Vanegas, Fernando, Campbell, Duncan, Roy, Nicholas, Gaston, Kevin J., & Gonzalez, Luis F.
(2017)
UAV tracking and following a ground target under motion and localisation uncertainty. In
Aerospace Conference, 2017 IEEE, 4-11 March 2017, Big Sky, Montana.

https://doi.org/10.1109/AERO.2017.7943775

UAV Tracking and Following a Ground Target under Motion and Localisation Uncertainty

Fernando Vanegas
Queensland University of Technology (QUT)
Australian Research Centre for Aerospace Automation
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 1772
fernando.vanegasalvarez@hdr.qut.edu.au

Duncan Campbell
Queensland University of Technology (QUT)
Australian Research Centre for Aerospace Automation
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 2179
da.campbell@qut.edu.au

Nicholas Roy
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology (MIT)
The Stata Center, Building 32 32 Vassar Street Cambridge, MA 02139 USA
+1 (617) 253-2517
nickroy@csail.mit.edu

Kevin J. Gaston
Environment & Sustainability Institute, University of Exeter
Penryn Campus, Penryn, Cornwall TR10 9FE, UK
+44 (0) 1326 259490
k.j.gaston@exeter.ac.uk

Felipe Gonzalez
Queensland University of Technology (QUT)
Australian Research Centre for Aerospace Automation
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 1363
felipe.gonzalez@qut.edu.au

Abstract— **Unmanned Aerial Vehicles (UAVs) are increasingly being used in numerous applications, such as remote sensing, environmental monitoring, ecology and search and rescue missions. Effective use of UAVs depends on the ability of the system to navigate in the mission scenario, especially if the UAV is required to navigate autonomously. There are particular scenarios in which UAV navigation faces challenges and risks. This creates the need for robust motion planning capable of overcoming different sources of uncertainty. One example is a UAV flying to search, track and follow a mobile ground target in GPS-denied space, such as below canopy or in between buildings, while avoiding obstacles. A UAV navigating under these conditions can be affected by uncertainties in its localisation and motion due to occlusion of GPS signals and the use of low cost sensors. Additionally, the presence of strong winds in the airspace can disturb the motion of the UAV. In this paper, we describe and flight test a novel formulation of a UAV mission for searching, tracking and following a mobile ground target. This mission is formulated as a Partially Observable Markov Decision Process (POMDP) and implemented in real flight using a modular framework. We modelled the UAV dynamic system, the uncertainties in motion and localisation of both the UAV and the target, and the wind disturbances. The framework computes a motion plan online for executing motion commands instead of flying to way-points to accomplish the mission. The system enables the UAV to plan its motion allowing it to execute information gathering actions to reduce uncertainty by detecting landmarks in the scenario, while making predictions of the mobile target trajectory and the wind speed based on observations. Results indicate that the system overcomes uncertainties in localisation of both the aircraft and the target, and avoids collisions into obstacles despite the presence of wind. This research has the potential of use particularly for remote monitoring in the fields of biodiversity and ecology.**

TABLE OF CONTENTS

1. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) are popular platforms being used in diverse civilian applications. Examples of such applications include search and rescue operations, environmental monitoring, ecology, mapping, remote sensing, surveying and crop inspection [1],[2],[3],[4],[5]. In most of these applications the UAV is operated either remotely by a pilot, or its path is planned using GPS waypoints programmed through a ground control station. The UAV relies on an accurate estimation of its location by using a Global Positioning System (GPS) on-board sensor in order to successfully accomplish these missions. However, there are particular scenarios in which the GPS signal could be occluded, affecting the positioning system and generating a high degree of uncertainty to the UAV location. Flying autonomously in GPS-denied environments becomes dangerous and the UAV needs alternative methods to overcome the large uncertainty in its localisation [6],[7].

An additional challenging further situation is presented when the UAV has the objective of finding a target whose location is unknown. This constitutes a multi-objective problem [8],[9],[10],[11]. Moreover, the mission increases in difficulty if this target is moving on the ground and the scenario is cluttered with obstacles that must be avoided by the UAV. This scenario can be even more intricate if the UAV has to navigate under windy conditions. In [12], an MDP approach was used to take advantage of the wind energy but only for UAV navigation, no multiple objectives were considered in the mission.

A solution to the UAV navigation problem in GPS-denied scenarios is to use on-board cameras for visual odometry and landmark detection [13] and a probabilistic motion planning approach that enables the UAV to overcome different sources of uncertainty to navigate safely without colliding into objects in the airspace.

Most of the research presented on target finding and tracking is based on simulated scenarios [14], [15], [16], [17], [18], which might oversimplify the real conditions of natural environments. In [19], the authors present a system that models the target's pose uncertainty using multi-modal Gaussian belief-states. They assume the uncertainty can be modelled using Gaussian probability distributions but also that there is no uncertainty in the UAV localisation. Other works focus on detecting targets in real scenarios but discretise the state space and rely on accurate GPS systems for UAV localisation [20].

In this paper, we present the formulation of the UAV Target Tracking mission as a sequential decision problem under multiple sources of uncertainty. We build upon the existing work presented in [13], [21] and increase the complexity of the UAV mission by including wind disturbances to the UAV in a Partially Observable Markov Decision Process (POMDP) formulation. We use on-board wind estimation to update the belief-state based on the wind speeds and orientation observations. The POMDP formulation updates the motion policy based on the estimated prevailing wind conditions. Furthermore, this paper presents a study in which different reward functions are proposed and tested in order to consider the multiple objectives of the mission and to take advantage of the wind conditions present in the flying airspace.

The framework uses a low cost quad-copter platform incorporated into a modular system running the Robotic Operating System (ROS).

The main contributions of this paper are:

1. A novel approach for planning the motion of the UAV in a target tracking mission under multiple sources of uncertainty, such as UAV and target localisation; UAV and target motion and wind disturbances affecting the UAV motion. This whole mission is formulated as a POMDP in continuous state space which includes a model of the wind conditions in the scenario that are updated based on onboard observations of the wind speed and direction.
2. A study of different reward functions for the UAV Target Tracking POMDP formulation.
3. An implementation of the POMDP formulation in real cluttered and GPS-denied scenarios under windy conditions.

This research has the potential to be used as a method for collecting data in challenging environments and for remote sensing and monitoring in the fields of biodiversity and ecology (including the tracking of moving animals).

This paper is organised as follows: Section 2 covers a description of POMDP and an online solver. Section 3 describes the system architecture. Section 4 describes the method used for calculating the update frequency of the POMDP solver. Section 5 describes the Navigation and Target Finding problem formulation as a POMDP. Results are discussed in Section 6, and Section 7 provides conclusions and future areas of research.

2. BACKGROUND

In real world scenarios, the UAV perception system is limited by the type of sensors and the environment in which it navigates and consequently there is uncertainty and error in the estimation of its localisation and orientation. Furthermore, depending on the mission scenario, there could be situations in which variables that compose the mission state space are not observable or can only be observed under certain conditions, such as the target location, speed and heading. This limitation in the sensory systems of a UAV and the inability to obtain perfect knowledge of the vehicle and mission state is also known as partial observability [22], [23], [24]. A method to model robotic and UAV sequential decision problems under uncertainty are Markov Decision Processes (MDP) and POMDP.

MDPs are used for robotic missions to generate a policy that allows the robot to decide what sequence of actions should be taken in order to maximise a return or cost function, taking into account the uncertainties in motion [25]. Plain MDPs assume that the states are completely observable which is not the case for a robot that has limitations in perception. POMDP, on the other hand, incorporates the uncertainties in sensing and the partial observability of the agent in the environment [26], [27].

POMDP

A POMDP is a tuple consisting of the following elements $(S, A, O, T, Z, R, \gamma)$ where S is the set of states in the environment, A is the set of actions the agent, in this case the UAV, can execute, O is the set of observations, T is the transition function between states after executing an action, Z is the distribution function describing the probability of observing o from state s after taking action a, R is the set of rewards for every state and γ is the discount factor. A feature of POMDPs is that the state of the process is not represented by a single value but by a probability distribution over all the possible states in its state-space representation at a particular time, this is known as a belief-state and is denoted by *belief* b.

In a POMDP the state of the agent cannot be observed exactly or completely, instead the agent receives observations $o \in O$. The perception of these observations could be represented by the probability distribution Z. The solution of a POMDP is a policy $\pi : \mathcal{B} \rightarrow \mathcal{A}$ that maps actions a to belief-states $b \in \mathcal{B}$, which is the set of possible belief-states. These belief-states are updated after receiving an observation based on the Bayes' theorem. Given the current belief-state b, the objective of a POMDP algorithm is to find an optimal policy π^* that maximizes a value function when following a sequence of actions and observations. The accumulated *discounted return* is the sum of the discounted rewards after executing every action in the sequence from time t onwards $R_t = \sum_{k=t}^{\infty} \gamma^{k-t} r_k$, where r_k is the immediate reward received at particular time step t for taking action a_t. The *Value function* is the expected return from belief-state b when following policy π, $V^{\pi}(b) = \mathbb{E}\left[\sum_{k=t}^{\infty} \gamma^{k-t} r_k | b, \pi\right]$. An optimal policy for the POMDP is the one that maximizes the value function $\pi^*(b) = \arg\max_{\pi} V^{\pi}(b)$.

The Adaptive Belief Tree

In order to test the target finding and tracking problem formulated as a POMDP, we implemented in hardware and software one of the fastest on-line POMDP algorithms to the authors knowledge, ABT [28]. This algorithm has been tested previously in simulations and onboard a UAV for navigation [13] and target finding [21] but not for a more challenging UAV mission such as Target Tracking in the presence of wind.

The Adaptive Belief Tree (ABT) is an online POMDP solver that uses Monte Carlo Simulations and a set of state particles to represent the belief-state. It generates a search tree to store the policy. The root of this tree is a node containing the state particles representing the initial belief-state of the environment.

In ABT, the policy is updated after receiving an observation. ABT does not clear the policy tree after receiving and matching an observation but instead improves the policy based on the observation received. Moreover, ABT provides a mechanism for accepting changes in the environment and adapts its policy accordingly. This algorithm was selected for the implementation of the Target Finding and Tracking mission based on its capacity to handle a large number of possible observations that result from a larger uncertainty in the target's location. ABT can model the observations using a dynamic approach in which the number of observations does not have to be fixed as a constant value for the whole episode. Thanks to this feature of ABT, the observations do not have to be discretised but instead a programmable function defines whether two observations are equal based on the value of their variables. The implementation of this function for the Target Finding and Tracking problem is explained in Section 4

3. SYSTEM ARCHITECTURE

The system is divided into four modules, shown in Fig. 1, one module for running the online POMDP solver, another module for controlling forward $\boldsymbol{V_{r_f}}$ and lateral velocity $\boldsymbol{V_{r_l}}$, yaw Ψ_r and altitude z_r of the multi-rotor, one module for calculating the multi-rotor position in the scenario $\boldsymbol{P_r} = (x_r, y_r, z_r, \Psi_r)$ based on the forward and lateral velocity, heading angle and altitude, estimated from the on-board sensors and the target position $\boldsymbol{P_t} = (x_t, y_t, z_t, \Psi_t)$, estimated by the target's pose in an image if detected by a downward looking camera. A fourth module, which is the Autonomy Lab driver for ROS [29], receives the commands for actuation in roll, pitch, and yaw angles as well as the altitude and sends the sensor readings upon request. All nodes running in parallel in different threads allow the system to have different update rates for each module and that permits different levels of controllers.

Online POMDP Module

The Online POMDP module executes an online POMDP solver algorithm, ABT in our example. This module implements the ABT algorithm in which the POMDP elements described in section IV are programmed using the object oriented programming interfaces available [30]. In this module the POMDP model and solver are initialised according to input parameters and an initial belief-state b_0. It first produces a policy based on b_0 and outputs the action a that

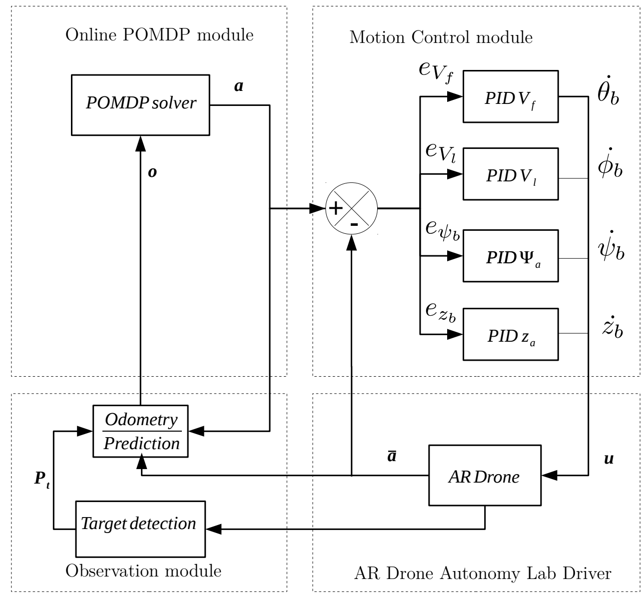


Figure 1. POMDP ROS System Architecture

maximises an expected return. The action is then executed by the motion control module, while the perception module calculates independently the robot position and heading angle and the target pose in case it is detected by the onboard downward camera. Based on the perception module data, an observation is created and received after the action is executed.

The online POMDP module updates the belief-state b, to match the obtained observation and replenishes particles until a time-out is reached. Based on the current belief-state b, the POMDP solver calculates and updates the policy and outputs the subsequent action based on the updated policy. The process repeats until a terminal state is reached, i.e. the target has been detected and followed for at least ten seconds.

Finding the right time-step duration for the implementation of the target finding and tracking problem as an online POMDP, depends on the dynamic capabilities of the UAV and on the complexity of the scenario. We set the online POMDP solver with a frequency of 1 Hz, since we know how the motion control system will respond within that time-frame and the online planner provides good policies within that time.

On one hand, having long time-steps for the policy update reduces the complexity for calculating a solution since the length of the planning horizon is reduced compared to a shorter time-step. On the other hand, having short time-steps reduces the uncertainty in motion and perception since the observation updates are more frequent. The POMDP online planner iteration time selected allows the UAV to perform the set of actions, while simultaneously it computes and updates the motion policy online based on perceived observations.

Motion Control Module

The block diagram presented in Fig. 2 shows the developed framework control structure. In this structure, there is one inner loop that controls the translational dynamics and which receives as input the action command vector $\boldsymbol{a}$ from the motion planner. The action command vector is also called the reference state vector $\boldsymbol{X}_b^* = \{\dot{x}_b, \dot{y}_b, \psi_b, z_b\}^*$ with the reference values for forward velocity $\dot{x}_b^*$, lateral velocity $\dot{y}_b^*$, heading angle Ψ_b^* and altitude z_b^*. The outer loop in the control structure is in charge of the motion planning which is calculated and updated by the motion planner module.

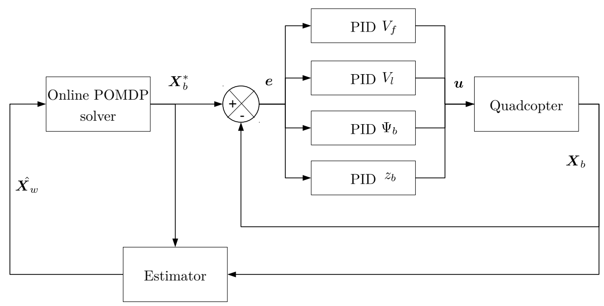


Figure 2. Structure of the Motion Control System.

The Motion Control module was designed to enable the framework to execute a set of motion commands that are produced by the motion planner module. This set of commands or actions emulate those that are given to a quadrotor UAV by an operator using a remote controller with a joystick. In this way the quadrotor UAV can move forward and backwards by controlling its forward velocity, it can move left or right by controlling its lateral velocity, it can ascend or descend by controlling its altitude and it can also rotate over the z axis by controlling its yaw angle. An on-board attitude controller maintains the UAV in a stable attitude which allows it to hover in a quasi-stationary position.

The set of commands that are used by the motion planner are performed by the UAV by actuating in four different states that are decoupled so that an independent controller can be designed to control them. The Motion Control module uses four independent PID controllers to control the following states of the UAV: forward velocity $\dot{x}_b$, lateral velocity $\dot{y}_b$, yaw angle (heading angle) ψ_b and altitude z_b. The Motion Control module receives the references for the UAV states $\boldsymbol{a} = \boldsymbol{X}_b^* = \{\dot{x}_b, \dot{y}_b, \psi_b, z_b\}^*$ from the motion planner module and subtracts the actual states $\boldsymbol{X}_b = \{\dot{x}_b, \dot{y}_b, \psi_b, z_b\}$ from the reference states $\boldsymbol{X}^*$ to generate error signals that are used by each of the PID controllers.

The output of the PID controllers is a control vector $\boldsymbol{u} = \{\theta_b, \phi_b, \dot{\psi}_b, \dot{z}_b\}$, where θ_b, ϕ_b are pitch and roll angles, respectively, $\dot{\psi}_b$ is yaw rate and $\dot{z}_b$ is rate of climb or descent. These outputs are sent to the AR Drone Autonomy Lab driver, which transforms them into control signals for the UAV onboard attitude controller.

Each of the PID controllers has a feedback loop in which the UAV on-board navigation system calculates the forward $\dot{x}_b$ and lateral $\dot{y}_b$ velocities using optical flow obtained from the downward looking camera. The UAV obtains the yaw angle ψ_b from IMU and magnetometer readings and the altitude z_b from on-board ultrasonic and barometric pressure sensors that are fused using a proprietary Kalman filter.

Perception Module

The perception module is also implemented in ROS and executes at 100 Hz. Current forward and lateral velocities are calculated by fusion of optical flow (using a downward looking camera) and inertial measurement unit readings, Yaw angle is calculated based on onboard IMU and altitude is sensed by combining IMU, ultrasonic sensors and barometric sensor.

This module is constantly calculating the current multi-rotor position based on the sensed forward V_{r_f} and lateral V_{r_l} velocities and the heading angle Ψ_r. It converts the forward and lateral velocities in the multi-rotor's frame to the world frame, and calculates x_r and y_r positions, using Equation (1). It also reads the altitude or z_r position from the on-board sensors.

In order to have an accurate source of global localisation the system uses Augmented Reality tags that can be detected by the front camera onboard the UAV. These tags are placed on one face of the obstacles and its position in the world frame is known to the system so that when detected they provide an accurate measuring of the UAV pose in the world frame. These tags can only be detected when the UAV is within 1 m in front of the tag. The perception module updates the observation once they are detected by the UAV front camera.

$$\begin{bmatrix} x_{r_{t+1}} \\ y_{r_{t+1}} \end{bmatrix} = \begin{bmatrix} x_{r_t} \\ y_{r_t} \end{bmatrix} + \begin{bmatrix} \cos(\Psi_{r_t}) & -\sin(\Psi_{r_t}) \\ \sin(\Psi_{r_t}) & \cos(\Psi_{r_t}) \end{bmatrix} \begin{bmatrix} V_{r_{f_t}} \\ V_{r_{l_t}} \end{bmatrix} \Delta t \quad (1)$$

4. TARGET FINDING AND TRACKING

Problem Description and Formulation

We illustrate the scenario for a target finding and tracking mission in Figure 3. A multi-rotor UAV flies in a confined 3D space filled with obstacles. The UAV does not have GPS localisation and there is wind in the scenario. The mission of the UAV is to search and find a ground moving target, and follow it for at least 10 s. The target moves on the ground in any direction at an irregular speed. A map of the environment with the obstacles locations is known to the system and is input as a parameter of the algorithm as a text file.

After take off, the UAV hovers for a few seconds to initialise its orientation with readings from its on-board sensors. There is some initial drift produced in this initial hovering position. A normal probability distribution with mean value around the take off position is used to model this initial uncertainty.

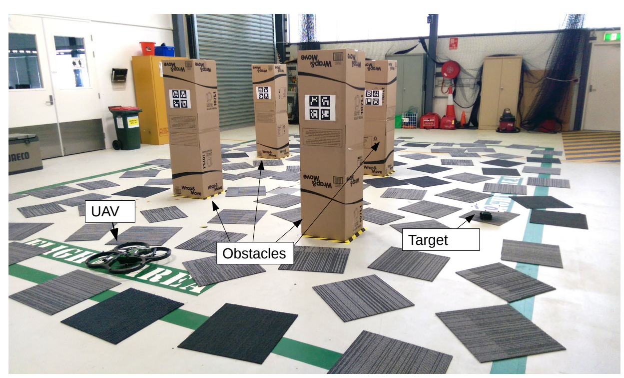


Figure 3. Target Finding and Tracking scenario

The problem is formulated as a POMDP that has the following elements: the state of the aircraft in the environment (S), the set of actions that the multi-rotor can execute (A), the transition function describing the state transition after applying a specific action (T), the observation model that represents the sensed state of the aircraft after taking an action (O), and the reward and cost function (R).

State Variables (S)

The state variables considered in the POMDP formulation are the quad-rotor position and heading $\boldsymbol{P_r} = (x_w, y_w, z_w, \Psi_w)$, the target planar position and heading $\boldsymbol{P_T} = (x_T, y_T, \Psi_T)$, target forward velocity $\boldsymbol{V_T}$ and the UAV's forward velocity $\dot{x_b}$ and lateral velocity $\dot{y_b}$ in the body frame. The wind speed $\boldsymbol{V_w}$ and orientation W_ψ are all measured in the world frame.

Actions (A)

The UAV actions are designed to be the set-points for the four PID controllers, thus the UAV actuates in four state variables, namely, forward $\dot{x_b}$ and lateral $\dot{y_b}$ velocities, heading angle Ψ_b and altitude z_b.

The set of actions consists of 7 actions. An action to keep the aircraft hovering, i.e. $\dot{x_b}, \dot{y_b} = 0\ m/s$; two actions to go forward and backward, with current heading angle Ψ_w, lateral velocity $\dot{y_b} = 0\ m/s$, and forward velocity, $\dot{x_b} = 0.6\ m/s$ and $\dot{x_b} = -0.6\ m/s$, respectively. Actions Up and Down, increase or decrease altitude z_w in $0.3\ m$, respectively, with multi-rotor velocity fixed at $0\ m/s$, and two actions to roll left and right with current heading angle Ψ_w and forward velocity at $\dot{x_b} = 0\ m/s$, and lateral velocity at $\dot{y_b} = \pm 0.6\ m/s$, respectively.

Transition Function (T)

The transition function (T) uses the discretised transient responses of the four PID controllers in order to predict the next state of the process. The actions are step inputs or references to the four states controllers. The transient responses of the PID controllers to step input commands are acquired experimentally and are incorporated into the kinematic model of the aircraft using a decoupled model.

The following UAV model was derived taking into account the four states in the UAV that are controlled by the Motion Control module and the different frames that take place in the UAV mission, which are the world frame, the UAV frame, the downward looking camera frame, the front looking camera frame and the image frame. The UAV dynamic model equations are derived in terms of the UAV state vector and the Motion Control vector. The UAV state vector $\boldsymbol{X_w} = \{x_w, y_w, z_w, \psi_w\}$ is composed by the UAV position and its heading and are measured in the world frame, the roll and pitch angles are not relevant for the localisation of the UAV since they are not used in the calculation for the prediction of the next states. The control input vector is composed of the UAV forward and lateral velocities, the heading angle and the altitude, all measured in the body frame.

$$\begin{bmatrix} \Delta x_{w_t} \\ \Delta y_{w_t} \\ \Delta z_{w_t} \end{bmatrix} = \begin{bmatrix} \cos(\psi_{b_t} + \sigma_{b_t}) & -\sin(\psi_{b_t} + \sigma_{b_t}) & 0 \\ \sin(\psi_{b_t} + \sigma_{b_t}) & \cos(\psi_{b_t} + \sigma_{b_t}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} (\dot{x_{b_t}} + V_{w_x})\Delta t \\ (\dot{y_{b_t}} + V_{w_y})\Delta t \\ \Delta z_b \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} x_{w_{t+1}} \\ y_{w_{t+1}} \\ z_{w_{t+1}} \end{bmatrix} = \begin{bmatrix} x_{w_t} \\ y_{w_t} \\ z_{w_t} \end{bmatrix} + \begin{bmatrix} \Delta x_{w_t} \\ \Delta y_{w_t} \\ \Delta z_{w_t} \end{bmatrix} \tag{3}$$

Where x_{w_t}, y_{w_t} and z_{w_t} are the x, y and z aircraft world coordinates at time t, x_{b_t} and y_{b_t} are forward and lateral velocities in the body frame at time t, and ψ_{b_t} and σ_{b_t} are heading and heading deviation at time t.

A Gaussian distribution is used to model the uncertainty in motion. The UAV can deviate from a commanded heading due to external disturbances and error in the heading control system caused by drift in the magnetometer and heading angle sensor. Eq. (4) calculates the probability of the distribution.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{4}$$

The transient responses for the four controlled states are discretised and included in the transition function as lookup tables in order to approximate the dynamics of the quadrotor.

We used a radio controlled toy car moving on the ground with a marker on top as the target. The target's motion is described by equations 5 and 6. The target's velocity V_{T_t} has an uncertainty of $0.2\ m/s$ around an average value $\mu = 0.5\ m/s$. The target motion is model using two modes which are 1) Loop clockwise around the obstacles in the flying area which changes the heading of the target Ψ_{T_t}, alternating among north, east, south and west, and 2) Continue with same heading with a 50% probability and 50% probability of changing its orientation Ψ_{T_t} to the opposite direction of the UAV in order to go away from it.

$$x_{T_{t+1}} = x_{T_t} + V_{T_t}\cos(\Psi_{T_t})\Delta t \tag{5}$$

$$y_{T_{t+1}} = y_{T_t} + V_{T_t}\sin(\Psi_{T_t})\Delta t \tag{6}$$

Observation Model (O)

An observation is composed of the following variables: 1) the UAV position in the world frame, 2) the target's position and orientation if it is detected by the downward looking camera, and 3) boolean variables indicating whether the target has been detected and whether or not an AR tag is detected by the front camera. The system receives an odometry reading from the perception module with an uncertainty that is caused by the accumulating error and drift in the yaw angle reading. This uncertainty is represented by adding noise to the UAV position using a gaussian distribution as in equation (4), with mean value μ around the perceived position and standard deviation $\sigma = 0.5\ m$.

This uncertainty can be reduced, i.e. $\sigma = 0.05m$ if the UAV flies in front of an obstacle where it can detect an AR tag and can reset the error inherent to the odometry and onboard sensors.

If the target is detected by the onboard downward looking camera, the detection system provides the target position within the image. This position is transformed into a position in the world frame. A model of the downward camera field of view is also included in order to extract the target pose in the world frame from the target's position and orientation in the image.

In order to verify if two observations are equal, the euclidean distance between the UAV position of both observations should be less than a threshold value that represents the error in the odometry system and the boolean variable that indicates whether a target has been found should have the same value for both observations.

$$Obs_1 = Obs_2 \text{ if } \|P_{r_{Obs_1}} - P_{r_{Obs_2}}\| \leq \epsilon \quad (7)$$
$$\text{and } TO_{Obs_1} = TO_{Obs_2} \quad (8)$$

where $P_{r_{Obs_1}}$ is the quadrotor position in observation 1 and TO_{Obs_1} is the variable that indicates whether a target has been found in observation 1.

The UAV generates estimates of the wind speed and orientation based on the compensation executed in order to maintain its state. This wind estimation generates an observation which updates the belief of the prevailing wind in the scenario.

Rewards and Costs Function (R)

Two different reward function structures were tested. In the first structure R_1 shown in Eq.9 the multi-rotor receives a high reward if it detects the target within the downwards looking camera field of view. Hitting an obstacle or going out of the scenario incur a penalty and every other movement will carry a small cost with the purpose of generating shorter sequences of actions and thus shorter paths. The second reward function R_2 is shown in Eq. 10. It differs from R_1 in that a reward for being close to the target is included in the function. The values of the reward and cost functions were selected as a result of tuning the system to be able to accomplish the mission faster.

$$R_1 = \begin{cases} -10 & \text{for every UAV action} \\ -70 & \text{if UAV out of scenario} \\ -70 & \text{if UAV collides} \\ 500 & \text{if target detected} \end{cases} \quad (9)$$

$$R_2 = \begin{cases} -10 & \text{for every UAV action} \\ -70 & \text{if UAV out of scenario} \\ -70 & \text{if UAV collides} \\ \frac{1}{d_g} & d_g \text{ is distance from UAV to target} \\ 500 & \text{if target detected} \end{cases} \quad (10)$$

5. Results and Discussion

We conducted simulation and real flight tests in order to analyse the performance of the system.

Simulation

The UAV target finding and tracking mission was tested in simulation for 100 runs for each of the three cases shown in table 1 to get an average for the discounted return, the number of steps and the success in accomplishing the mission, with a maximum of 120 steps which corresponds to 2 min flight time.

Table 1. Simulation results for Target Tracking problem

Target motion mode	**Success rate (%)**	**Flight time to target (s) (Number of steps)**	**Reward structure**
Loop mode	100	41	R_1
Loop mode	100	32	R_2
Loop mode (Wind)	100	56	R_1
Loop mode (Wind)	100	45	R_2
Escaping mode	81	93	R_1
Escaping mode	87	84	R_2
Escaping mode (Wind)	78	93	R_1
Escaping mode (Wind)	86	84	R_2

Simulation results indicate that the system is able to find a target and follow for all trials in the loop target mode. If the target motion is mostly deterministic, as in the loop mode, the UAV spends fewer steps to achieve a terminal state and the return is higher. On the other hand, when the uncertainty in the target motion increases, such as in the escaping mode, the system takes more steps to accomplish the mission and its success ratio decreases.

Moreover, the results also show that the reward structure R_2 performs better than R_1. This structure rewards the UAV for getting close to the believed target position and after the target has been detected, it ensures that the UAV stays close to the target in order to track its motion.

We use the graphical tool RVIZ for visualisation (Fig. 4). Point clouds are used to visualise the distribution of the belief-state particles representing the target location (red particles) and UAV location (white particles). The target starts moving from four possible locations. The UAV is initially located at coordinates $(-2.9, 0, 0.0)$ and takes off and hovers for some seconds which increases the uncertainty in its position. Fig. 4 shows a trial in which there is no wind disturbance affecting the UAV motion. In Fig. 5 the UAV motion gets affected by the wind blowing in the northerly direction. It can be seen in the figure that the UAV takes a shorter path towards the upper right corner taking advantage of the wind influence in order to anticipate the target location. In Fig. 6 the wind is heading west. The influence of the wind direction can be seen in the UAV trajectory which avoids colliding into obstacles while tracking the target.

In the target escaping mode, the system and POMDP solver guide the multi-rotor towards the corners of the flying space, see Fig. 7. This happens because the POMDP model predicts that the target is escaping in the opposite direction of the UAV, which increases the probability of the target to go towards the corners.

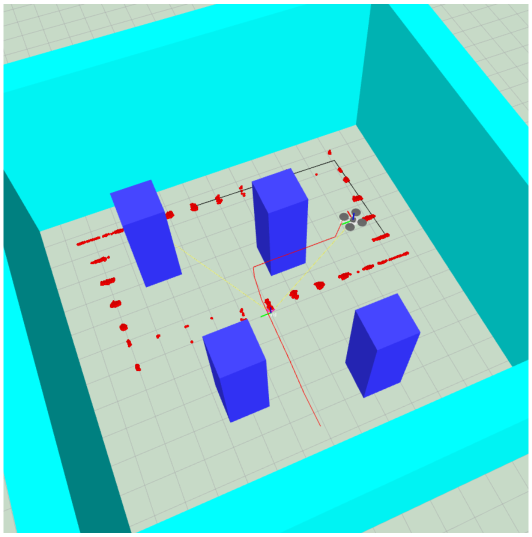


Figure 4. **(a) Belief-state. UAV trajectory (orange) and target trajectory (black).**

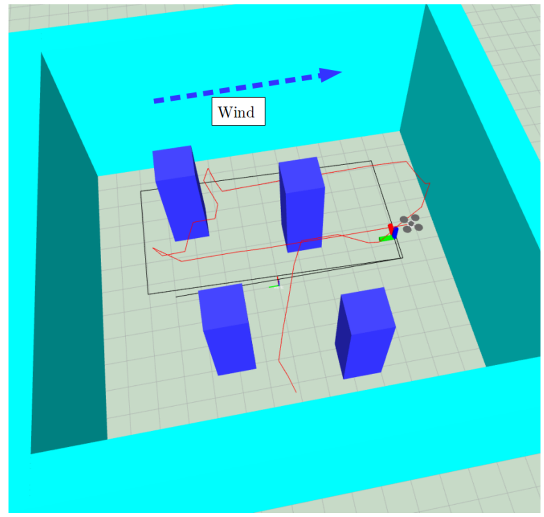


Figure 6. **(c) Wind blows east. UAV trajectory (orange) and target trajectory (black).**

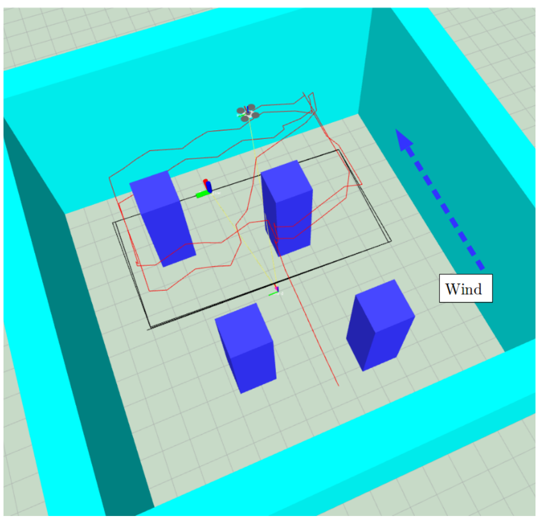


Figure 5. **(b) Wind blows north. UAV trajectory (orange) and target trajectory (black).**

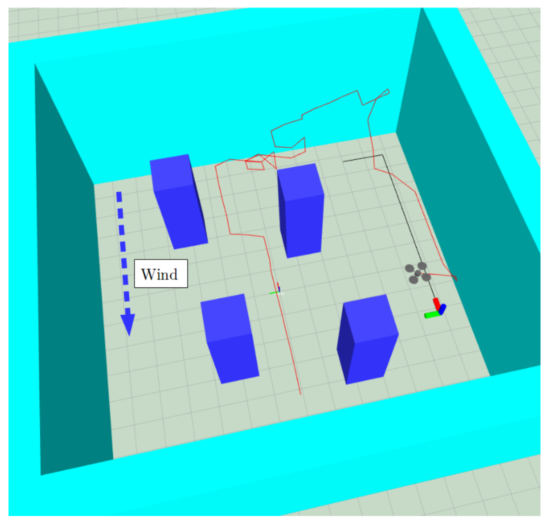


Figure 7. **(d) Target escaping, wind to south. UAV trajectory (orange) and target trajectory (black).**

Real Flight Tests

Real flight tests were conducted indoors to explore the performance of the system. Industrial fans were used to emulate outdoor windy conditions. The experiments were conducted using a low cost commercial platform, the Parrot AR Drone, with four obstacles, Fig. 3.

Real flight tests were conducted 20 times for each of the target modes. Table 2 shows that the system successfully completes the mission 100% of the time for the loop mode. On the other hand, the system has more difficulty finding a mobile target that is trying to escape and completes the mission 48% of the time. The difference in success rate compared to the simulation results is due to the nature of the target motion. The target motion gets affected by the tiles that are placed on the ground to generate texture. Moreover, the wind speed is assumed to be constant throughout the whole flying scenario. This creates a larger stochastic component on the target motion and the wind that are difficult to model. Furthermore, the target is a toy car that is driven remotely by a person which also contributes to generating an unpredictable component to its motion.

An example of the paths followed by the UAV and target for the loop mode is shown in Fig. 8. In this case the UAV flies first near an obstacle to reduce the uncertainty in its position and then it ascends to have a wider FOV and hovers above the obstacle waiting to detect the target that will eventually pass under the camera. Afterwards, the UAV follows the target until it is detected 10 times and finishes the mission.

Once the system has found the target it is also able to recapture it if the target goes out of the field of view of the camera. The POMDP formulation allows the system to predict the target's next possible location and to navigate towards it to detect it. This situation is shown in Fig. 8. It can be seen that the UAV was able to find, detect and track the target in a real flight scenario.

Table 2. Real tests results for Target Tracking problem

Target motion mode	Success rate (%)	Flight time to target (s) (Number of steps)	Reward structure
Loop mode	100	54	R_1
Loop mode	100	43	R_2
Loop mode (Wind)	87	71	R_1
Loop mode (Wind)	92	69	R_2
Escaping mode	48	112	R_1
Escaping mode	58	98	R_2
Escaping mode (Wind)	42	121	R_1
Escaping mode (Wind)	45	106	R_2

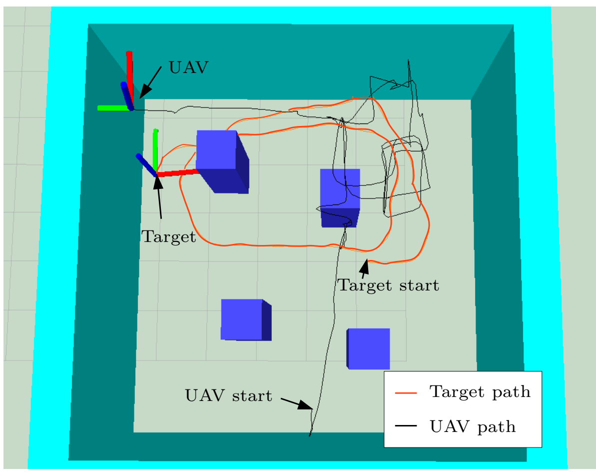


Figure 8. Real flight UAV trajectory (black) and Target trajectory (orange)

6. Conclusions

A system for target finding and tracking with a UAV using POMDP in GPS-Denied and cluttered environments was presented and demonstrated in simulation and flight test. This system is able to find, detect and track a target with a UAV flying within a confined space using an onboard downwards looking camera and in the presence of obstacles and wind.

The system presented uses a state-of-the-art POMDP on-line algorithm that outputs actions instead of waypoints, and only relies on on-board sensors for the localisation of both the UAV and the target. This allows modelling the system's dynamics using the motion controller responses for each decoupled state, in this case forward and lateral velocities, altitude and yaw angle.

The system is also able to find, track and follow a target in a real scenario when the target's motion model is mostly deterministic. The system is capable of following a target when it is pursued by an UAV for cases where the target motion is mostly deterministic. In cases where the target is lost, the UAV can predict the target motion and reacquire the target's position.

A new reward structure was presented as a function of the distance from the UAV to the target, assigning a positive reward to the UAV for staying close to the target. This new reward structure outperforms a former reward structure that only assigns a positive reward for detecting the target under the onboard camera FOV.

Current work focuses on testing the system in dynamic environments where there is uncertainty in the obstacles locations.

Future work will concentrate on the use of this enabling technology to enhance the use of a UAV to navigate in challenging and remote scenarios for data collection, remote sensing and monitoring of biodiversity and ecology at the Sanford Ecological Research Facility (SERF), QLD, Australia.

Acknowledgments

The authors thank the Australian Research Centre for Aerospace Automation for its resources and indoor flying facilities. We thank Dr. Hanna Kurniawati (UQ), for support given on POMDP and the open source software ABT.

References

[1] A. Malaver, F. Gonzalez, and N. Motta, "Towards the development of a gas sensor system to monitoring pollutant gases in the low troposphere using small unmanned aerial vehicles," *Environmental Monitoring*, pp. 9–11, 2012.

[2] F. Gonzalez, M. P. G. Castro, P. Narayan, R. Walker, and L. Zeller, "Development of an autonomous unmanned aerial system to collect time-stamped samples from the atmosphere and localize potential pathogen sources," *Journal of Field Robotics*, vol. 28, no. 6, pp. 961–976, 2011. [Online]. Available: http://dx.doi.org/10.1002/rob.20417

[3] S. Waharte and N. Trigoni, "Supporting search and rescue operations with UAVs," in *Proceedings - EST 2010 - 2010 International Conference on Emerging Security Technologies, ROBOSEC 2010 - Robots and Security, LAB-RS 2010 - Learning and Adaptive Behavior in Robotic Systems*. IEEE, sep 2010, pp. 142–147.

[4] K. Anderson and K. J. Gaston, "Lightweight unmanned aerial vehicles will revolutionize spatial ecology," *Frontiers in Ecology and the Environment*, vol. 11, no. 3, pp. 138–146, 2013.

[5] L. F. Gonzalez, G. A. Montes, E. Puig, S. Johnson, K. Mengersen, and K. J. Gaston, "Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife monitoring and conservation," *Sensors (Switzerland)*, vol. 16, no. 1, 2016.

[6] P. Lommel, M. W. McConley, and N. Roy, "Robust path planning in GPS-denied environments using the Gaussian augmented markov decision process," *Navigation*, vol. 12, no. 3, p. 1.

[7] R. He, S. Prentice, and N. Roy, "Planning in information space for a quadrotor helicopter in a GPS-denied environment," in *Proceedings - IEEE International Conference on Robotics and Automation*. IEEE, 2008, pp. 1814–1820.

[8] L. F. Gonzalez, "Robust evolutionary methods for multi-objective and multdisciplinary design optimisation in aeronautics," *University of Sydney*, 2005.

[9] D. S. Lee, L. F. Gonzalez, and E. J. Whitney, "Multi-obejctive," *Multidisciplinary Multi-fidelity Design tool: HAPMOEA-User Guide*, 2007.

[10] D. Lee, J. Periaux, and L. F. Gonzalez, "UAS Mission Path Planning System (MPPS) Using Hybrid-Game Coupled to Multi-Objective Optimizer," *Journal of Dynamic Systems, Measurement, and Control*, vol. 132, no. 4, p. 041005, 2010.

[11] D. Lee, L. F. Gonzalez, J. Periaux, K. Srinivas, and E. Onate, "Hybrid-Game Strategies for multi-objective design optimization in engineering," *Computers and Fluids*, vol. 47, no. 1, pp. 189–204, 2011.

[12] W. Al-Sabban, L. Gonzalez, and R. Smith, "Wind-Energy based Path Planning For Unmanned Aerial Vehicles Using Markov Decision Processes," in *ICRA*, 2013, pp. 1–6.

[13] F. Vanegas and F. Gonzalez, "Enabling UAV navigation with sensor and environmental uncertainty in cluttered and GPS-denied environments," *Sensors (Switzerland)*, vol. 16, no. 5, 2016.

[14] S. Zhu, D. Wang, and C. B. Low, "Ground target tracking using UAV with input constraints," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 69, no. 1-4, pp. 417–429, jan 2013.

[15] D. Hsu, W. S. Lee, and N. Rong, "A point-based POMDP planner for target tracking," in *ICRA*. IEEE, 2008, pp. 2644–2650.

[16] J. Hu, L. Xie, J. Xu, and Z. Xu, "Multi-agent cooperative target search," *Sensors (Switzerland)*, vol. 14, no. 6, pp. 9408–9428, may 2014.

[17] U. Zengin and A. Dogan, "Real-time target tracking for autonomous UAVs in adversarial environments: A gradient search algorithm," *IEEE Transactions on Robotics*, vol. 23, no. 2, pp. 294–307, apr 2007.

[18] N. Farmani, L. Sun, and D. Pack, "Tracking multiple mobile targets using cooperative Unmanned Aerial Vehicles," in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2015, pp. 395–400.

[19] R. He, A. Bachrach, and N. Roy, "Efficient planning under uncertainty for a target-tracking micro aerial vehicle," in *Proceedings of IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 1–8.

[20] C. C. Chanel, "Multi-Target Detection and Recognition by UAVs Using Online POMDPs," *AAAI. 2013*, pp. 1381–1387, 2013.

[21] F. Vanegas and F. Gonzalez, "Uncertainty based online planning for UAV target finding in cluttered and GPS-denied environments," *IEEE Aerospace Conference Proceedings*, vol. 2016-June, 2016.

[22] N. Dadkhah and B. Mettler, "Survey of motion planning literature in the presence of uncertainty: Considerations for UAV guidance," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 65, no. 1-4, pp. 233–246, jan 2012.

[23] H. Kurniawati, Y. Du, D. Hsu, and W. S. Lee, "Motion planning under uncertainty for robotic tasks with long time horizons," in *Springer Tracts in Advanced Robotics*, vol. 70, no. STAR, 2011, pp. 151–168.

[24] S. C. W. Ong, Shao Wei Png, D. Hsu, and Wee Sun Lee, "Planning under Uncertainty for Robotic Tasks with Mixed Observability," *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1053–1068, 2010.

[25] C. H. Papadimiitriou and J. N. Tsitsiklis, "The Complexity of Markov Decision Processes," pp. 441 – 450, 1987.

[26] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT press Cambridge, MA, 2005, vol. 1.

[27] J. Pineau, G. Gordon, and S. Thrun, "Anytime Point-Based Approximations for Large POMDPs." *Jounal of Artificial Intelligence Research*, vol. 27, no. 1, pp. 335–380, 2006.

[28] H. Kurniawati and V. Yadav, "An Online POMDP Solver for Uncertainty Planning in Dynamic Environment," *International symposium of robotics research*, pp. 1–16, 2014.

[29] M. Monajjemi, "AR Drone Autonomy Lab," 2014. [Online]. Available: http://ardrone-autonomy.readthedocs.io/en/latest/

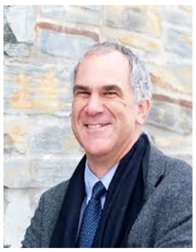
[30] D. Klimenko, J. Song, and H. Kurniawati, "TAPIR: A Software Toolkit for Approximating and Adapting POMDP Solutions Online," *Proceedings of the 2014 Australasian Conference on Robotics and Automation*, 2014.

BIOGRAPHY

Fernando Vanegas *received his B.S. in Mechatronics Engineering from UMNG in 2004 and M.Sc. in Electrical Engineering from Halmstad University in 2008. He is currently a Ph.D. candidate in Robotics and Autonomous Systems at The Australian Research Centre for Aerospace Automation and Queensland University of Technology. His current research activities include motion planning for UAV in cluttered and uncertain environments modelled as POMDPs.*

Duncan Campbell *is the Director of QUTs Australian Research Centre for Aerospace Automation (ARCAA) in the Robotics and Autonomous Systems Discipline with the School of Electrical Engineering and Computer Science. He has 25 years of research leadership in control and automation. Duncan serves as Treasurer on the Board of the Australian Association for Unmanned Systems (AAUS) and was the IEEE Queensland Section Chapter Chair of the Control Systems/Robotics and Automation Society Joint Chapter (2008/2009). He is the Chair of CDIO Australian and New Zealand regional group of the global CDIO collaboration in engineering education, and was President of the Australasian Association for Engineering Education in 2011.*

Nicholas Roy *is an associate professor in the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL). His research has focussed specifically on the problems that result from uncertainty in the world, such as sensor noise or unpredictable action outcome. Probabilistic, decision-theoretic models have proven to be ideally suited for state estimation in the face of uncertainty. He has formulated algorithms for a class of planning models called Partially Observable Markov Decision Processes (POMDPs).*



Kevin Gaston *is Professor of Biodiversity and Conservation, and founding director of the Environment and Sustainability Institute in the University of Exeter. He leads basic, strategic and applied research in ecology and conservation biology, with particular emphases at present including common ecology, ecosystem goods and services, land use strategies, and urban ecology.*

Felipe Gonzalez *is an Associate Professor (UAV) in the Science and Engineering Faculty, QUT and Theme leader for UAV Remote Sensing at ARCAA. He holds a BEng (Mech) and a PhD from the University of Sydney. His research explores bioinspired optimization, uncertainty based UAV path planning and UAV for environmental monitoring. He leads the CRC Plant Biosecurity project Evaluating Unmanned Aerial Systems for Deployment in Plant Biosecurity and the Green Falcon Solar Powered UAV. Felipe is a Chartered Professional Engineer and member of professional organizations including the RAeS,IEEE and AIAA.*