and relative UKF estimate RMSE comparison from 100 simulations. Axes are N(orth), E(east) and D(down). Positions are measured in m, velocities in m/s and angles in degrees.

Implementation

The estimator was tested in a high fidelity simulated environment [67] where conditions are repeatable and the ground truth is known. The simulation was run 100 times and the results are summarised numerically in table 1. Compelling performance improvements were observed when compared to both the raw relative estimate and the benchmark vision-only pose estimation algorithm, POSIT. Particularly large improvements in horizontal position and ψ can be attributed to the relative inaccuracy of the GPS and magnetometers. As expected, the gains over POSIT are less but remain notable which is likely due to a smoothing effect of the vehicle inertial measurements.

The algorithm was also testing in ground based experiments on a dual-UAV system to isolate the relative navigation problem, and demonstrate the estimation framework functioning in real-time on an embedded system. This system consists of two fixed-wing UAVs, an autopilot and formation flight computer onboard each aircraft, LED markers on the leader, and a camera on the follower.

Relative position and attitude estimates from one such experiment are shown in figures 17a and 17b. Here, we can see good agreement between the vision-only POSIT algorithm and the output of the relative estimator. A slight bias can be observed in the east and ψ components which indicate a slight error in the camera extrinsic calibration. Additionally, we see that POSIT fails between 23-29 seconds because less than four points are available. Here, the relative estimator is able to utilise information from even a single visual marker and only slowly degrades to the raw relative estimate when no visual measurements are available.

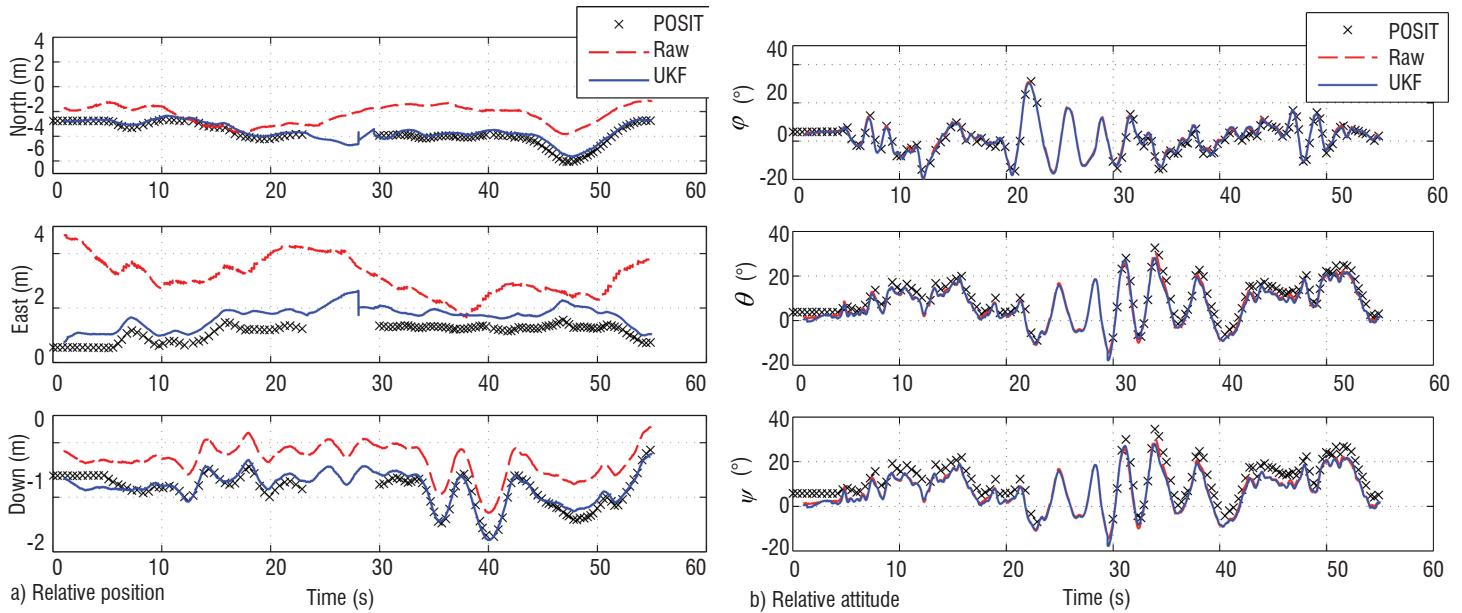


Figure 17 - Vision based relative estimation results from the dual-UAV system

Conclusion

The work presented here is aimed at extending the flight duration of fixed-wing UAVs. This paper highlighted some of the work performed at the University of Sydney towards reaching this goal. We examined the basic mechanisms of soaring flight and used these equations to derive utility functions for planning. This was extended to planning in unknown wind fields using GP regression for wind map building, and for use in an RL framework to develop energy gaining trajectories without specifying control strategies. Further, soaring was integrated

into target search problems multiple UAVs. The paper also looked at coordination for aerial refueling with small UAVs.

Future work will aim to bring some of these methods together for long-duration UAV missions. The focus will be on autonomy and decision-making, ideally for a system with multiple heterogeneous vehicles to perform an externally driven mission. The system should make decisions about the energy available for soaring and the requirement for inflight refueling to allocate vehicles based on their demands for energy and mission utility ■

Acknowledgements

This work is supported by the Australian Centre for Field Robotics at the University of Sydney. J.J. Acevedo visited the University of Sydney through the framework of the MUAC-IREN (FP7-PEOPLE-295300) project of the European Commission. The author J.J. Acevedo acknowledges too the CLEAR (DPI2011-28937-C02-01) Spanish National Research project and the project of excellence of the Junta de Andalucía WSAN-UAV (P09-TEP-5120).

Acronyms

DLS	(Depth-Limited (tree) Search)
DP	(Dynamic Programming)
EKF	(Extended Kalman Filter)
MCTS	(Monte Carlo Tree Search)
RL	(Reinforcement learning)
RMSE	(Root Mean Square Error)
TD	(Temporal difference)
UAV	(Unmanned Aerial Vehicle)
UKF	(Unscented Kalman Filter)

SBXC Glider Parameters

Parameter	Value	Units	Explanation
$C_{D,O}$	0.012		Parasitic drag coefficient
S	0.957	m^2	Wing reference area
$\mathcal{A}R$	19.54		Wing aspect ratio
e	0.85		Oswald's efficiency factor
m	5.44	kg	Vehicle mass
N_{\max}	2.0		Maximum load factor
N_{\min}	0		Minimum load factor
$C_{L,\max}$	1.2		Maximum lift coefficient
$\frac{d\phi}{dt}_{\max}$	30	$^{\circ}/s$	Maximum roll rate
$\gamma_{\alpha,\max}$	50	$^{\circ}$	Max. air relative climb
$\left(\frac{L}{D}\right)_{est}$	50		Approximate glide ratio

Table 2 - The aerodynamic and geometric properties of the SB-XC glider model

References

- [1] J. J. ACEVEDO, B. N. C. ARRUE, I. MAZA, and A. OLLERO - *Cooperative Large Area Surveillance with a Team of Aerial Mobile Robots for Long Endurance Missions*. Journal of Intelligent and Robotic Systems 70 (2013), 329–345.
- [2] H. AIRY - *The Soaring of Birds*. Nature 27, 703 (19 April 1883), 590–592.
- [3] Z. ÁKOS, M. NAGY, S. LEVEN and T. VICSEK - *Thermal Soaring Flight of Birds and Unmanned Aerial Vehicles*. Bioinspiration & Biomimetics 5, 4 (2010), 045003.
- [4] M. J. ALLEN - *Autonomous Soaring for Improved Endurance of a Small Uninhabited Air Vehicle*. 43rd AIAA Aerospace Sciences Meeting and Exhibit (Reno, Nevada, 2005). AIAA Paper 2005-1025.
- [5] M. J. ALLEN, and V. LIN - *Guidance and Control of an Autonomous Soaring UAV*. Technical Memorandum NASA/TM-2007-214611, NASA Dryden Flight Research Center, February 2007.
- [6] Jr. J. D. ANDERSON - *Fundamentals of Aerodynamics*. Fourth ed. McGraw-Hill Series in Aeronautical and Aerospace Engineering. McGraw-Hill, New York, 2007.
- [7] R. BARATE, S. DONCIEUX and J.-A. MEYER - *Design of a Bio-Inspired Controller for Dynamic Soaring in a Simulated Unmanned Aerial Vehicle*. Bioinspiration & Biomimetics 1, 3 (2006), 76–88.
- [8] M. B. BOSLOUGH - *Autonomous Dynamic Soaring Platform for Distributed Mobile Sensor Arrays*. Tech. Rep. SAND2002-1896, Sandia National Laboratories, June 2002.
- [9] C. B. BROWNE, E. POWLEY, D. WHITEHOUSE, S. M. LUCAS, P. I. COWLING, P. ROHLFSHAGEN, S. TAVENER, D. PEREZ, S. SAMOTHRAKIS and S. COLTON - *A Survey of Monte Carlo Tree Search Methods*. IEEE Trans. Comput. Intell. AI in Games 4, 1 (2012), 1–43.
- [10] J. J. CHUNG, N. R. J. LAWRENCE, and S. SUKKARIEH - *Gaussian Processes for Informative Exploration in Reinforcement Learning*. 2013 IEEE International Conference on Robotics and Automation (2013), IEEE, pp. 2633–2639.
- [11] J. J. CHUNG, N. R. J. LAWRENCE and S. SUKKARIEH - *Resource Constrained Exploration in Reinforcement Learning*. Robotics: Science and Systems Workshop on Robotic Exploration, Monitoring, and Information Collection: Nonparametric Modeling, Information-based Control, and Planning under Uncertainty (Berlin, Germany, 2013).
- [12] J. J. CHUNG, M. A. TRUJILLO and S. SUKKARIEH - *A New Utility Function for Smooth Transition Between Exploration and Exploitation of a Wind Energy Field*. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems Conference Proceedings (2012), pp. 4999–5005.
- [13] J. CRASSIDIS AND F. MARKLEY - *Unscented Filtering for Spacecraft Attitude Estimation*. AIAA Guidance, Navigation, and Control Conference (2003), AIAA.
- [14] D. F. DEMENTHON and L. S. DAVIS - *Model-Based Object Pose in 25 Lines of Code*. International Journal of Computer Vision 15, 1-2 (1995), 123–141.
- [15] N. T. DEPENBUSCH and J. W. LANGELAAN - *Receding Horizon Control for Atmospheric Energy Harvesting by Small UAVs*. AIAA Guidance, Navigation and Control Conference (Toronto, Ontario, Canada, 2010). AIAA Paper 2010-8180.
- [16] J. DOEBBLER, T. SPAETH, J. VALASEK, M. J. MONDA and H. SCHAUD - *Boom and Receptacle Autonomous Air Refueling Using a Visual Pressure Snake Optical Sensor*. AIAA Atmospheric Flight Mechanics Conference and Exhibit (2006), American Institute of Aeronautics and Astronautics.
- [17] S. DONCIEUX, J.-B. MOURET and J.-A. MEYER - *Soaring Behaviors in UAVs : Animat Design Methodology and Current Results*. 7th European Micro Air Vehicle Conference (MAV07) (Toulouse, France, 2007), pp. 1–10.
- [18] D. J. EDWARDS and L. M. SILVERBERG - *Autonomous Soaring : the Montague Cross-Country Challenge*. Journal of Aircraft 47, 5 (2010), 1763 – 1769.
- [19] T. FIORENZANI, C. MANES, G. ORIOLO and P. PELITI - *Comparative Study of Unscented Kalman Filter and Extended Kalman Filter for Position/Attitude Estimation in Unmanned Aerial Vehicles*. Tech. rep., Istituto Di Analisi Dei Sistemi ed Informatica, 2008.
- [20] A. FOSBURY and J. CRASSIDIS - *Relative Navigation of Air Vehicles*. Journal of Guidance, Control, and Dynamics 31, 4 (2008), 824–834.
- [21] C. GIAMPIERO, F. MARIO LUCA, F. ANTONIO, N. MARCELLO, S. BRAD and P. MARIO - *Autonomous Aerial Refueling for UAVs Using a Combined GPS-Machine Vision Guidance*. AIAA Guidance, Navigation, and Control Conference and Exhibit (2004), Guidance, Navigation, and Control and Co-located Conferences, American Institute of Aeronautics and Astronautics.
- [22] Y. GU, B. SEANOR, G. CAMPA, M. NAPOLITANO, L. ROWE, S. GURURAJAN and S. WAN - *Design and Flight Testing Evaluation of Formation Control Laws*. Control Systems Technology, IEEE Transactions on 14, 6 (2006), 1105–1112.
- [23] E. A. HAERING - *Airdata Calibration of a High-Performance Aircraft for Measuring Atmospheric Wind Profiles*. Technical memorandum NASA-TM-101714, NASA, January 1990.
- [24] P. IDRAC - *Experimental Study of the “Soaring” of Albatrosses*. Nature 115, 2893 (1925), 532–532.
- [25] 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 (2007), 707–738.
- [26] 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 (1995), vol. 3, pp. 1628–1632.
- [27] S. M. KHANSARI-ZADEH and F. SAGHAFI - *Vision Based Navigation in Autonomous Close Proximity Operations Using Neural Networks*. Aerospace and Electronic Systems, IEEE Transactions on 47, 2 (2011), 864–883.
- [28] J. A. KYLE - *Optimal Soaring by a Small Autonomous Glider*. PhD thesis, Oregon State University, 2006.
- [29] J. W. LANGELAAN - *Gust Energy Extraction for Mini and Micro Uninhabited Aerial Vehicles*. Journal of Guidance, Control, and Dynamics 32, 2 (2009), 464–473.
- [30] J. W. LANGELAAN, J. SPLETZER, C. MONTELLA and J. GRENSTEDT - *Wind Field Estimation for Autonomous Dynamic Soaring*. Robotics and Automation(ICRA), 2012 IEEE International Conference on (May 2012), pp. 16 –22.
- [31] N. R. J. LAWRENCE - *Autonomous Soaring Flight for Unmanned Aerial Vehicles*. PhD thesis, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, 2011.
- [32] N. R. J. LAWRENCE and S. SUKKARIEH - *A Guidance and Control Strategy for Dynamic Soaring with a Gliding UAV*. "Robotics and Automation (ICRA), 2009 IEEE International Conference on" (Kobe, Japan, 2009), pp. 3632–3637.
- [33] N. R. J. LAWRENCE and S. SUKKARIEH - *Wind Energy Based Path Planning for a Small Gliding Unmanned Aerial Vehicle*. AIAA Guidance, Navigation and Control Conference (Chicago, Illinois, 2009). AIAA Paper 2009-6112.
- [34] N. R. J. LAWRENCE and S. SUKKARIEH - *Autonomous Exploration of a Wind Field with a Gliding Aircraft*. Journal of Guidance, Control, and Dynamics 34, 3 (2011), 719–733.

- [35] N. R. J. LAWRENCE and S. SUKKARIEH - *Path Planning for Autonomous Soaring Flight in Dynamic Wind Fields*. IEEE International Conference on Robotics and Automation (Shanghai, China, 2011), pp. 2499 – 2505.
- [36] V. LEPETIT, F. MORENO-NOGUER and P. FUA - *Epnlp: An Accurate o(n) Solution to the PNP problem*. International Journal of Computer Vision 81, 2 (2009), 155–166.
- [37] P. LISSAMAN - *Wind Energy Extraction by Birds and Flight Vehicles*. 43rd AIAA Aerospace Sciences Meeting and Exhibit (Reno, Nevada, 2005). AIAA Paper 2005-241.
- [38] C. P. LU, G. D. HAGER and E. MJOLSNESS - *Fast and Globally Convergent Pose Estimation from Video Images*. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22, 6 (2000), 610–622.
- [39] P. B. MACCREADY - *Optimum Airspeed Selector*. Soaring 9 (March 1954), 8.
- [40] P. B. MACCREADY - *Optimum Airspeed Selector*. Soaring January–February (1958), 10–11.
- [41] Z. MAHBOUBI, Z. KOLTER, T. WANG and G. BOWER - *Camera Based Localization for Autonomous UAV Formation Flight*. Proceedings of the AIAA@ Infotech Conference (2011).
- [42] M. MAMMARELLA, G. CAMPA, M. R. NAPOLITANO and M. L. FRAVOLINI - *Comparison of Point Matching Algorithms for the UAV Aerial Refueling Problem*. Machine Vision and Applications 21, 3 (2010), 241–251.
- [43] M. MAMMARELLA, G. CAMPA, M. R. NAPOLITANO, M. R. FRAVOLINI, M. L., 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 (2008), 791–801.
- [44] C. MARTINEZ, T. RICHARDSON and P. CAMPOY - *Towards Autonomous Air-to-Air Fuelling for UAVs Using Visual Information*. Robotics and Automation (ICRA), 2013 IEEE International Conference on (2013), pp. 5756–5762.
- [45] J. L. NGUYEN, N. R. J. LAWRENCE, N. R. J. FITCH and S. SUKKARIEH - *Energy-Constrained Motion Planning for Information gathering with autonomous aerial soaring*. Robotics and Automation (ICRA), 2013 IEEE International Conference on (2013), IEEE, pp. 3825–3831.
- [46] J. L. NGUYEN, N. R. J. LAWRENCE and S. SUKKARIEH - *Nonmyopic Planning for long-Term Information Gathering With an Aerial Glider*. Robotics and Automation (ICRA), 2014 IEEE International Conference on (2014), 6573–6578.
- [47] S. OH and E. N. JOHNSON - *Relative Motion Estimation for Vision-Based Formation Flight Using Unscented Kalman Filter*. AIAA Guidance, Navigation and Control Conference and Exhibit (2007), AIAA 2007-6866.
- [48] S. PARK, J. DEYST and J. P. HOW - *A New Nonlinear Guidance Logic for Trajectory Tracking*. Proceedings of the AIAA Guidance, Navigation and Control Conference (2004), Citeseer.
- [49] S. PEAL - *Soaring of birds*. Nature 23, 575 (4 November 1880), 10–11.
- [50] C. PENNYCUICK - *Field Observations of Thermals and Thermal Streets, and the Theory of Cross-Country Soaring Flight*. Journal of Avian Biology 29, 1 (1998), 33–43.
- [51] C. PENNYCUICK, T. ALERSTAM and B. LARSSON - *Soaring Migration of the Common Crane Grus Crus Observed by Radar and from an Aircraft*. Ornis Scandinavica 10, 2 (1979), 241–251.
- [52] L. RAYLEIGH - *The Soaring of Birds*. Nature 27, 701 (5 April 1883), 534–535.
- [53] R. M. ROGERS - *Applied Mathematics in Integrated Navigation Systems*, 3rd ed. AIAA Education Series. AIAA, 2007.
- [54] G. A. RUMMERY and M. NIRANJAN - *On-line Qlearning Using Connectionist Systems*. Tech. rep., Cambridge University Engineering Department, 1994.
- [55] G. SACHS - *Minimum Shear Wind Strength Required for Dynamic Soaring of Albatrosses*. Ibis 147, 1 (2005), 1–10.
- [56] R. J. SATTIGERI - *Adaptive Estimation and Control with Application to Vision-Based Autonomous Formation Flight*. PhD thesis, School of Aerospace Engineering, Georgia Institute of Technology, 2007.
- [57] R. SPAAR and B. BRUDERER - *Optimal Flight Behavior of Soaring Migrants: a Case Study of Migrating Stepp Buzzards, Buteo Buteo Vulpinus*. Behavioral Ecology 8, 3 (1997), 288–297.
- [58] R. SUTTON and A. BARTO - *Reinforcement Learning: An Introduction*. The MIT press, 1998.
- [59] A. L. TREASTER and A. M. YOCUM - *The Calibration and Application of Five-Hole Probes*. Tech. Rep. TM-78-10, Pennsylvania State University Applied Research Laboratory, 1978.
- [60] V. TUCKER - *Flight Energetics in Birds*. American Zoologist 11, 1 (1971), 115–124.
- [61] J. VALASEK, K. GUNNAM, J. KIMMETT, J. L. JUNKINS, D. HUGHES and M. D. TANDALE - *Vision-Based Sensor and Navigation System for Autonomous Air Refueling*. Journal of Guidance, Control, and Dynamics 28, 5 (2005), 979–989.
- [62] S. Walkden - *Experimental Study of the "Soaring" of Albatrosses*. Nature 116, 2908 (25 July 1925), 132–134.
- [63] H. WEIMERSKIRCH, T. GUIONNET, J. MARTIN, S. A. SHAFFER and D. P. COSTA - *Fast and Fuel Efficient? Optimal use of Wind by Flying Albatrosses*. Proceedings of the Royal Society of London - Biological Sciences 267, 1455 (2000), 1869 – 1874.
- [64] J. WHARINGTON - *Autonomous Control of Soaring Aircraft by Reinforcement Learning*. PhD thesis, Royal Melbourne Institute of Technology, 1998.
- [65] J. WHARINGTON and I. HERSZBERG - *Control of a High Endurance Unmanned Aerial Vehicle*. 21st Congress of International Council of the Aeronautical Sciences (Melbourne, Australia, September 1998). ICAS-98-3.7.1.
- [66] 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 (2009), 262–275.
- [67] D. B. WILSON, A. H. GÖKTOĞAN and S. SUKKARIEH - *UAV Rendezvous: From Concept to Flight Test*. Australasian Conference on Robotics and Automation (ACRA) (2012).
- [68] E. WONG, F. BOURGAULT and T. FURUKAWA - *Multivehicle Bayesian Search for Multiple Lost Targets*. Proc. of "Robotics and Automation (ICRA), 2005 IEEE International Conference on".
- [69] C. J. WOOD - *The Flight of Albatrosses (a computer simulation)*. Ibis 115, 2 (1972), 244–256.
- [70] Y. J. ZHAO - *Optimal Patterns of Glider Dynamic Soaring*. Optimal Control Applications and Methods 25, 2 (2004), 67–89.
- [71] Y. J. ZHAO and Y. C. QI - *Minimum Fuel Powered Dynamic Soaring of Unmanned Aerial Vehicles Utilizing Wind Gradients*. Optimal Control Applications and Methods 25, 5 (2004), 211–233.
- [72] J.J. ACEVEDO, N.R.J. LAWRENCE, B.C. ARRUE, S. SUKKARIEH, and A. OLLERO - *Persistent monitoring with a team of autonomous gliders using static soaring*. 2014 IEEE/RS J. International Conference on Intelligent Robots and Systems Conference Proceedings (2014), to appear.

AUTHORS



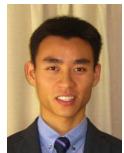
Nicholas R. J. Lawrence completed his B.E. in Aeronautical Space Engineering in 2007 and PhD in Robotics in 2011, both at the University of Sydney. He is currently a research associate at the Australian Centre for Field Robotics at the University of Sydney. His research interests include autonomous soaring for unmanned aerial vehicles, aircraft performance, reinforcement learning and informative path planning.



Jose Joaquin Acevedo received the Telecommunication Engineer Degree in 2007 and the Electrical Energy Systems M.Sc Degree in 2012, both from the University of Seville (Spain). From 2007 to 2010, he worked for different ICT companies. He is currently a PhD Student at the Systems Engineering and Automation Department of the University of Seville, at the Robotic, Vision and Control Group. He is working in different national and European research projects and has published several papers in International journals and conferences. His research interests include coordination and cooperation in distributed multi-robot and multi-UAV systems and task and resource allocation.



Jen Jen Chung is currently completing her Ph.D. at the Australian Centre for Field Robotics at the University of Sydney where she also received her B.E. in Aeronautical Space Engineering. Her research focuses on the development of information-based exploration strategies that can be applied within reinforcement learning frameworks to improve learning performance.



Joseph Nguyen is a Ph.D candidate at the Australian Centre for Field Robotics (ACFR) at The University of Sydney. He received a BE degree in Aeronautical (Space) Engineering from The University of Sydney in 2010. His research interests include aerial robotics and path planning for long-endurance autonomous flight.



Daniel B. Wilson received the B.E. degree in mechanical (space) engineering from The University of Sydney, in 2009. He is currently working towards the Ph.D. degree in the aerospace group at the Australian Centre for Field Robotics (ACFR), at The University of Sydney. In 2010, he worked at the Martin Aircraft Company, developing a flight control system for the Martin jetpack. His research interests include relative navigation for autonomous formation flight and guidance and control systems for aerial vehicles.



Salah Sukkarieh is the Professor of Robotics and Intelligent Systems at the University of Sydney, and the Director of Research and Innovation at the Australian Centre for Field Robotics. He has supervised over 10 research fellows, and graduated over 30 PhDs, 5 Masters and 75 honours students. Salah is on the editorial board for the Journal of Field Robotics, Journal of Autonomous Robots, and Transactions of Aerospace Systems, and has over 300 academic and industry publications in robotics and intelligent systems.