Robust Multi-Sensor Fusion for Micro Aerial Vehicle Navigation in GPS-Degraded/Denied Environments

Andrew Chambers, Sebastian Scherer, Luke Yoder, Sezal Jain, Stephen Nuske, Sanjiv Singh

***Abstract*— State estimation for Micro Air Vehicles (MAVs) is challenging because sensing instrumentation carried on-board is severely limited by weight and power constraints. In addition, their use close to and inside structures and vegetation means that GPS signals can be degraded or all together absent. Here we present a navigation system suited for use on MAVs that seamlessly fuses any combination of GPS, visual odometry, inertial measurements, and/or barometric pressure. We focus on robustness against real-world conditions and evaluate performance in challenging field experiments. Results demonstrate that the proposed approach is effective at providing a consistent state estimate even during multiple sensor failures and can be used for mapping, planning, and control.**

I. Introduction

Micro aerial vehicles can fly close to and under buildings and vegetation which makes them especially useful in environments where GPS is unreliable or unavailable. In these environments it is necessary to estimate velocities and position using exteroceptive sensors because available proprioceptive sensors (inertial measurement units) are not accurate enough to enable velocity and position control.

The demand for micro aerial vehicles that can carry out missions in outdoor remote environments is ever increasing. With little or no prior information of the area and relying on limited on-board sensing capabilities, micro air vehicles can carry out these difficult missions if they are able to estimate their pose (position, orientation) and rates (velocity, angular rates) to enable control and mapping of the environment. In this paper we consider the mission of autonomous river mapping with micro aerial vehicles. Rivers are often challenging visual and GPS-degraded environments that require a filter that can handle drop-outs and bad measurements.

There has been some prior work in state estimation for GPS-denied flight of micro aerial vehicles, however, it has focussed on showing the feasibility of the filtering approach. Long missions in GPS-degraded environments are only possible if the filtering approach is robust to outliers, outages, and is able to successfully fuse a wide variety of sensors.

Here we propose a robust unscented kalman filter (UKF) framework that can fuse and accept/reject measurements from multiple sensor modalities such visual odometry (VO), GPS, barometric pressure, and inertial measurement units (IMU) to create a smooth state estimate. To convey the performance of our system we show significant autonomous test flights in outdoor environments.

The authors are with The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A., {basti}@cmu.edu

Fig. 1: The micro aerial vehicle used in the state estimation and control experiments. The vehicle is equipped with an inertial measurement unit, GPS and a stereo camera. These sensors are used to estimate the state of the vehicle.

II. Related Work

There is a rich field of work in the area of combined inertial and visual state estimation for air and ground vehicles. The low cost of commodity IMUs and cameras have also encouraged research in the area. Research into combined inertial-visual state estimation can be broadly split into two categories: loosely and tightly coupled approaches.

Tightly coupled systems calculate expected visual measurements directly in the state estimation by estimating the position of visual landmarks and the vehicle pose at the same time [1], [2]. By incorporating the tracked feature points into the state vector, the covariance between the vehicle pose and the feature points is maintained to improve estimation results at a computational and latency cost. On the other hand, loosely coupled systems allow visual estimation subsystems to calculate pose information which is then combined in a higher level filter or optimization.

Visual odometry is a pose estimation method that computes relative camera motion by comparing sequential images. These algorithms take a structure from motion (SfM) approach and track visual features over two or more camera frames to compute the relative motion [3], [4], [5]. Monocular cameras can be used if additional sensors are used to disambiguate scale [6] or the vehicle operates within a constrained environment [7].

Konolidge et al. [8] present results over large distance for a system combining inertial and feature-point stereo visual odometry measurements. Their low-level IMU filter assumes zero average acceleration to identify the gravity vector and

is too restrictive for MAV applications. Rehder [9] combined frame-to-frame stereo visual odometry, inertial sensing, and extremely sparse GPS readings in a graph optimization to determine state, however the approach is not optimized for a low latency.

A challenge when using visual odometry measurements is determining the correct method for integrating the relative pose measurements into the filtering architecture. Various techniques such as numerical differentiating for average velocity [10], [11], pseudo-absolute measurements [12], and the introduction of delayed states have been explored in the literature. This work makes use of the last approach by using the relative delayed state (or Stochastic cloning) technique introduced by Mourikis and Roumeliotis [13] and therefore is most similar to approaches by [14], [15] who combine visual odometry and inertial sensing in an extended Kalman filter (EKF) for state estimation. However, this work extends state estimation to an Unscented Kalman filter (UKF).

Historically, EKFs are the typical choice for low-latency, efficient computation of state estimation. To propagate Gaussian probability density functions through non-linear processes or measurement functions, the EKF performs a 1st order Taylor series approximation. In cases where the functions show a high degree of local nonlinearity, Iterative Extended Kalman filters (IEKFs) have been utilized [2]. Recent work by Van Der Merwe et al. [16], and more recently by Arasaratnam and Haykin [17] demonstrated the improved accuracy of Sigma-Point Kalman filters and Cubature Kalman Filters over traditional EKFs with similar computational complexity. EKFs still perform well with functions which display high local linearity or when the variance over the state estimate is small. However, as observed by Voigt et al. [15], large variance for unobservable variables can cause EKF filter divergence. Unscented Kalman filters [18], a subclass of Sigma-Point Kalman filters, are better able to handle highly nonlinear functions or larger state variance.

In our work we exploit the prior work, however we combine the real-time filter with a delayed filter state for visual odometry and show signifcant flight test results where actual GPS and visual odometry outages occured.

III. Problem

In this paper we address the problem of fusing multiple redundant motion sensors to generate a smooth consistent state estimate for control and mapping. Our estimator emphasizes smoothness and latency over global accuracy to prevent sudden unexpected control inputs to the vehicle and map jumps after accurate global measurement appear from reaquired GPS. We need the estimator to be able to cope with sparse GPS measurements, and visual odometry dropouts because of signal occlusions and adverse scene geometry. The varying availability and redundancy of the sensors requires careful checking before measurements can be integrated. In addition, since the estimate is used for control onboard a micro aerial vehicle we require a low(<5ms) latency.

IV. Approach

A. Overview

Our filter robustly fuses measurements from a varity of sensor subsystems into a consistent, smooth estimate of the vehicle's state (See Fig. 2). The Unscented Kalman filter (UKF) was chosen based on the demonstrated improvements over the Extended Kalman filter (EKF) for better estimation and filter consistency of non-linear systems. For details on how sigma points are created and used to calculate the estimate state and covariance, the interested reader should refer to [19], [18], [20]. In brief, an EKF propagates the multivariate Gaussian distribution representing the vehicle's state through the nonlinear system by using a first-order linearization around the mean. In contrast, an unscented Kalman filter (UKF) deterministically chooses a set of sigma points that capture the same multivariate Gaussian distribution. These sigma points are propagated through the nonlinear system dynamics and combined to approximate the posterior mean and covariance of the distribution. This derivative-free method achieves at least a 2nd order approximation of the nonlinear system dynamics.

The problem is defined using the following nonlinear equations:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \quad (1)$$

$$z_k = h(x_k, v_k) \quad (2)$$

where the unobserved system state x_k evolves over time as a nonlinear function $f(\cdot)$ of the previous system state x_{k-1}, control inputs u_{k-1}, and process noise w_{k-1}. Observable sensor measurements z_k are a nonlinear function $h(\cdot)$ of the vehicle's state corrupted with observation noise v_k. The process and observation noise are assumed to be normally distributed random variables with zero mean and covariance Q_k and R_k respectively.

$$w_k \sim N(\mathbf{0}, Q_k) \quad (3)$$

$$v_k \sim N(\mathbf{0}, R_k) \quad (4)$$

For vehicle control and mapping, the primary variables of interest are the vehicle's pose in the world frame and the translational velocity in the vehicle body frame. In addition to pose estimation, the filter needs to estimate state variables for specific sensors. MEMS IMUs suffer from a drifting bias along each axis of the accelerometer and gyroscope, which must be estimated to reduce error in the motion prediction. Over long missions, barometric air pressure will drift with changing temperature and weather conditions. A drifting bias term is tracked for the barometric sensor. Finally, to incorporate relative frame-to-frame visual odometry measurements, a delayed state is tracked which captures the vehicle's pose at the previous camera frame. By adding these bias terms and previous pose vectors to the state vector, we attempt to satisfy the Markov assumption that the state vector contains all the information needed to predict the next state given the control input. The vehicle's state is represented as the vector

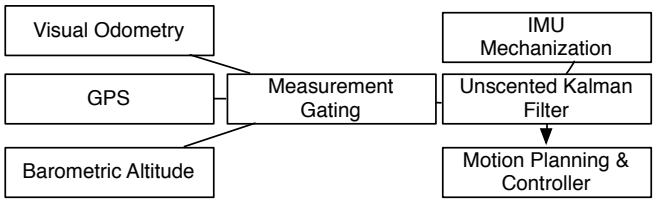


Fig. 2: The Overall Filter Design. The filter combines measurements from several sources and robustifies the integration by gating the measurements to produce a smooth estimate for mapping, motion planning, and control.

$$x = \begin{bmatrix} p^T & \bar{q}^T & v^T & b_a^T & b_g^T & b_p & p_d^T & \bar{q}_d^T \end{bmatrix}^T, \quad (5)$$

where p is the 3-dimensional vehicle's position in the world coordinate system, $\bar{q}$ is a unit quaternion representing the rotation of the world frame to the vehicle's body frame, v is the translational velocity in body coordinates. For simplicity, the vehicle's body frame is defined as the IMU's frame. The bias terms b_a and b_g are 3-by-1 vectors that track the accelerometer and gyroscope drifting bias along each axis. The barometric pressure drifting bias is b_p and the elements p_d and $\bar{q}_d$ represent delayed state vectors, which capture the delayed position and quaternion. The delayed states will be explain in Sect. IV-C.

B. Process Model

The system process model predicts forward in time using IMU mechanization for pose. IMU mechanization use the IMU measurements instead of vehicle control inputs to predict the motion of the vehicle. This approach has the advantage of being vehicle agnostic since new dynamics models are not required for different vehicles or changing vehicle weights or configurations. The IMU provides a measurement of acceleration and angular rate but these measurements have additional terms which must be removed. The following model is used to express the measured acceleration (a_m) and measured rotation rate (ω_m) in the IMU body frame:

$$\begin{aligned} a_m &= a + R_{\bar{q}}^{-1} g + b_a + w_a \quad (6) \\ \omega_m &= \omega + b_g + w_g \quad (7) \end{aligned}$$

The true acceleration a and angular rate ω of the IMU are corrupted by the walking bias terms (b_a and b_g) and wide-band noise (w_a and w_g). The wide-band sensor noise is assumed to be independent between axes of the sensor and is distributed with normal probability around a zero mean. The measured acceleration a_m has an additional contribution from gravity g rotated from the inertial frame into the IMU frame with the rotation matrix $R_{\bar{q}}^{-1}$ defined by the current rotation quaterion.

The true acceleration and angular rate are the inputs for the state propagation equations:

$$\begin{aligned} \dot{p} &= R_{\bar{q}} v \quad (8) \\ \dot{\bar{q}} &= \frac{1}{2}\Omega(\omega)\bar{q} \quad (9) \\ \dot{v} &= a \quad (10) \\ \dot{b}_a &= w_{b_a} \quad (11) \\ \dot{b}_g &= w_{b_g} \quad (12) \\ \dot{b}_p &= w_{b_p} \quad (13) \\ \dot{p}_d &= 0_{3\times 1} \quad (14) \\ \dot{q}_d &= 0_{3\times 1} \quad (15) \end{aligned}$$

$R_{\bar{q}}$ is a rotation matrix formed from the unit rotation quaternion $\bar{q}$.

$$R_{\bar{q}} = \begin{bmatrix} 2\left(q_w^2 + q_x^2\right) - 1 & 2\left(q_x q_y - q_w q_z\right) & 2\left(q_x q_z + q_w q_y\right) \\ 2\left(q_x q_y + q_w q_z\right) & 2\left(q_w^2 + q_y^2\right) - 1 & 2\left(q_y q_z - q_w q_x\right) \\ 2\left(q_x q_z - q_w q_y\right) & 2\left(q_y q_z + q_w q_x\right) & 2\left(q_w^2 + q_z^2\right) - 1 \end{bmatrix} \quad (16)$$

$\Omega(\omega)$ is the quaternion kinematic matrix determined from the 3-by-1 angular rotation rates vector ω measured in IMU coordinates [21].

$$\begin{aligned} \Omega(\omega) &= \begin{bmatrix} 0 & -\omega^T \\ \omega & -[\omega\times] \end{bmatrix} \quad (17) \\ [\omega\times] &= \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \quad (18) \end{aligned}$$

$[\omega\times]$ is the skew-symmetric matrix to perform cross-product operations. The IMU and barometric pressure bias terms evolve over time according to a random walk with a driving noise sources w_{b_a}, w_{b_g}, and w_{b_p} for the accelerometer, gyroscope, and barometric pressure respectively. Since delayed states p_d and $\bar{q}_d$ represent the vehicle's pose at a previous point in time, these estimates do not change in the process model. The noise parameters for the wide-band and walking bias of the accelerometer and gyroscope can be found by performing an Allan variance analysis on the sensors [22].

a) Quaternion Representation: For a singularity-free rotation representation, we parameterize the vehicle's attitude using a unit quaternion. Special consideration must be made when using the quaternion in the UKF since the barycentric mean computed by sigma points does not necessarily represent the correct mean unit quaternion [23]. During prediction and correction steps, a local error state quaternion $\delta\bar{q}$ is defined [21], which aligns the estimated quaternion $\hat{\bar{q}}$ with the true quaternion $\bar{q}$ according to the relation

$$\bar{q} = \delta\bar{q} \otimes \hat{\bar{q}} \quad (19)$$

During prediction and correction steps, the rotation is parameterized by the local error quaternion

$$\delta\bar{q} \simeq \begin{bmatrix} 1 & \frac{1}{2}\delta\theta^T \end{bmatrix}^T \quad (20)$$

$\delta\theta$ is a three dimensional angular error vector which describes attitude errors in a minimal representation. The full state vector (5) is reduced from a 24 element vector to a 22 element error state vector. The error state vector and the resulting 22-by-22 error state covariance matrix are used for filter prediction and correction cycles. The error state vector is defined as

$$x_{se} = \begin{bmatrix} p & \delta\theta & v & b_a & b_g & b_p & p_d & \delta\theta_d \end{bmatrix} \quad (21)$$

Before each prediction/correction cycle, the angular error states are set to zero.

$$\delta\theta = 0_{3x1} \quad (22)$$

$$\delta\theta_d = 0_{3x1} \quad (23)$$

The $0th$ sigma point which represents the mean of the multivariate state distribution will retain zeros for the angular error state. The other $2n$ sigma points will have non-zero angular error based on the Cholesky decomposition of the error state covariance matrix. The attitude quaternion for the remaining sigma points is calculated based on Eq. 19 and 20.

$$^j\bar{q} = ^j\delta\bar{q} \otimes \bar{q}, \, j = 1 \cdots 2n \quad (24)$$

At the end of the prediction step, the weighted mean and variance of the sigma points is used for the predicted mean and covariance of the error state vector and error state covariance respectively. The local error quaternion for each sigma point is found as

$$^j\delta\hat{\bar{q}} = ^j\hat{\bar{q}} \otimes \left(^0\hat{\bar{q}}\right)^{-1}, \, j = 1 \cdots 2n \quad (25)$$

where $^0\hat{\bar{q}}$ is the quaternion found at the $0th$ sigma point.

C. Observation Models

Measurements from barometric pressure, visual odometry, and GPS provide corrections for the filter's predictions. Fig. 3 shows the varying frequencies and asynchronous arrival times of different sensor measurements. As new measurements arrive, the process model is predicted forward in time with the latest IMU measurements and then a measurement correction step is performed. A common problem in real systems is measurement latency caused by processing and communication delays. In our system, the time between the stereo camera pair capturing the images and availability of the stereo visual odometry results is roughly 100 ms [24]. Rather than waiting for latent measurements, the system is controlling based on the latest IMU data (50 Hz with roughly 18 ms lag) to allow for a high frequency, low latency control loop. To integrate latent measurements into the filter, recent incoming measurements and the corresponding filter state are stored in a ring buffer. When a new measurement arrives, the filter is rewound back to the last filter state before the new measurement's time stamp. The filter is then re-run on the new measurement and all subsequent data. Measurements with latency greater than the ring buffer size (in our case, 2 seconds) are discarded. There is additional computational cost for re-running the filter but this method provides the most accurate estimation.

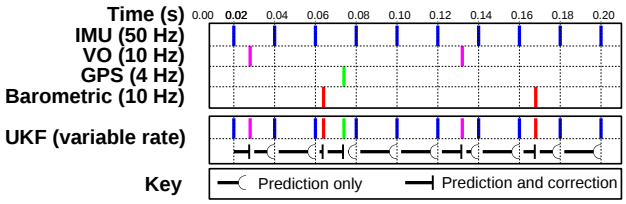


Fig. 3: Asynchronous prediction and correction updates

Stereo Visual Odometry: In our loosely coupled system, stereo visual odometry is an independent subsystem which provides a measurement for the relative translation and orientation between the stereo camera pair at two points in time. Visual features are detected and matched between four images (previous left, previous right, current left, and current right) and the 3D location of matched visual features are triangulated in space using the known stereo camera calibration. The motion of the stereo pair is computed by finding motion parameters which minimize the sum of squared reprojection error for matched features between current and previous image pairs. To provide robustness against outliers such as mismatched features and tracked features on moving objects in the environment, the minimization is wrapped in a RANSAC loop. For our stereo visual odometry system, we make use of the open-source library LIBVISO2 [11]. The state estimation filter is agnostic with respect to the underlying visual odometry implementation. As visual odometry algorithms improve, the newer algorithms can be used.

A challenge when using the relative motion computed by visual odometry is how to best integrate this measurement into the filtering architecture. During the correction step, the filter must be able to derive an estimate based on the current state vector to compare against the actual measurement. The constant acceleration EKF estimation bundled with LIBVISO2 and Oskiper et al. [10] are examples of filters that convert visual odometry measurements to an average translational and rotational velocity over the measurement time period. This average velocity is used to correct the instantaneous velocity estimate of the filter. Unfortunately, the average velocity may be a poor approximation for the instantaneous velocity especially during aggressive changes in direction. An alternative is the relative delayed state method.

Relative Delayed State: Since visual odometry is providing a relative measurement between two instances in time (current and previous camera shutter times), we can augment the state vector to include the previous pose. The visual odometry measurement can then provide a correction for the relative difference between the current and delayed poses. This approach is known as Stochastic cloning and has been used successfully to model the relative measurements provided by wheel odometry and visual odometry [13], [14], [15]. The augmented terms of the state vector are $\begin{bmatrix} p_d & \bar{q}_d \end{bmatrix}$

and represent the delayed position and orientation at the time of the previous camera frame.

The expected measurement is created from the state vector as

$$\hat{z} = h(\hat{x}) = \begin{bmatrix} R_{\hat{\bar{q}}_d}^T (\hat{p} - \hat{p}_d) \\ \hat{\bar{q}}_d^{-1} \otimes \hat{\bar{q}} \end{bmatrix} \tag{26}$$

which is the relative difference in pose between the delayed state and the current state. $R_{\hat{\bar{q}}_d}$ is the rotation matrix formed from the quaternion $\hat{\bar{q}}_d$.

Directly after a measurement update, the delayed portion of the state vector $\begin{bmatrix} p_d & \bar{q}_d \end{bmatrix}$ is set equal to the current pose vector $\begin{bmatrix} p & \bar{q} \end{bmatrix}$. The state covariance matrix is updated by copying the covariance blocks from the current to the delayed pose covariance. This can most easily be shown as a matrix operation [14]:

$$P = TP_rT^T \tag{27}$$

$$T = \begin{bmatrix} I_{3x3} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{3x3} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{3x3} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{3x3} & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{3x3} & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{1x1} \\ I_{3x3} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{3x3} & 0 & 0 & 0 & 0 \end{bmatrix} \tag{28}$$

The new covariance matrix P is found by applying a transformation matrix T to a sub-block of the full error state covariance matrix that we define as P_r immediately after a measurement update. The sub-block of the covariance matrix is only a 16-by-16 matrix and does not include the cloned delayed states. The variance of the delayed state and covariance between the current and the delayed states is set equal to the variance of current state.

GPS: GPS measurements of vehicle's global position and velocity correct the filter's prediction. The expected measurement is created from the state vector as

$$\hat{z} = h(\hat{x}) = \begin{bmatrix} \hat{p} \\ R_{\hat{\bar{q}}}\hat{v} \end{bmatrix} \tag{29}$$

with body velocities v rotated into the global coordinate system. Measurement noise is obtained by the GPS subsystem's self-reported error.

Barometric Pressure: Barometric pressure acts as another loosely coupled subsystem. Altitude is calculated by the subsystem using pressure, temperature, and humidity reported to the UKF for filtering. The expected measurement is

$$\hat{z} = h(\hat{x}) = \begin{bmatrix} \hat{p}_z - \hat{b}_p \end{bmatrix} \tag{30}$$

where $\hat{p}_z$ is the altitude (z) component of the global position. The pressure bias b_p is initialized at vehicle turn-on time as the difference between GPS altitude and barometric pressure altitude estimates. Noise parameters for barometric pressure measurement and the random walk driving noise are determined by performing an Allen variance analysis on a stationary pressure sensor during an offline calibration process.

D. Robustification

Two screening gates are used to reject erroneous or outlying sensor measurements. The first gate rejects sensor measurements with a self-reported error greater than a fixed threshold value. Both the visual odometry and the GPS subsystems estimate their measurement error based on the feature point distribution and the satellite configuration respectively. For example if the estimated horizontal error in GPS position is greater than 5 meters, the measurement is discarded. The second screening gate evaluates the innovation (or measurement residual) using a Chi-squared test. The innovation covariance matrix S_k is the sum of the predicted measurement covariance matrix $P_{\hat{z}_k,\hat{z}_k}$ and the measurement covariance matrix R_k. The predicted measurement covariance matrix is calculated during the UKF measurement update process [18] and the measurement covariance matrix is either generated by the sensor subsystem or found during offline calibration.

$$S_k = P_{\hat{z}_k,\hat{z}_k} + R_k \tag{31}$$

The Chi-squared test:

$$\chi^2 = (z_k - \hat{z}_k)^T S_k^{-1} (z_k - \hat{z}_k) \tag{32}$$

with z_k as the actual measured output of the sensor and $\hat{z}_k$ as the predicted measurement. If χ^2 exceeds the Chi-squared value for a desired confidence level, the measurement update will not complete and the measurement is skipped over. We keep the failed measurement in the measurement ring buffer and will attempt to apply the measurement again if additional (latent) measurements arrive to change the state of the filter. Repeated innovation gate failures of an individual sensor can signal a higher-level health monitoring process that the sensor has failed and should be disabled.

This test performs well at detecting failures caused by sudden jumps in the measurement data such as incorrect visual odometry optimization convergence caused by degenerative scene geometry or large GPS jumps. However, this test will not detect errors caused by drifting biases or sensors that fail in a gradual manner. For example, many commercial GPS units apply a motion model to their output position and a GPS failure will appear to be a smoothed position drift rather than a position jump.

V. Experiments

A. Experimental Setup

The following experiments were performed using a custom built MAV with all sensing and computing on-board. The airframe is an off-the-shelf Mikrokopter octocopter with custom mounts for the added hardware. The vehicle's sensors include a wide baseline stereo camera, spinning LIDAR, barometric pressure sensor, and a MEMS IMU. The visual

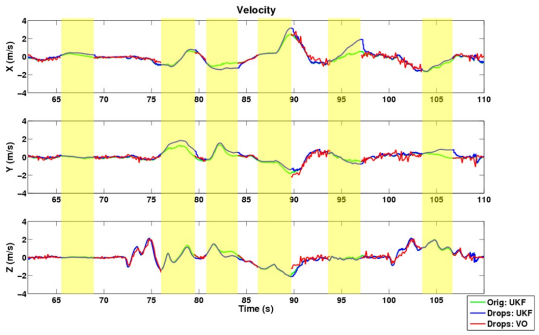


Fig. 4: Velocity error for GPS-denied data set with manually removed visual odometry sections.

odometry, states estimation, mapping, planning, and control algorithms are run in real time on a single board computer with an Intel Core 2 Duo processor.

B. Robustness Tests

Given the noisy MEMS IMU measurements, we cannot expect the state estimate to remain error-free during sensor failures. However, the state covariance matrix P must capture the uncertainty growth during these periods and remain consistent and if GPS or visual odometry should drop out the estimate should remain stable.

We show three sets of results to test the robustness of our filtering approach. In the first experiment, multiple segments of visual odometry measurements are manually removed from a GPS-denied flight data set. The UKF is run over the trimmed data set to compare the performance of the state filter during visual odometry outages versus the performance of the filter without any missing visual odometry data. In the second experiment the octocopter is manually flown from a bright, sunny outdoor environment into a completely dark warehouse, and then back out again. This experiment examines the effect of the slow degradation of visual odometry and GPS to the point of complete failure. Ground truth from GPS is not available for most of warehouse data set since GPS is lost when entering the building. In the third set of results, we present autonomous flights where GPS or visual odometry failed and the filter was able to maintain a consistent estimate.

Fig. 4 plots the velocity error for a GPS-denied data set. The data set contains IMU, barometric, and visual odometry measurements. To test robustness again the lose of both GPS and visual odometry, a total of eleven three-second segments were removed from the visual odometry log. The green line plots the original state estimate. The blue line shows the estimated velocity with the trimmed log. The red line plots the velocity estimate from the VO subsystem alone as a comparison. As expected, the esimated velocity is identical up until the first visual odometry cut at roughly 25 seconds. Z velocity remains the most accurate since barometric pressure provides redundant observations of altitude.

Fig. 5 shows the satellite view for the warehouse data set. The warehouse is completely dark inside and will result

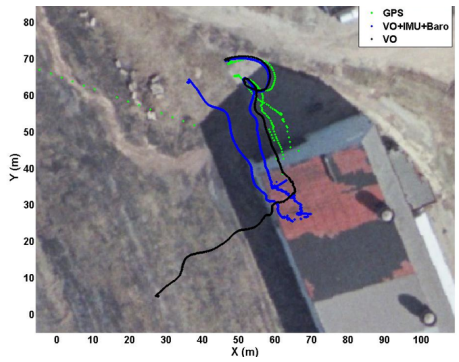
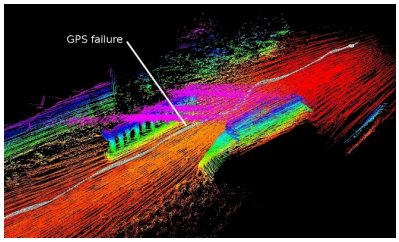
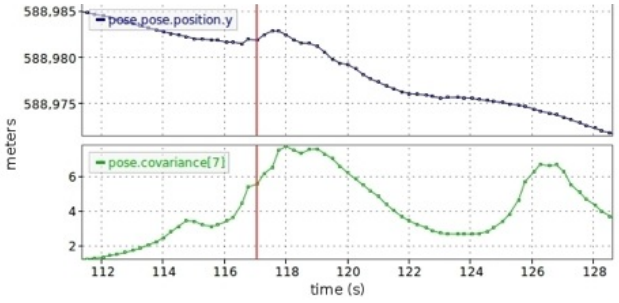


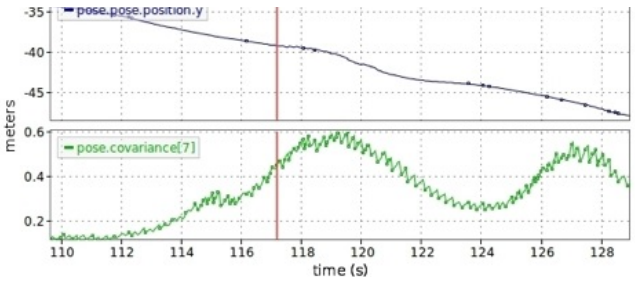
Fig. 5: Satellite view of warehouse data set. Here the vehicle was flown into a dark building to force a GPS and visual odometry failure. The filtered solution (blue) is significantly close to the GPS track (green) than visual odometry alone (black).



(a) Point cloud and flight path



(b) GPS Output

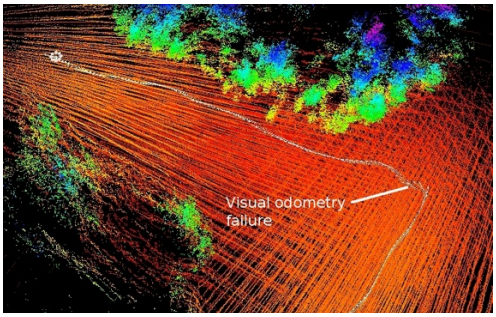


(c) UKF Filter Output

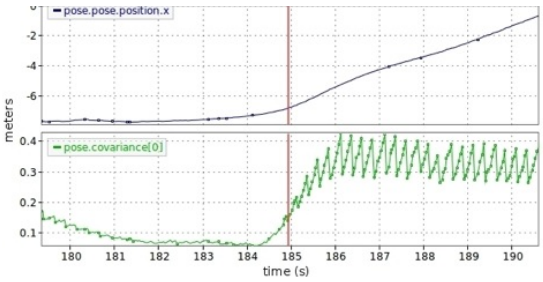
Fig. 6: An example GPS-degraded autonomous flight test. The vehicle autonomously flew under a bridge and the vertical line (in Fig. b and c) shows approximately when the robot enters the area under the bridge. Note that the filter maintains a consistent estimate while the GPS position jumps.

Global Axis	Final Position Error (m)	UKF 3σ Error Estimate (m)
X	-1.73	$\pm$14.78
Y	-12.61	$\pm$14.82
Z	0.10	$\pm$1.25

TABLE I: Final position error of warehouse data set



(a) Point cloud and flight path.



(b) A single axis UKF position and associated covariance.

Fig. 7: An example transition from VO+GPS+IMU+Baro to GPS+IMU+Baro. In this example the visual odometry system stopped working during autonomous flight and the system continued to fly without noticable performance degredation.

in failure for both GPS and VO. GPS is valid during the beginning and end of the flight but unavailable while the rotorcraft is inside the building. Several instances of complete visual odometry failure in the dark warehouse cause any position estimate from visual odometry alone to be grossly incorrect. Even with the visual odometry failures, the UKF is able to maintain a reasonable position estimate using IMU predictions and barometric corrections. Table I reports the final position error of the UKF estimate by using GPS at the landing spot. As demonstrated with other experiments, the filter remains consistent by accurately estimating the upper bound on its position error.

The last set of experiments present autonomous flight tests where the micro aerial vehicle was controlled using the UKF filter output and experienced sensor failures. The filter output is also used to register the onboard lidar data to create a consistent 3D map that is color coded by height in the plots.

Fig.6a shows part of a mission where the vehicle flies under a small train bridge. As the vehicle enters the area under the bridge, Fig. 6b shows the position reported by the GPS jump by several meters. The UKF filter output is not significantly effected by this GPS failure, producing the trajectory in Fig.6c which matches the flight path.

Fig. 7 shows a mission where visual odometry fails permanently. In Fig.7a the vehicle can be seen turning a corner as it navigates towards a distant goal. During the turn, a computer process terminates causing the visual odometry to fail for the rest of the mission. The UKF filter output shown in Fig.7b makes a smooth, stable transition from (VO+GPS+IMU+Baro) on the left of the vertical line to (GPS+IMU+Baro) on the right. The UKF filter covariance shows a transition away from the high local accuracy of visual odometry to the high global accuracy yet poor local accuracy of the GPS. Even with a sustained visual odometry failure the robot is able to complete the mission.

VI. Conclusions and Future Work

In this research we achieved accurate, low-latency state estimation by combining multiple redundant sensors to estimate motion. We use an an Unscented Kalman filter (UKF) to capture the non-linearities in the motion estimates and combined it with robustness tests to enable smooth estimates in visually challenging and degraded GPS environments. The use of IMU mechanization rather than a dynamical vehicle model allows the state estimation system to be transplanted easily and without modification to other vehicles (aerial, ground, or underwater). Overall we demonstrated significant flight test results of GPS-denied operation of a micro aerial vehicle in challenging environments with varying availability of different sensor modalities.

Acknowledgments

The authors gratefully acknowledge Lyle Chamberlain for his help with the vehicle design and experimentation. The work described in this paper is funded by the Office of Naval Research under grant number N00014-10-1-0715.

References

[1] J. Langelaan, *State estimation for autonomous flight in cluttered environments*. PhD thesis, Stanford University, 2006.

[2] D. Strelow and S. Singh, "Motion estimation from image and inertial measurements," *The International Journal of Robotics Research*, vol. 23, pp. 1157 – 1195, December 2004.

[3] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.

[4] G. Chowdhary, E. N. Johnson, D. Magree, A. Wu, and A. Shein, "Gps-denied indoor and outdoor monocular vision aided navigation and control of unmanned aircraft," *Journal of Field Robotics*, vol. 30, no. 3, pp. 415–438, 2013.

[5] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics*, vol. 23, no. 1, pp. 3–20, 2006.

[6] S. Roumeliotis, A. Johnson, and J. Montgomery, "Augmenting inertial navigation with image-based motion estimation," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, vol. 4, pp. 4326–4333, IEEE, 2002.

[7] B. M. Kitt, J. Rehder, A. D. Chambers, M. Schonbein, H. Lategahn, and S. Singh, "Monocular visual odometry using a planar road model to solve scale ambiguity," in *Proc. European Conference on Mobile Robots*, September 2011.

[8] K. Konolige, M. Agrawal, and J. Sola, "Large-scale visual odometry for rough terrain," *Robotics Research*, pp. 201–212, 2011.

[9] J. Rehder, "Optimal state estimation for a river mapping robot," Master's thesis, Technical University Hamburg-Harburg, November 2011.

[10] T. Oskiper, Z. Zhu, S. Samarasekera, and R. Kumar, "Visual odometry system using multiple stereo cameras and inertial measurement unit," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–8, IEEE, 2007.

[11] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *IEEE Intelligent Vehicles Symposium*, (Baden-Baden, Germany), June 2011.

[12] B. Hoffman, E. Baumgartner, T. Huntsberger, P. Schenker, and B. Bhanu, "Improved rover state estimation in challenging terrain," *Autonomous Robots*, vol. 2, pp. 1–20, 1992.

[13] A. Mourikis and S. Roumeliotis, "On the treatment of relative-pose measurements for mobile robot localization," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pp. 2277–2284, IEEE, 2006.

[14] J. Tardif, M. George, M. Laverne, A. Kelly, and A. Stentz, "A new approach to vision-aided inertial navigation," in *IROS*, pp. 4161–4168, IEEE, 2010.

[15] R. Voigt, J. Nikolic, C. Hurzeler, S. Weiss, L. Kneip, and R. Siegwart, "Robust embedded egomotion estimation," in *IROS*, pp. 2694–2699, IEEE, 2011.

[16] R. Van Der Merwe, E. Wan, and S. Julier, "Sigma-point kalman filters for nonlinear estimation and sensor-fusion-applications to integrated navigation," in *Proceedings of the AIAA Guidance, Navigation & Control Conference*, pp. 16–19, 2004.

[17] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *Automatic Control, IEEE Transactions on*, vol. 54, no. 6, pp. 1254–1269, 2009.

[18] S. Julier, "The scaled unscented transformation," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6, pp. 4555–4559, IEEE, 2002.

[19] R. Van Der Merwe, *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. PhD thesis, University of Stellenbosch, 2004.

[20] J. Crassidis, "Sigma-point kalman filtering for integrated gps and inertial navigation," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 2, pp. 750–756, 2006.

[21] N. Trawny and S. Roumeliotis, "Indirect kalman filter for 3d attitude estimation," *University of Minnesota, Dept. of Comp. Sci. & Eng., Tech. Rep*, vol. 2, 2005.

[22] H. Hou, *Modeling inertial sensors errors using Allan variance*. University of Calgary, Department of Geomatics Engineering, 2004.

[23] J. Kelly and G. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[24] J. Haines, "Vision-based control of a multi-rotor helicopter," Master's thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, December 2011.