© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Distance Function based 6DOF Localization for Unmanned Aerial Vehicles in GPS Denied Environments

James Unicomb[1], Lakshitha Dantanarayana*[1], Janindu Arukgoda[1],
Ravindra Ranasinghe[1], Gamini Dissanayake[1], and Tomonari Furukawa[2]

Abstract— **This paper presents an algorithm for localizing an unmanned aerial vehicle (UAV) in GPS denied environments. Localization is performed with respect to a pre-built map of the environment represented using the distance function of a binary mosaic, avoiding the need for extraction and explicit matching of visual features. Edges extracted from images acquired by an on-board camera are projected to the map to compute an error metric that indicates the misalignment between the predicted and true pose of the UAV. A constrained extended Kalman filter (EKF) framework is used to generate an estimate of the full 6-DOF location of the UAV by enforcing the condition that the distance function values are zero when there is no misalignment. Use of an EKF also makes it possible to seamlessly incorporate information from any other system on the UAV, for example, from its auto-pilot, a height sensor or an optical flow sensor. Experiments using a hexarotor UAV both in a simulation environment and in the field are presented to demonstrate the effectiveness of the proposed algorithm.**

I. INTRODUCTION

Small autonomous Unmanned Aerial Vehicles (UAVs) such as quadrotors or hexarotors are becoming increasingly common in a broad range of applications. Many of these applications require a UAV to follow a path or move to a location defined in the 3D Cartesian space. Typically, the UAV location obtained using a GPS receiver is therefore adequate in many situations. However, incorporating robustness to GPS failure is widely accepted as important, particularly in law enforcement and defence applications. UAVs have been built with navigation units that are able to fuse information from one or more of the multitude of sensors including an on-board inertial measurement unit, a magnetometer, a height sensors, a barometer and optical flow sensors, typically using an extended Kalman filter (EKF). Such a system can generate sufficiently accurate estimates of the attitude of the UAV making it feasible to control and stabilise the UAV. The Cartesian position of the UAV become observable only when a GPS receiver is connected to the navigation unit. The challenge addressed in this paper is how to provide position when information from the GPS receiver is momentarily lost.

Integration of information from an on-board camera and a body mounted inertial measurement system is one of the most promising approaches to navigate in environments without GPS information. In this paper, we focus on using images from a camera looking at the terrain below. We consider the situation that the UAV is flying at a sufficient height so that we can assume that the terrain is locally planar. We also assume that an a-priori geo-referenced map of the environment is available. Although the UAV attitude with an accuracy adequate for control purposes is available from its navigation system, a small angular error results in a large change to the image captured from a camera. Therefore, we focus on solving the complete six degrees-of-freedom (DOF) localisation problem to estimate the UAV state. This work extends the distance function based 3-DOF localisation on 2D occupancy grid maps presented in [1].

The localisation strategy used in this paper is as follows. We use a distance function based technique for representing the map of the environment [2], [1]. This is achieved by first obtaining a binary edge map of the terrain and computing its signed distance function. We use a cubic-spline approximation to extract high resolution information from the map and its spatial derivatives in a compact form. Although computing distance function and its derivatives is computationally expensive, this needs to be done only once. We extract edges from the image captured by the on board camera and compute value of the distance function and its uncertainty at each edge pixel, using the map and the latest available 6-DOF state of the UAV. This is done at run-time but consists of simply a few function look ups. We then use the fact that the distance function value at each edge pixel should be zero at the correct alignment in an EKF framework to compute the 6-DOF state estimate of the UAV. We implement the pseudo-observation technique to enforce this constraint and use the sum of squares of distance function values at each edge pixel to reduce the dimension of the resulting observation equation. We demonstrate that the EKF can function using a motion model based on the velocities and angular velocities of the UAV or with the UAV motion modelled as a random walk.

We argue that the method proposed in this paper has a number of significant advantages over image based techniques proposed in literature. Use of edges extracted from images make the strategy robust to illumination changes as well as makes it computationally light, as tracking, association, and explicit matching of visual features in the environment is not required. The observation equation has the same computational cost, independent of the height from which the image is taken, in contrast to template matching techniques that require resizing images and therefore can be costly to implement. Furthermore, only the edge pixels need to be processed, rather than the whole image. In addition,

---

[1] Centre for Autonomous Systems, University of Technology Sydney, Australia.

[2] Computational Multi-physics Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

* Corresponding Author: `laki (at) ieee.org`

the availability of sound information on the uncertainty associated with the measurements makes it possible to exploit conventional EKF strategies to reject measurements resulting from noise due to the edge extraction process or from edges that are visible in the camera image but not present in the pre-built map. Thus when building the map, only edges that are always likely to be visible can be used.

The paper is organised as follows. Section II reviews the related literature. Section III presents a brief description of the UAV platform and the specific task for which the proposed algorithm was first developed. Section IV details the methodology including the formulation of the EKF. Experimental results are presented in Section V, while Section VI concludes the paper.

II. RELATED WORK

As mentioned in Section I, vision based sensing techniques for mapping and navigation of UAVs using information from an on-board camera and inertial information from a body mounted inertial measurement system are quite common.

Yol et. al [3] proposed a template based image registration method in which the UAV camera motion obtained be maximising a mutual information based similarity function between the images acquired by the on-board camera and geo-referenced images. Amidi et al. [4] presented a vision-based odometer based on a stereo pair pointed at the ground together with a gyroscopes to determine the UAV pose. In order to address significantly high drift in traditional GPS aided inertial navigation system in estimating the UAV pose during extended GPS outages, Madison et al. [5] proposed a method to augment the navigation system with an onboard camera to track visual landmarks to infer vehicle motion. Using an EKF framework the vehicle states and the inertial locations of 3D features are estimated. Langelaan [6] has presented a framework to compute the UAV state (position, orientation, and velocity) and the positions of obstacles in the environment using a sigma point Kalman filter to enable control and navigation of small autonomous UAVs operating in cluttered environments, using a monocular camera and inertial measurements as sensing. Here the state estimation problem was formulated as bearings only Simultaneous Localisation And Mapping (SLAM).

An inertial-aided vision-based localization and mapping algorithm is proposed by Yang et al. [7]. However, this is specifically developed for a UAV operating in riverine environments. The algorithm exploits a two-view geometry formulation using captured features surrounding rivers and their corresponding points reflected in the river using a light weight monocular camera to triangulate locations, and an Inertial Measurement Unit (IMU) fitted to the UAV is used to compliment the estimates.

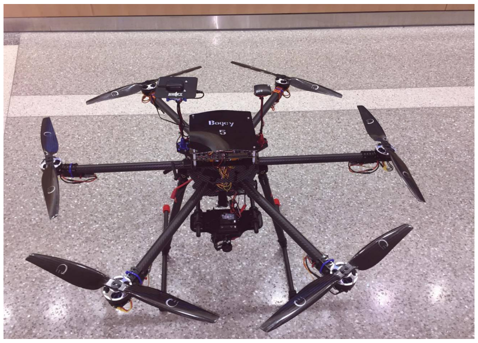


Fig. 1: Hexarotor UAV

Work reported in Wu et al. [8] uses Harris corner detector based features extracted from a monocular camera that are then fused with information from an IMU in an EKF framework to address navigation and mapping problems. At the beginning, when GPS is available, UAV uses feature point observations of landmarks in the surrounding environment to map out the landmarks in 3D inertial space. Once accurate feature locations are available, observations from the camera can be subsequently used to determine the pose of the aircraft in the absence of GPS. This, however, requires accurate data association.

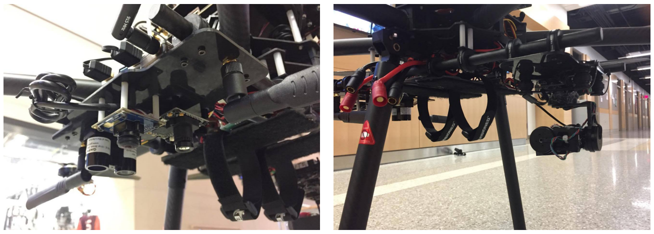
III. DESCRIPTION OF UAV PLATFORM

This work was motivated by the need to develop a reliable localisation framework for hexarotor UAVs that are used by team VICTOR in the Mohamed Bin Zayed International Robotics Challenge (MBZIRC) competition in March, 2017. The competition is composed of three stages: stage one and three requires a UAV to autonomously land on a truck moving on the figure-of-eight track drawn in the arena, and for a team of UAVs to autonomously and cooperatively pick up objects that are scattered around the arena.

The UAV platform of choice for team VICTOR is a custom-built hexa-rotor helicopter (Fig. 1). This UAV is designed for a much higher payload capacity and longer endurance than the commercially available hexa-rotor helicopters to be able to complete the challenges posed at the competition. The platform is equipped with a Pixhawk$^{\mathrm{TM}}$ flight controller unit and an on-board ARM computer. This flight controller comes with an in-built IMU and several other sensors such as an on-board GPS with a compass. An RTK GPS module, a laser range finder, a PX4FLOW$^{\mathrm{TM}}$ optical sensor and two cameras connected externally to the flight controller. The first of these cameras is a 5 MP perspective camera mounted on the UAV body using a gimbal to maintain its optical axis normal to the ground plane while the second camera is a fish-eye camera with 180° field of view fixed to the UAV body pointing downwards.

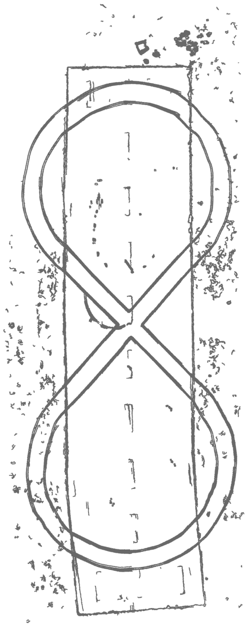
The images captured by either the perspective camera or the fish-eye camera are to be used as inputs for localisation.

We run all high-level software modules using the on-board NVIDIA$^{\mathrm{TM}}$ Jetson TX1 computer. We rely on Robot Operating System (ROS), which is an open source set of software libraries and tools to build the software framework



(a) (b)

Fig. 2: Visual sensors mounted on the platform. (a). LiDar, Px4Flow and fisheye camera mounted on UAV . (b). Gimbal used to mount the perspective camera



(a) (b)

Fig. 3: (a) Map of the arena created by stitching UAV images. (b) Cleaned-up map used for localization.

to interface low level drivers and high-level algorithms such as localisation.

IV. METHODOLOGY

A. Map of the Environment

In order to localise, the UAV uses a pre-built map represented using the distance function of a binary mosaic that correspond to the edges that are present in the environment. Fig. 3a shows the initial map built using the images acquired using the on board camera. This map has been pruned to remove spurious noise from the map while retaining the prominent edge image (Fig. 3b) to use as the final map.

B. Sensing the Environment

Once an image is captured through the on board camera (perspective camera or the fish-eye camera), it is processed to correct each pixel location for radial and tangential lens distortions. Subsequently, the undistorted image is passed through an edge detector to identify edge pixel points $(\lambda_i,\ \mu_i)$ that are used in the observation model. Fig. 4 shows some typical camera images captured during the real experiment. It can be seen that lighting conditions of the environment varies across images, but, the edge detection process can be made robust to these changes.

Detected pixels belonging to an edge form the observation vector, $\mathbf{z}$. Each of the edge pixels on the image plane are transformed into their original location on the ground plane with respect to the camera coordinate frame using the camera projection model. The blue dots on Fig. 5 represent a set of edge pixels of the image (Fig. 5) captured from the perspective camera. This process only takes approximately 6ms to complete.

C. Distance Function Based Observation Model

Distance functions are implicit shape representations commonly used within the computer vision community. For a

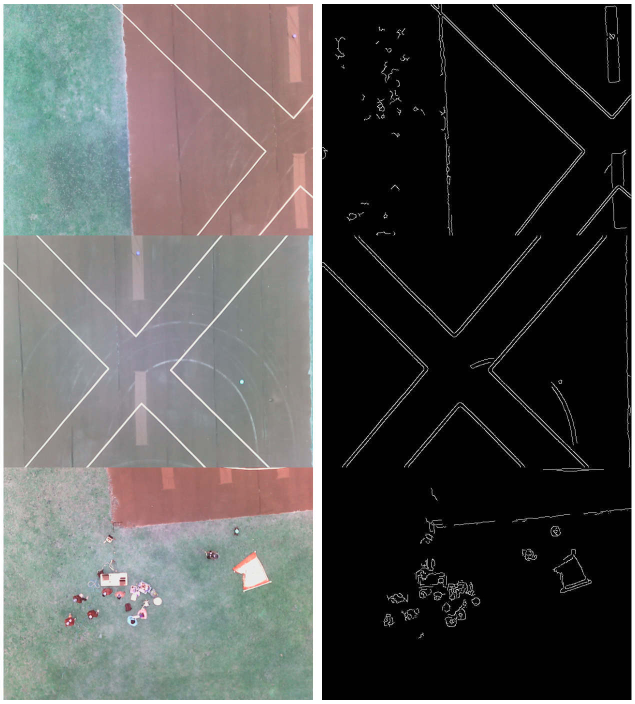


Fig. 4: Typical images seen by the UAV camera at an approximate altitude of 20 metres, and corresponding edge images

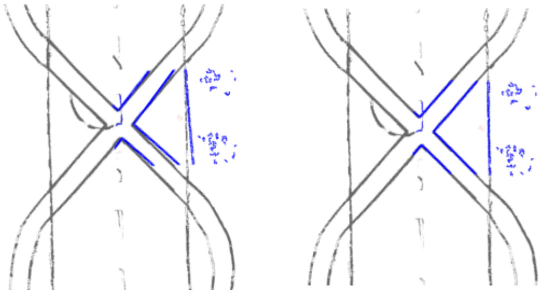


Fig. 5: Selected edge pixels projected on the ground plane. Before and after alignment.

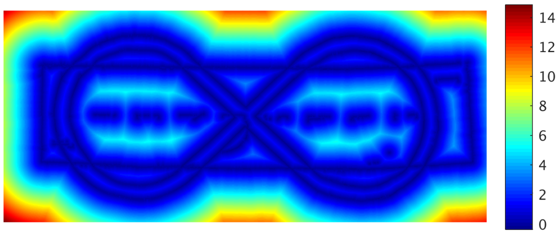


Fig. 6: Distance field of the arena

given binary image with the set of boundaries V, the distance function value at a given location $\mathbf{x}$, is computed via (1), which specifies the Euclidean distance from that pixel to the nearest boundary pixel v_j in V [9]. Fig. 6 shows the distance field of the synthetically generated MBZIRC arena in Gazebo simulation environment.

$$DF(\mathbf{x}) = \min_{\mathbf{v_j} \in V} \|\mathbf{x} - \mathbf{v_j}\| \tag{1}$$

The distance function as described in (1) quantises these distances into cell numbers. Furthermore, derivatives of the distance functions are not continuous at points which belong to the edges of the binary map or to the cut-locus [10]. As the purpose of the exercise is to use the distance functions as the basis for an observation model within an EKF framework, a cubic spline approximation based interpolation algorithm is used to compute the distance function and its derivatives at any given location in the map. The parameters of the spline function are precomputed and stored so that the computational effort required during runtime is minimised. Future references to distance function in this paper refer to the interpolated continuous version.

We further define the sign of the distance function based on the colour of the initial binary map. Distances corresponding to points in black regions hold a positive value while points in white region holds a negative value. This guarantees that the gradient of distance function is continuous at the edges of the binary map.

The observation vector $\mathbf{z}_i \in \mathbf{z}$; $\mathbf{z}_i = (\lambda_i, \mu_i)$, of a single binary image obtained by the camera, consisting of n edge points with coordinates $(\lambda,\ \mu)$ on the image plane can be projected from the current estimate of the 6DOF UAV pose $\mathbf{x} = (x,\ y,\ z(\text{altitude}),\ \psi(\text{roll}),\ \theta(\text{pitch}),\ \phi(\text{yaw}))^{\top}$, using (2) to obtain the observation vector in 2D space $\mathbf{x}_o$ on the ground plane.

$$\mathbf{x}_{o_i} = \begin{Bmatrix} x_{o_i} \\ y_{o_i} \end{Bmatrix} = \begin{Bmatrix} x + z\frac{\lambda_i R_{1,1} + \mu_i R_{1,2} - f R_{1,3}}{\lambda_i R_{3,1} - \mu_i R_{3,2} + f R_{3,3}} \\ y + z\frac{\lambda_i R_{2,1} + \mu_i R_{2,2} - f R_{2,3}}{\lambda_i R_{3,1} - \mu_i R_{3,2} + f R_{3,3}} \end{Bmatrix} \tag{2}$$

where, R is the 3D rotation matrix representing the UAV pose, $R(\psi,\ \theta,\ \phi)$, and f is the focal length of the camera.

When the projected edge points and the map are fully aligned, the sum of squared distance function values at these points is expected to be zero. Therefore, setting the expected measurement to zero in (3) yields the measurement equation (4) that is suitable for robot localisation.

$$h(\mathbf{x}, \mathbf{z}) = \sum_{i=0}^{n-1} DF(\mathbf{x}_{o_i})^2 = d_{ss} \tag{3}$$

$$h(\mathbf{x}, \mathbf{z}) = 0 \tag{4}$$

D. Formulation of the Extended Kalman Filter

Traditional formulation of the EKF requires an observation equation of the form $\mathbf{z} = h(\mathbf{x})$. The alternative formulation that is proposed below can directly deal with an implicit form of the measurement equation.

1) Prediction: Let the estimate of the UAV pose $\mathbf{x}_{k|k-1} = (x_{k|k-1},\ y_{k|k-1},\ z_{k|k-1},\ \psi_{k|k-1},\ \theta_{k|k-1},\ \phi_{k|k-1})^{\top}$, be subjected to a control command of $\mathbf{u}_k = (v_k, \omega_k)^{\top}$, where $\mathbf{v}_k$ is the linear velocity (in x, y, and z directions) and ω_k is the angular velocity (in ψ, θ, and ϕ directions) over a period of Δt.

Then the predicted location of the UAV is given by (5) and its covariance by (6).

$$\mathbf{x}_{k|k-1} = F(\mathbf{x}_{k-1|k-1},\ \mathbf{u}_k \Delta t) \tag{5}$$

$$P_{k|k-1} = \nabla F_{\mathbf{x}} P_{k-1|k-1} \nabla F_{\mathbf{x}}^{\top} + \nabla F_u Q_k \nabla F_u^{\top} \tag{6}$$

where $\nabla F_{\mathbf{x}}$ and ∇F_u are respectively the Jacobian of the control function F with respect to $\mathbf{x}$ and $\mathbf{u}$, obtained by linearising about the UAV pose estimate $\mathbf{x}_{k-1|k-1}$, while Q_k is the control noise covariance matrix.

2) Observation: Equation (4) provides the observation function, $h(\mathbf{x}, \mathbf{z}) = 0$.

Assuming that each selected pixel point of the edge image $\mathbf{z}$, corrupted by noise η with $\mathcal{N}(0, \sigma^2)$ in both λ and μ directions, the covariance of the measurement vector is given by the diagonal matrix, $\Sigma_{\mathbf{z}} = diag(\sigma^2,\ \sigma^2)$.

3) Update: Update equations can be written as follows, where the filter gain K is given by,

$$K = P_{k|k-1} \nabla h_{\mathbf{x}}^{\top} (\nabla h_{\mathbf{x}} P_{k|k-1} \nabla h_{\mathbf{x}}^{\top} + \nabla h_{\mathbf{z}} \Sigma_{\mathbf{z}} \nabla h_{\mathbf{z}}^{\top})^{-1} \tag{7}$$

The state update:

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + K(-h(\mathbf{x}_{k|k-1}, \mathbf{z})) \tag{8}$$

while the covariance update is,

$$P_{k|k} = (I - K \nabla h_{\mathbf{x}}) P_{k|k-1} \tag{9}$$

The Jacobians $\nabla h_{\mathbf{x}}$ and $\nabla h_{\mathbf{z}}$ at the appropriate linearisation points can be easily calculated using the DF.

As previously mentioned, DF and its derivatives can be precomputed and stored to be used during the runtime to efficiently compute the gradients of DF and the Jacobians above. The remaining components of the gradient can then be analytically derived from (2).

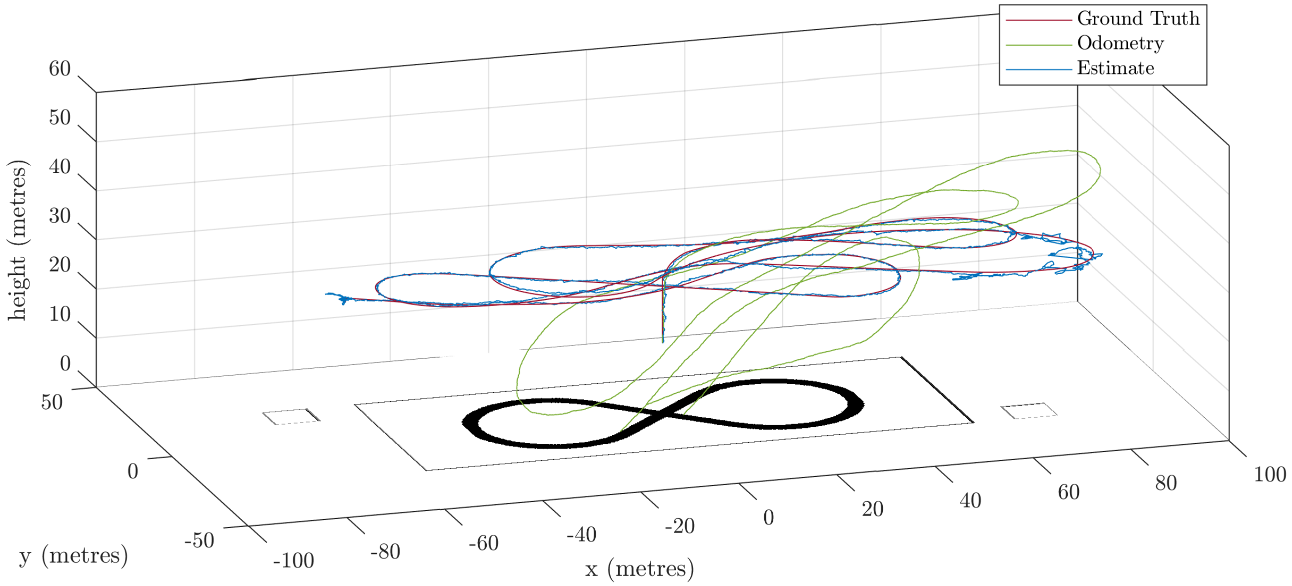


Fig. 7: Flight path of the UAV during Gazebo simulation. Estimation without the use of odometry.

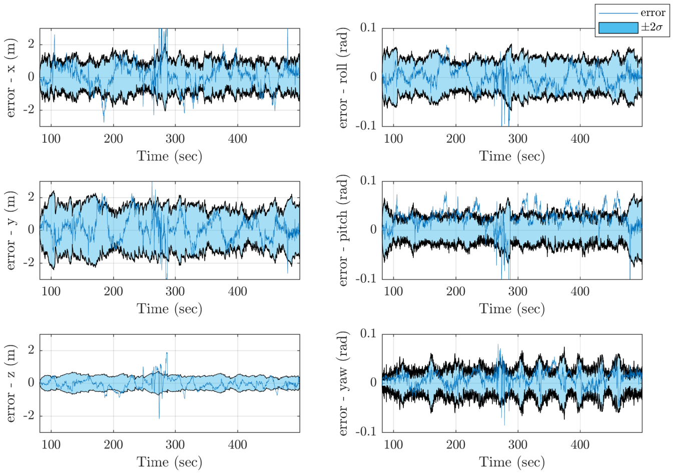


Fig. 8: Pose error and the relevant covariances during simulation without using odometry.

Fig. 4 and Fig. 5 shows spurious measurements resulting from the edge extraction process. This is to be expected in a practical situation as the objects that did not exist during the mapping process can exist in the environment. Changes to the illumination and shadows can also contribute to this. An innovation gate (with a 2σ bound) is used to rejects these measurements before they reach the update stage.

V. EXPERIMENTAL RESULTS

In this section we validate the proposed algorithm using simulation and experimental data.

For simulations, we use the Gazebo robot simulation software provided with ROS Kinetic Kame. The simulated UAV flies at varying altitudes and captures images using a monocular camera that is mounted on the robot's body. The Pixhawk controller firmware estimates the UAV velocities and odometry based on physics properties of the simulation. Ground-truth is known.

For the real experiment we used images obtained from the normal camera. The UAV is controlled by a Pixhawk Autopilot device. RTK GPS unit is also present on the UAV and is used only for ground-truth. Map shown in Fig. 3b is used for localization.

A. Simulation Results

During the simulation, the UAV is flown in altitudes varying between 20 - 40 metres and travels a total distance of 1km. Fig. 7 shows a comparison between ground-truth and the estimated poses. Odometry generated from the UAV flight system is also plotted. This odometry is used within the EKF framework during the prediction phase. However, as the flight system cannot estimate the x and y components of the UAV linear velocity in the absence of GPS, we consider those components to be zero, and use a larger uncertainty to capture the transition.

It can be seen that the estimation closely follows the ground-truth trajectory. However, when the UAV moves out of the range of the arena (275 seconds into simulation, Fig. 8), the errors increase when the image captured by the camera is completely blank, as soon as the UAV is back in range, the estimate continues to track the trajectory well.

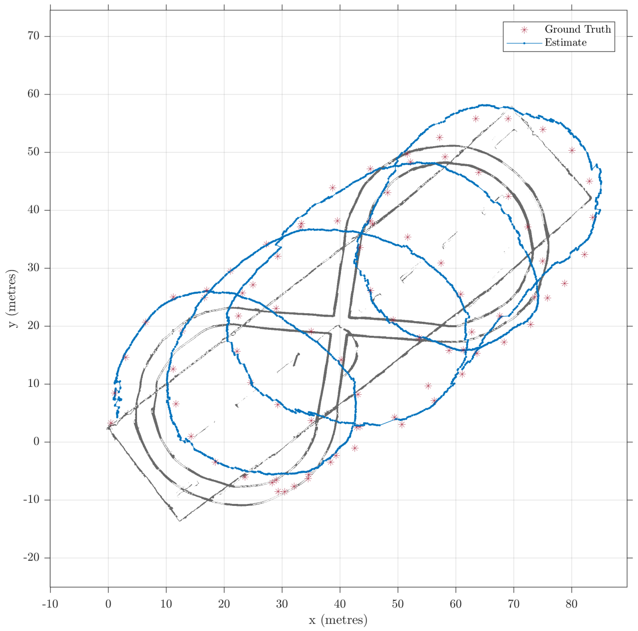


Fig. 9: Flight path of the UAV during the field trials.

B. *Field Trials*

Results from estimation during field trials are shown in Fig. 9. Position updates from RTK-GPS system is only available at a 1Hz frequency and are plotted against the estimated pose. However, it is not reliable and depends on environment conditions. We use these GPS readings for ground-truth purposes.

It can be seen that the estimates closely follow the ground-truth trajectory.

VI. Conclusion

In this paper, we presented an algorithm for localizing a UAV in absence of GPS signals. We use experiments conducted both in simulation and field trials to demonstrate the effectiveness of the proposed algorithm. The results reported demonstrate that the estimated pose closely follow ground-truth trajectories even when an odometry estimate of the UAV is not present. It is also seen that the algorithm is able to reject observations from edges extracted from the images that contain noise and artefacts that are not present in the map.

Our future work will focus on incorporating information from other sensors mounted on the UAV for example the optical flow sensor. We also aim to explore the utility of tightly coupling the information from an IMU and other sensors in a single EKF framework.

One major limitation of the work presented is the assumption that the map of the environment that the UAV traverses is planar. Relaxing this assumption using a full three dimensional distance map or exploring whether maintaining only a planar map and relying on the outlier rejection capability of the EKF is adequate, are also of interest.

References

[1] L. Dantanarayana, G. Dissanayake, R. Ranasinghe, and T. Furukawa, "An extended Kalman filter for localisation in occupancy grid maps," in *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 12 2015, pp. 419–424. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7399048

[2] L. Dantanarayana, G. Dissanayake, and R. Ranasinge, "C-LOG: A Chamfer Distance Based Algorithm for Localisation in Occupancy Grid-maps," *CAAI Transactions on Intelligence Technology*, 10 2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S2468232216300555

[3] A. Yol, B. Delabarre, A. Dame, J.-E. Dartois, and E. Marchand, "Vision-based absolute localization for unmanned aerial vehicles," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 9 2014, pp. 3429–3434. [Online]. Available: http://ieeexplore.ieee.org/document/6943040/

[4] O. Amidi, T. Kanade, and K. Fujita, "A visual odometer for autonomous helicopter flight," *Robotics and Autonomous Systems*, vol. 28, no. 2-3, pp. 185–193, 8 1999. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0921889099000160

[5] R. Madison, G. Andrews, P. DeBitetto, S. Rasmussen, and M. Bottkol, "Vision-Aided Navigation for Small UAVs in GPS-Challenged Environments," in *AIAA Infotech@Aerospace 2007 Conference and Exhibit*. Reston, Virigina: American Institute of Aeronautics and Astronautics, 5 2007. [Online]. Available: http://arc.aiaa.org/doi/10.2514/6.2007-2986

[6] J. W. Langelaan, "State Estimation for Autonomous Flight in Cluttered Environments," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 5, pp. 1414–1426, 9 2007. [Online]. Available: http://arc.aiaa.org/doi/10.2514/1.27770

[7] J. Yang, A. Dani, S.-J. Chung, and S. Hutchinson, "Inertial-Aided Vision-Based Localization and Mapping in a Riverine Environment with Reflection Measurements," in *AIAA Guidance, Navigation, and Control (GNC) Conference*. Reston, Virginia: American Institute of Aeronautics and Astronautics, 8 2013. [Online]. Available: http://arc.aiaa.org/doi/10.2514/6.2013-5246

[8] Allen D. Wu, Eric N. Johnson, Michael Kaess, Frank Dellaert, and Girish Chowdhary, "Autonomous Flight in GPS-Denied Environments Using Monocular Vision and Inertial Sensors," *Journal of Aerospace Information Systems*, vol. 10, no. 4, pp. 172–186, 4 2013. [Online]. Available: http://arc.aiaa.org/doi/10.2514/1.I010023

[9] M. Y. Liu, O. Tuzel, A. Veeraraghavan, and R. Chellappa, "Fast directional chamfer matching," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 6 2010, pp. 1696–1703. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5539837

[10] M. Jones, J. Baerentzen, and M. Sramek, "3D distance fields: a survey of techniques and applications," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 4, pp. 581–599, 7 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16805266 http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1634323