A Framework for UAV Navigation and Exploration in GPS-denied Environments

Fernando Vanegas
Queensland University of Technology (QUT)
Robotics and Autonomous Systems
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 4593
f.vanegasalvarez@qut.edu.au

Kevin J. Gaston
Environment & Sustainability Institute, University of Exeter
Penryn Campus, Penryn, Cornwall TR10 9FE, UK
+44 (0) 1326 255810
k.j.gaston@exeter.ac.uk

Jonathan Roberts
Queensland University of Technology (QUT)
Robotics and Autonomous Systems
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 2437
jonathan.roberts@qut.edu.au

Felipe Gonzalez
Queensland University of Technology (QUT)
Robotics and Autonomous Systems
2 George St
Brisbane, QLD 4000, Australia
+61 (0) 7 3138 1363
felipe.gonzalez@qut.edu.au

Abstract— **Unmanned Aerial Vehicles (UAV) are increasingly used in a wide range of applications such as civil infrastructure inspection, agriculture, ecology and remote sensing. UAV autonomous operation relies on the use of GPS in order to localise and plan a mission, however there are places in which GPS positioning is limited or not available, such as in urban and natural canyons or below the canopy. This paper presents the development of a framework that enables a drone to navigate in an unknown, unstructured and GPS-denied environment. The aim of this research is to combine the use of localisation algorithms such as Simultaneous Localisation and Mapping (SLAM) with Partially Observable Markov Decision Processes (POMDP) algorithms into a framework in which the navigation and exploration tasks are modelled as sequential decision problems under uncertainty. The framework tested in simulation allows the UAV to navigate safely, avoiding collisions whilst guiding exploration in order to create an occupancy map of the UAV's surroundings. The proposed system guides the UAV through a series of actions in order to maximise the information gain about the unknown environment. The implementation of the proposed framework will enable the use of UAV for autonomous navigation and exploration in challenging environments where GPS positioning is not available or limited.**

TABLE OF CONTENTS

1. INTRODUCTION

Unmanned Aerial Vehicles (UAV) are increasingly being used in a wide range of applications such as civil infrastructure inspection, agriculture, ecology and remote sensing. For these types of applications, a UAV can be programmed to autonomously fly a predefined path composed by a series of GPS way-points, provided the UAV has a clear view of the sky in order to obtain a reliable GPS positioning. However, there exist challenging scenarios in which the UAV is required to fly under structures, such as bridges, within mines and buildings or below dense canopy, or in natural or urban canyons in which the GPS signal is partially or completely obstructed.

The aim of this research is to combine the use of localisation algorithms such as SLAM [1], [2] and/or Visual Odometry (VO) [3], [4] with POMDP [5], [6] algorithms into a framework in which the navigation and exploration tasks are modelled as sequential decision problems under uncertainty. The framework models the UAV operation in challenging GPS-denied and unstructured environments. In this regard, POMDPs can generate a motion plan for the UAV navigation, considering the uncertainty generated by the absence of GPS positioning and the motion model of the UAV. This leads to feasible and safe trajectories in order to create an occupancy map. The planning algorithm outputs motion commands in order to guide the UAV to explore unknown regions whilst avoiding collisions with obstacles.

The octomap library [7] is used in order to create and update an occupancy map that is queried by the POMDP planner in order to plan a sequence of motion commands. These motion commands are sent to the UAV flight controller using mavlink through mavros messages using the robotic operating system ROS.

The framework is simulated using the simulation environment Gazebo, with a simulated 3DR iris UAV platform using a PX4 dronecode firmware.

This research builds upon previous work [8], [9] and extends the capabilities with the aim to have a UAV navigation system that can be deployed in an unstructured and GPS-denied environment.

978-1-5386-6854-2/19/$31.00

2. BACKGROUND

POMDP

In this work, we use the POMDP framework to model UAV navigation and exploration in GPS-denied environments as a sequential decision problem under uncertainty. This formulation consists of the following elements $(S, A, O, T, Z, R, \gamma)$ where S is the set of states the UAV can have while performing its mission. A is the set of actions or motion commands that the UAV can execute. O is the set of observations the UAV can perceive. 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. In a POMDP, the state of the process is not defined by a single estimation at any given moment, but instead, it is represented by a probability distribution over possible states, which is known as a belief-state and is denoted by b, with $\mathcal{B}$ the set of all possible beliefs.

The solution of a POMDP is an optimal policy $\pi^* : \mathcal{B} \rightarrow \mathcal{A}$ that maps actions a to belief-states $b \in \mathcal{B}$. 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 onward $R_t = \sum_{k=t}^{\infty} \gamma^{k-t} r_k$, 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 $\pi^*(b) = \arg\max_\pi V^\pi(b)$.

POMDP solver software

We use the TAPIR [10] software which uses the ABT [11] POMDP algorithm to model the UAV navigation and exploration sequential decision problem. We selected ABT because of its capacity to model changes in the environment of a POMDP and to adjust the computed policy to a new environment. In our formulation, we use an occupancy map that changes and is updated after every step iteration with the new information gathered by a depth sensor. These changes in the map and the detection of obstacles and unknown spaces are modelled in the POMDP formulation.

3. PROBLEM FORMULATION

A possible scenario for application of the proposed system is shown in Fig. 1. In this scenario a UAV has the mission to navigate and generate an occupancy map in order to avoid collisions. The objective of the mission is to maximise the information gain about the collapsed structure whilst navigating safely through it. Furthermore, the UAV navigates with certain degree of uncertainty in its localisation. This degree of uncertainty depends on the positioning system used and the type of environment that is being explored. In this formulation, we aim to model the uncertainty in the positioning system which for GPS-denied environments could be a SLAM or VO algorithm solution.

State variables (S)

The POMDP formulation considers the following variables to represent the state of the process:

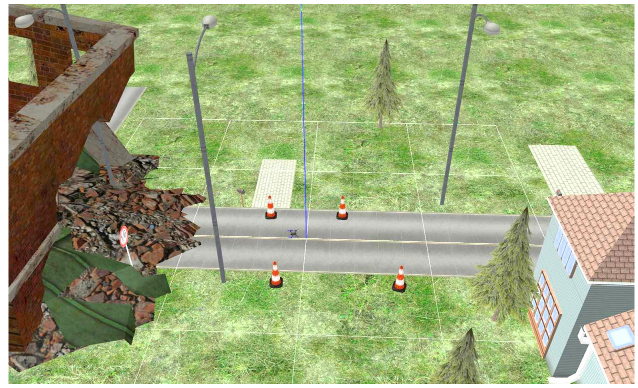


Figure 1. Simulation environment

Table 1. UAV actions in POMDP model

Action	$\dot{x_b}(m/s)$	$\dot{y_b}(m/s)$	$\Delta\Psi_b(\text{deg})$	$\Delta z_b(m)$
Hover	0	0	0	0
Forward	1	0	0	0
Turn left	0	0	45	0
Turn right	0	0	-45	0
Up	0	0	0	0.3
Down	0	0	0	-0.3

► the quad-rotor position and heading angle measured in a local frame $\boldsymbol{P_{UAV}} = (x_L, y_L, z_L, \Psi_L)$,
► the UAV's forward velocity $\dot{x_b}$ and lateral velocity $\dot{y_b}$ in the body frame,
► and an obstacle detection variable ob in front of UAV provided by a depth sensor

In this formulation, we assume that the UAV navigates in an environment where the wind disturbances are minimal. The UAV is also flying at more than 1 m above the ground where ground effects can be neglected.

Actions (A)

The actions in the POMDP formulation are selected based on the dynamic capabilities of the UAV. In this regard, they are designed as motion commands to set four state variables, namely, forward $\dot{x_b}$ and lateral $\dot{y_b}$ velocities, heading angle Ψ_b and altitude z_b. The motion commands are set-points or step inputs to PID controllers for each controlled variable. They are designed in a way that the controller reaches a steady state within the duration of action execution. The action execution duration is 1 second, during this time the POMDP planner can recompute the policy after receiving observations and update the map of the environment.

The set of actions consists of seven actions, namely, Hover, Forward, Turn left, Turn right, Up and Down. Table 1 describes the set-points to the UAV motion control system for each of these.

Observations (O)

The observations of the environment perceived by the UAV are represented by two variables. The first variable is its on-board estimated pose measured in the local frame $\boldsymbol{\hat{P}_{UAV}}$. Internal sensors such as IMU, GPS, compass, barometric pressure, optical flow and point lidar are used in this estimation. In environments where no GPS is present, the estimation of the local pose can benefit from depth sensors and cameras and a SLAM algorithm that uses landmarks or texture-rich features in the scene.

A second boolean variable $\hat{ob}$ indicates when a landmark is detected by the UAV front camera and/or SLAM system and is part of the observation. The detection of a landmark allows the UAV to reduce the uncertainty in its localisation. This detection is modelled in Algorithm 1. In this algorithm, we model how the depth sensor is used to discover obstacles in front. The algorithm uses the *castray* function [7] to detect occupied spaces in front of the UAV. An origin point, direction and range are given to the function in order to check whether the space in front of the sensors is marked as occupied or not in the map.

Algorithm 1 ExploreSpace($\boldsymbol{P_{UAV}}$, OccupancyMap)

```
T_W^b = ObtainTransform(P_UAV)
Ψ_UAV_b = TransformOrientation(P_UAV, T_W^b)
if CastRay(P_UAV, Ψ_UAV_b, range) then return ob̂ = true
else
    if not node in octomap then return ob̂ = false
    end if
end if
```

Transition function (T)

The transition function models the UAV dynamic response to motion commands that are represented as step inputs. The actions in the POMDP formulation are reference points for four independent controllers on board the UAV. Characteristic responses for each of these controllers are found experimentally and these responses are discretised into look-up tables. These look-up tables are then used to predict the response for each of the controlled states. However, a stochastic component is added into Equation 2 as a deviation from the commanded yaw angle. A Gaussian distribution (see Eq. 1) is used to represent this deviation. This error is presented in real systems as a drift in the heading angle caused.

Initial uncertainty—The exploration and mapping mission starts when the UAV takes off from a known initial position and heading with respect to a local frame $\boldsymbol{P_w} = \{x_w, y_w, z_w, \psi_w\}$. After taking off, the UAV experiences a drift in its initial position and heading denoted as $\boldsymbol{P_{w_d}} = \{x_{w_d}, y_{w_d}, z_{w_d}, \psi_{w_d}\}$. This drift in the initial position and heading has some uncertainty which is modelled as a deviation from the initial UAV position and heading. This uncertainty is modelled using a Gaussian probability distribution that follows:

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

where $P(x)$ is the probability of the variable x, μ is the expectation value of the distribution and σ is the standard deviation.

The POMDP algorithm uses particles to represent belief-states. In this representation, an initial belief-state b_0 is created using particles generated through sampling using probability distributions. The particles must contain the initial UAV position and heading after take off $\boldsymbol{P_{w_0}}$. The UAV position and heading for each of these particles is calculated using on Eq. 2.

$$\boldsymbol{P_{w_0}} = \boldsymbol{P_w} + \boldsymbol{P_{w_d}} \tag{2}$$

UAV motion uncertainty—Another source of uncertainty for the target tracking mission is in the dynamics and the motion of the UAV. That is, there is uncertainty in the fact that a commanded action may not result in the desired response. Four PID controllers were designed to control four states of the UAV. Characteristic step responses in time were obtained in order to include them in the transition model as the dynamic model of the UAV.

The dynamic response to the actions generated by the POMDP motion planner are included as discretised look-up tables. The standard deviation of the UAV step responses for each of the controlled states were measured in flight trials and characterised using a Gaussian distribution.

Occupancy map—The map of the environment is modelled using the octomap 3D occupancy library which stores the information in a data structure as occupancy voxels with a predetermined size. The size of voxels is the resolution of the map and is a parameter that can be set according to the needs of precision depending on the application. This occupancy map stores the probabilities of a voxel being occupied. A depth sensor producing point clouds is used to create and update the map once new areas are explored. Selecting a coarse voxel resolution increases the obstacle size in the map and thus generates a conservative approach for obstacle avoidance to the detriment of precision. An image of a created 3D octomap is shown in Fig. 2.

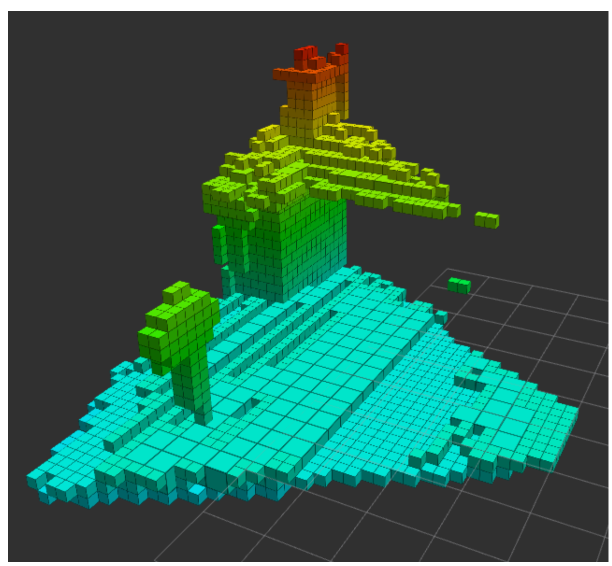


Figure 2. Occupancy map of scene

Reward function (R)

The reward function models the objectives of the mission. In the exploration and mapping mission the UAV must avoid collisions with obstacles and must navigate within an operational air-space. Furthermore, the UAV needs to explore the unknown space in order to update an occupancy map that can later be used for navigation. Additionally, the UAV must perform this mission as fast as possible to take advantage of its flight time. The reward function can be expressed as:

$$R_T = \alpha - \beta - \gamma - \epsilon \tag{3}$$

where α represents the reward of exploring an unknown space, β is the cost of colliding with obstacles, γ is the cost of going out of the operational air-space and ϵ is the energy cost for every action executed.

4. SYSTEM ARCHITECTURE

The framework is implemented using a modular system shown in Fig. 3, which is composed of the following elements:

► a POMDP solver algorithm with the implementation of the problem formulation that outputs actions in the form of position commands to the UAV. This module is programmed in C++ and has a ROS node to interface and communicate with the other modules through ROS.
► A module to simulate the UAV flying in a simulated world using the Gazebo software. It has a sub-module called SITL that executes the flight controller firmware. In addition, a sub-module computes the data acquired by the depth sensor in the form of point clouds. This module publishes the data from the UAV pose and depth sensor as ROS messages. This module uses open source software and is implemented following the documentation from the PX4 development guide [12].
► Another module, called the octomap server receives the messages from the depth sensor as point clouds and the UAV pose in order to create, update and publish an occupancy map as a ROS message.
► An observation module reads the updated map along with the estimated pose and produces an observation which is sent to the POMDP planner in order to update the belief and recompute the motion policy based on the new map.

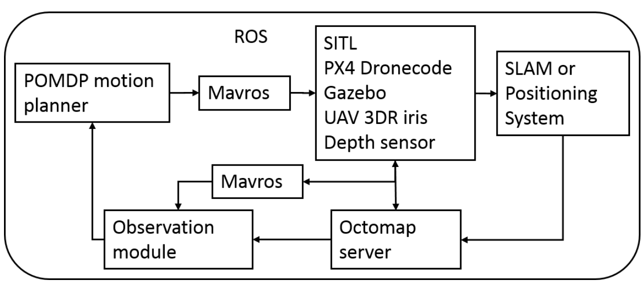


Figure 3. System and software architecture for simulation

The pose of the UAV is given as a frame measure with respect to an initial local frame. In a GPS-enabled environment this pose comes from the internal estimation of the flight controller, based mostly on GPS positioning and compass. In a GPS-denied scenario, the pose will come from a SLAM or visual odometry (VIO) [13], [14], [4] system which uses on-board sensors such as depth cameras and or INS to estimate the pose of the UAV in the local frame.

5. RESULTS

The framework was tested in simulation using software in the loop (SITL). This program simulates the flight controller firmware (PX4 Dronecode) and exchanges mavlink messages with a ground control station for motion commands and on-board sensor readings. We use a model of an off-the-shelf UAV (3DR Iris) flying in different simulated worlds. The UAV has the mission to explore the environment and generate an occupancy map whilst avoiding obstacles within an operational air-space.

In the first scenario the UAV takes off from the ground in front of a collapsed building as shown in Fig. 4(a). When the UAV is on the ground there is a prior map that is generated based on the point cloud produced by the depth sensor with the view in front of the UAV. Once the UAV takes off, it then flies towards the collapsed building entrance, it ascends, descends and turns exploring the environment, creating and updating the map, whilst avoiding obstacles. In this case the UAV does not enter the collapsed building due to the high risk of collision. Fig. 4(b) shows an initial map of the collapsed building scenario when the UAV starts its mission. Fig. 4(c) shows the final map created by the framework after executing the motion policy. A video of another episode simulated in this same scenario can be seen at https://youtu.be/DjuF2ZKtJ5Q.

A second simulation scenario is presented in Fig. 5. In this scenario, the UAV needs to navigate safely and produce a map of the space under the gas station ceiling. This is a situation where the GPS signal will be occluded and the estimation of the UAV pose will come from a SLAM or VIO system.

Figure 6 shows the generated occupancy map after the UAV has explored the GPS-denied space under the gas station ceiling. The operational air-space of the UAV is reduced to a limited height so that it is forced to fly under the ceiling for the purpose of testing. Once the UAV enters the GPS-denied space, it performs a sequence of turns in order to maximise the information gain whilst keeping a safe distance from the walls.

Furthermore, once the UAV has acquired more knowledge of the surrounding below the ceiling structure, it keeps exploring further, doing more turns and then returning.

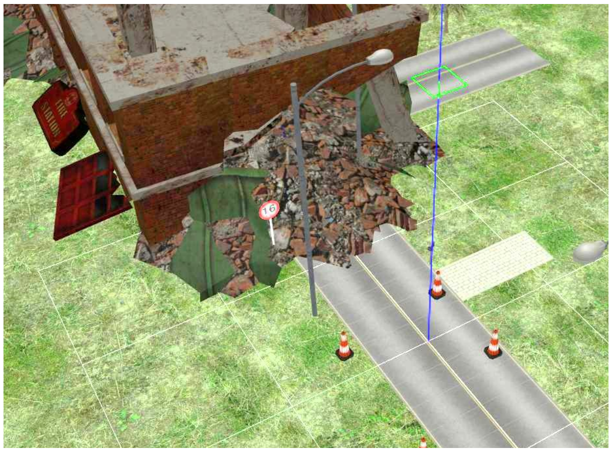
6. CONCLUSIONS

This paper presented a framework to formulate a UAV navigation and exploration mission in unknown GPS-denied environments as a sequential decision process under uncertainty using POMDP. The framework consist of a POMDP formulation, and various modules that communicate with each other through the robotic operating system.

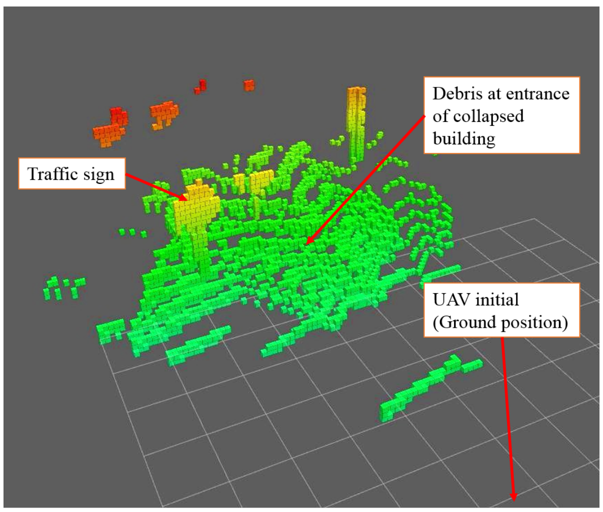
The framework is evaluated in simulation and results indicate that the UAV can safely navigate in an unknown environment where there is partial or total occlusion of GPS signal. The system commands the UAV through a sequence of actions in order to maximise the information gain about the unknown environment. The system creates and updates an occupancy map that enables the UAV to fly safely.

Ongoing work focuses on implementation of the framework in a real scenario using an Intel aero drone equipped with a depth sensor. This implementation will use the same flight controller firmware and ROS environment but in addition, visual inertial odometry systems and SLAM algorithm will be tested.

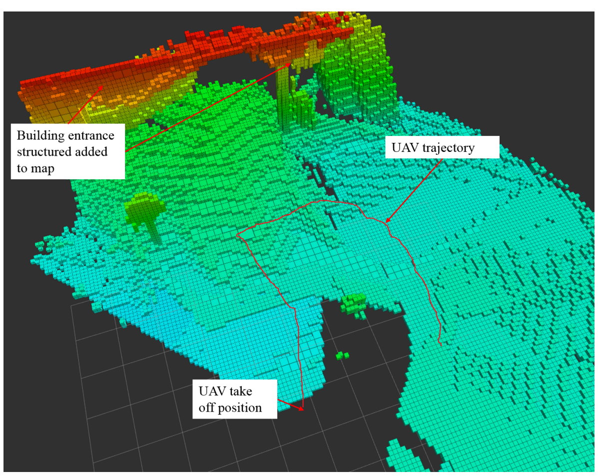
Furthermore, ongoing work focuses on exploring different



(a) Simulated environment with collapsed building



(b) Prior or initial occupancy map of collapsed building in front of UAV.



(c) Updated occupancy map of collapsed building after UAV exploration.

Figure 4. Render of occupancy maps generated before a) and after b) exploration.

reward function structures that maximise exploring unknown spaces.

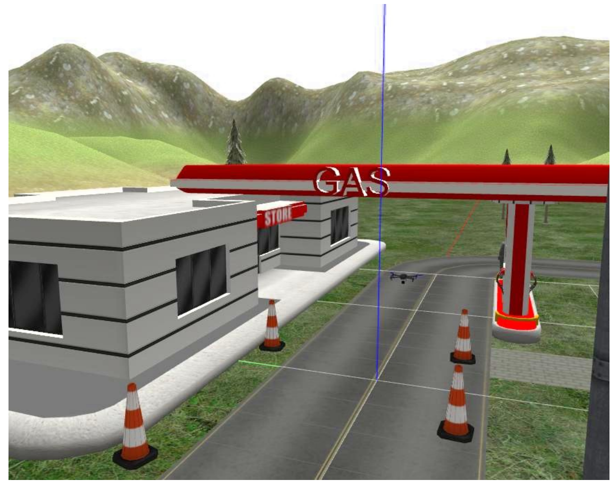


Figure 5. **Second simulated scenario of a GPS-denied place. In this, the UAV needs to navigate and map the space under the gas station ceiling.**

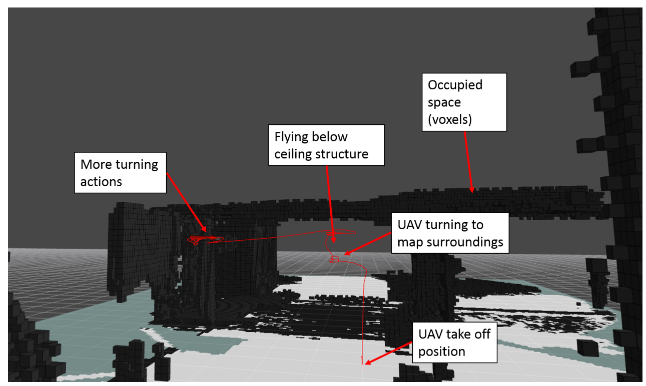


Figure 6. **Render of the occupancy map generated after UAV exploration.The UAV trajectory is shown in red.**

APPENDICES

A. REWARD FUNCTION PARAMETERS

The values used for each of the terms in the reward function are shown in Eqs. 4 to 7.

$$\alpha = \begin{cases} 15 & \text{new space discovered} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\beta = \begin{cases} 70 & \text{UAV collides with obstacle} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$\gamma = \begin{cases} 70 & \text{UAV flies out of operating air-space} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$\epsilon = \begin{cases} 2 & \text{Energy cost per time step} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

ACKNOWLEDGMENTS

The authors thank the Australian Research Council and Queensland University of Technology for funding this project through an ARC Discovery Project grant.

REFERENCES

[1] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[2] M. Labbé and F. Michaud, "Online global loop closure detection for large-scale multi-session graph-based SLAM," *IEEE International Conference on Intelligent Robots and Systems*, pp. 2661–2666, 2014.

[3] A. S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, and N. Roy, "Visual odometry and mapping for autonomous flight using an RGB-D camera," in *Springer Tracts in Advanced Robotics*, vol. 100, 2017, pp. 235–252.

[4] G. Zhou, L. Fang, K. Tang, H. Zhang, K. Wang, and K. Yang, "Guidance: A visual sensing platform for robotic applications," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2015-October, pp. 9–14, 2015.

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

[6] A. Guez and J. Pineau, "Multi-tasking SLAM," in *Proceedings - IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 377–384.

[7] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[8] 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.

[9] F. Vanegas, D. Campbell, N. Roy, K. J. Gaston, and F. Gonzalez, "UAV Tracking and Following a Ground Target under Motion and Localisation Uncertainty," in *2017 IEEE Aerospace Conference Proceedings*, 2017.

[10] 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.

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

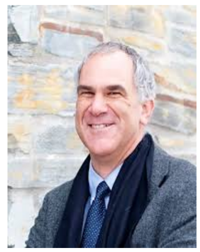
[12] "PX4 development guide," https://dev.px4.io/en/, accessed: 2018-10-18.

[13] T. Schneider, M. T. Dymczyk, M. Fehr, K. Egger, S. Lynen, I. Gilitschenski, and R. Siegwart, "maplab: An open framework for research in visual-inertial mapping and localization," *IEEE Robotics and Automation Letters*, 2018.

[14] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated extended kalman filter based visual-inertial odometry using direct photometric feedback," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, 2017.

BIOGRAPHY

Fernando Vanegas *is a Postdoctoral Research Fellow in Robotics at Queensland University of Technology at QUT. He holds a M.Sc. in Electrical Engineering from Halmstad University and a PhD in Aerial Robotics QUT. His current research interests include motion planning for UAV exploration in GPS-deniend and cluttered environments, POMDP, SLAM and VO.*



Kevin Gaston *is Professor of Biodiversity and Conservation, and founding director of the Environment and Sustainability Institute at the University of Exeter. He leads basic, strategic and applied research in ecology and conservation biology, with an emphasis on using novel technological approaches.*

Jonathan Roberts *is Professor in Robotics at Queensland University of Technology (QUT). His main research interest are in the area of Field Robotics and Medical Robotics. Jonathan was a co-inventor of the UAV Challenge, an international flying robot competition. Jonathan is a Past President of the Australian Robotics & Automation Association. He currently serves as a Senior Editor of the IEEE Journal of Robotics and Automation Letters, and an Associate Editor of the Journal of Field Robotics.*

Felipe Gonzalez *is an Associate Professor (UAV) in the Science and Engineering Faculty, QUT and the QUT UAV Remote Sensing Group. He holds a BEng (Mech) and a PhD from the University of Sydney. His research explores bio-inspired optimization, uncertainty based UAV path planning and UAV for environmental monitoring. He currently leads the ARC DP project UAV flying under the canopy and urban jungle. Felipe is a Chartered Professional Engineer and member of professional organizations including the RAeS,IEEE and AIAA.*