

Self-Calibrating Multi-Sensor Fusion with Probabilistic Measurement Validation for Seamless Sensor Switching on a UAV

Karol Hausman¹ Stephan Weiss² Roland Brockers³ Larry Matthies³ Gaurav S. Sukhatme¹

Abstract—Fusing data from multiple sensors on-board a mobile platform can significantly augment its state estimation abilities and enable autonomous traversals of different domains by adapting to changing signal availabilities. However, due to the need for accurate calibration and initialization of the sensor ensemble as well as coping with erroneous measurements that are acquired at different rates with various delays, multi-sensor fusion still remains a challenge.

In this paper, we introduce a novel multi-sensor fusion approach for agile aerial vehicles that allows for measurement validation and seamless switching between sensors based on statistical signal quality analysis. Moreover, it is capable of self-initialization of its extrinsic sensor states. These initialized states are maintained in the framework such that the system can continuously self-calibrate. We implement this framework on-board a small aerial vehicle and demonstrate the effectiveness of the above capabilities on real data. As an example, we fuse GPS data, ultra-wideband (UWB) range measurements, visual pose estimates, and IMU data. Our experiments demonstrate that our system is able to seamlessly filter and switch between different sensors modalities during run time.

I. INTRODUCTION

Multi-rotor Unmanned Aerial Vehicles (UAVs) are inherently unstable platforms and therefore require uninterrupted and accurate state estimation at a high frequency to facilitate safe control. Using multiple sensors provides redundancy, improves robustness of state estimation and enables the platform to perform cross-domain missions. However, all sensors must be calibrated with respect to each other, and their contributions properly merged, in order to optimally perform state estimation. In addition, the system should be able to automatically detect and ignore faulty measurements or sensors. All these requirements still render multi-sensor fusion a challenge. Nonetheless, if implemented correctly, such a multi-sensor fusion framework can significantly contribute to the robustness, versatility, and usability of the platform. For example, a UAV could detect the moment it enters a GPS-obstructed environment and stop including the GPS measurements that otherwise can negatively influence the state estimate.

In this paper, we propose a self-calibrating multi-sensor state-estimation system where additional sensors can be added in a modular fashion. In particular, we focus on fusing vision, GPS, and UWB range measurements with no need for



Fig. 1. The Asctec Hummingbird UAV used during the experiments.

external calibration or initialization. Moreover, we include a probabilistic sensor switching ability in the system and propose a statistical measurement validation algorithm that finds the optimal sensor suite for any given condition. We argue that the ability to dynamically make decisions about sensor information on the fly is essential for deploying UAVs in uncertain environments, and the contributions of this paper are a step towards this goal.

The key contributions of our approach are: a) a continuously self-calibrating modular multi-sensor fusion framework, b) a probabilistic switching ability and statistical measurement validation that selects the correct sensor suite in the given conditions, c) self-initialization and calibration of the range sensor states using data gathered during the initial phase of the flight. We evaluate our approach experimentally using the Asctec Hummingbird UAV depicted in Fig. 1. As a use case, we consider take-off and precision landing on a recharging station where, in addition to the IMU, the UAV has access to three sensor modalities: GPS, vision and range. The experimental sensor setup for this scenario is depicted in Fig. 2.

II. RELATED WORK

Multi-sensor fusion for aerial vehicles has recently received increased attention. Typically, indirect formulations of Extended (EKF) [7], or Unscented (UKF) Kalman Filters [13] are used as theoretical frameworks for sensor fusion. Although in most cases the underlying technique is similar, the sensor suite for a UAV varies across different approaches. Researchers have used cameras [15], laser-range finders [12], range modules [11], GPS [2], event-based cameras [10] and

¹Karol Hausman and Gaurav S. Sukhatme are with the Department of Computer Science, University of Southern California, Los Angeles, CA 90089, USA. hausman@usc.edu

²Stephan Weiss is with the Alpen-Adria Universitat Klagenfurt, 9020 Klagenfurt, Austria.

³Roland Brockers and Larry Matthies are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA.

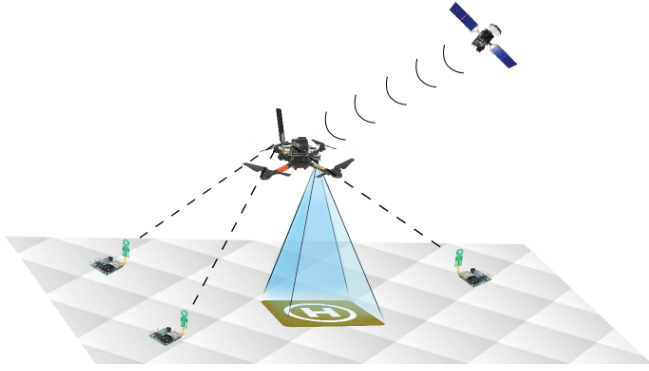


Fig. 2. Experimental setup for precision landing. The quadrotor obtains the measurements from three UWB range modules, a visual fiducial and a GPS sensor.

inertial measurement units (IMU) to estimate the state of a flying vehicle.

Although impressive state estimation results have been shown using the above sensors, state-of-the-art algorithms do not make use of the potential for self-initialization, self-calibration, and intelligent sensor switching when combining different sensor modalities. For example, extrinsic calibration of the camera or other proprioceptive and exteroceptive sensor calibration can be performed during run time as shown for individual sensors in [4, 5, 14]. In this paper, we achieve continuous self-calibration of the full system by estimating additional auxiliary states that are relevant to the sensor suite.

One of the contributions of this paper is the addition of a range sensor without the need of initialization or external calibration. Range sensors have been used previously in various sensor fusion approaches. Hol et al. [3] proposed a tightly coupled UWB-IMU pose estimation system that is able to handle outliers caused by multipath effects. In [11] the authors fuse the UWB range measurements with IMU measurements to estimate the pose and velocity of a quadrotor. Gross et al. [2] presented an approach where a differential GPS sensor is tightly coupled with an UWB range module for aircraft formation flights. However, none of the above approaches performs a full self-calibration and initialization of the system (including the position of the range modules in the world frame). In this work, we present an initialization method to approximate the range module position in the world frame, and we include all the calibration states in the state vector so that they can be estimated continuously during flight.

Probabilistic techniques have been used to perform sensor switching and sensor measurement validation. Leutenegger and Siegwart [6] proposed a back-up system to handle GPS outages for a fixed wing aircraft system. Mourikis et al. [9] presented an approach where GPS and visual measurements were used in different parts of the experiment. In [7] the authors showed how relative and absolute measurements can be handled within one EKF-based framework. In this paper, we extend the approach from Lynen et al. [7] to be able to detect erroneous sensors and measurement outliers. The technique we propose performs probabilistic sensor switching and measurement validation based on the current

conditions of both the environment and the system to select an optimal sensor suite.

III. EXTENDED KALMAN FILTER

In order to optimally fuse multiple types of measurements, we use a framework based on the indirect formulation of an iterated EKF where the state prediction is driven by IMU measurements. This treats the IMU as a core sensor, allowing the division of the state vector into a computationally efficient core part and an augmented part that includes all additional sensors.

A. Core State

The core state consists of the following states:

$$\mathbf{x}_{core}^T = [\mathbf{p}_w^T, \mathbf{v}_w^T, \mathbf{q}_w^T, \mathbf{b}_\omega^T, \mathbf{b}_a^T] \quad (1)$$

where \mathbf{p}_w^i , \mathbf{v}_w^i and \mathbf{q}_w^i are the position, velocity and orientation (represented as a quaternion) of the IMU in the world frame. The variables \mathbf{b}_ω , \mathbf{b}_a correspond to gyroscope and accelerometer biases. Since the IMU biases can change over time they are modeled as a random process:

$$\dot{\mathbf{b}}_\omega = \mathbf{n}_{b_\omega}, \quad \dot{\mathbf{b}}_a = \mathbf{n}_{b_a} \quad (2)$$

where \mathbf{n}_{b_ω} and \mathbf{n}_{b_a} are each assumed to be zero-mean white Gaussian noise.

The core state is governed by the following differential equations:

$$\dot{\mathbf{p}}_w^i = \mathbf{v}_w^i \quad (3)$$

$$\dot{\mathbf{v}}_w^i = \mathbf{C}_{(\mathbf{q}_w^i)}^T (\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) - \mathbf{g} \quad (4)$$

$$\dot{\mathbf{q}}_w^i = \frac{1}{2} \Omega(\omega_m - \mathbf{b}_\omega - \mathbf{n}_\omega) \mathbf{q}_w^i \quad (5)$$

$$\dot{\mathbf{b}}_\omega = \mathbf{n}_{b_\omega}, \quad \dot{\mathbf{b}}_a = \mathbf{n}_{b_a} \quad (6)$$

where $\mathbf{C}_{(\mathbf{q})}$ corresponds to the rotation matrix obtained from the quaternion \mathbf{q} , $\Omega(\omega)$ represents the skew symmetric quaternion multiplication matrix of ω , \mathbf{a}_m is the measured acceleration and ω_m is the angular velocity with white Gaussian noise \mathbf{n}_a and \mathbf{n}_ω .

Starting from the initial state defined in Eq. 1, we define the error state as:

$$\tilde{\mathbf{x}}_{core}^T = [\Delta \mathbf{p}_w^i, \Delta \mathbf{v}_w^i, \delta \Theta_w^i, \Delta \mathbf{b}_\omega^T, \Delta \mathbf{b}_a^T] \quad (7)$$

where $\tilde{\mathbf{x}}$ is the error between the real state value \mathbf{x} and the state estimate $\hat{\mathbf{x}}$. For quaternions the error state is defined as: $\delta \mathbf{q} = \mathbf{q} \otimes \hat{\mathbf{q}} \approx [1 \quad \frac{1}{2} \delta \Theta^T]^T$.

B. State Augmentation

In addition to the core states, each sensor extends the state vector with additional auxiliary states. These states are sensor-dependent and correspond to the calibration parameters that need to be estimated to make the system fully self-calibrating. In the following, we present the state augmentation for the sensors presented in the example scenario shown in Fig. 2.

1) *GPS position sensor (3 DoF)*: For a GPS sensor, only the relative position \mathbf{p}_i^p of the GPS module with respect to the IMU is needed. Note that, since GPS provides a measurement of the global position, the attitude between the sensor and the IMU can be omitted. The corresponding error state is defined as:

$$\tilde{\mathbf{x}}_{gps}^T = \Delta \mathbf{p}_i^{pT}. \quad (8)$$

2) *Vision-based pose sensor (6 DoF)*: A vision-based pose sensor requires the relative transformation between the camera and the IMU $\mathbf{p}_i^c, \mathbf{q}_i^c$. In addition, since the vehicle is localized in the global frame of the GPS, the world-relative position and orientation of the vision frame $\mathbf{p}_w^v, \mathbf{q}_w^v$ are needed with respect to the world frame:

$$\tilde{\mathbf{x}}_{pose}^T = [\Delta \mathbf{p}_i^{cT}, \delta \Theta_i^{cT}, \Delta \mathbf{p}_w^{vT}, \delta \Theta_w^{vT}]. \quad (9)$$

3) *Range sensor (1DoF)*: In order to use an on-board range sensor in the system, the position of the range module antenna on the vehicle \mathbf{p}_i^u is required. In addition, the system needs to estimate the position of each external range module that is deployed in the field. Thus, an extra state that represents the position of each external range module antenna in the world frame \mathbf{p}_w^r is necessary:

$$\tilde{\mathbf{x}}_{range}^T = [\Delta \mathbf{p}_i^{uT}, \Delta \mathbf{p}_w^{rT}]. \quad (10)$$

4) *Sensor switching and observability*: The design of the system enables the user to easily include additional sensors as long as the auxiliary states are included and the measurement models are defined. In fact, it is this modular design and consequent inclusion of the required auxiliary states that enable the seamless switching between different sensor modalities. Each measurement sensor acts on its auxiliary states and the core state; the other auxiliary are not affected by it. Since the auxiliary states are constant over time, they are modeled to not be affected by process noise. Consequently, the mean and covariance of the auxiliary states that do not belong to the current sensor measurement are not changed. For example, the position of the vision frame \mathbf{p}_w^v remains unchanged if there is no visual measurement. In some configurations (e.g. given a single range measurement) there may be either core or auxiliary states that are unobservable. In this case, the covariance of the affected states grows until a suitable measurement is included. A non-linear observability analysis shows that the core state is fully observable given a GPS measurement, visual measurement, or measurements of three different range modules. We refer to the approach used in [5] for a detailed discussion about the observability of all states.

In this paper, we show the example of integrating one GPS sensor, one vision-based pose sensor, and three range sensors. The final error state consists of the core error states (Eq. 7) and all the auxiliary error states defined in Eqs. 8,9,10, thus:

$$\tilde{\mathbf{x}}^T = [\tilde{\mathbf{x}}_{core}^T, \tilde{\mathbf{x}}_{gps}^T, \tilde{\mathbf{x}}_{pose}^T, \tilde{\mathbf{x}}_{range1}^T, \tilde{\mathbf{x}}_{range2}^T, \tilde{\mathbf{x}}_{range3}^T]. \quad (11)$$

C. Measurement Models

In the following, we present the models for the measurements \mathbf{z} . Due to our indirect formulation of the EKF, the filter uses the measurement error $\tilde{\mathbf{z}}$ which is modeled as $\tilde{\mathbf{z}} = \mathbf{z} - \hat{\mathbf{z}}$. This can be linearized to $\tilde{\mathbf{z}} \approx \mathbf{H}\tilde{\mathbf{x}}$ where $\mathbf{H} = \frac{\partial h}{\partial \tilde{\mathbf{x}}}$ is the Jacobian of the measurement function h with respect to the error state $\tilde{\mathbf{x}}$. For brevity, we specify the measurement equations but omit the formulation of the \mathbf{H} matrices.

1) *GPS position*: The GPS sensor returns position in the world frame:

$$\mathbf{z}_{gps} = \mathbf{p}_w^i + \mathbf{C}_{(\mathbf{q}_w^i)}^T \mathbf{p}_i^p + \mathbf{n}_{z_{gps}} \quad (12)$$

where $\mathbf{n}_{z_{gps}}$ is white Gaussian measurement noise.

2) *Vision-based pose*: The vision-based sensor measures the 6 DoF pose of the camera with respect to the vision coordinate frame. In our case, the origin of the vision coordinate frame is defined by an artificial marker whose world-relative position is estimated as well (see \mathbf{q}_w^v in Eq. 9). The measurement consists of the translational part \mathbf{z}_{pose_p} and the rotational part \mathbf{z}_{pose_q} :

$$\begin{bmatrix} \mathbf{z}_{pose_p} \\ \mathbf{z}_{pose_q} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{(\mathbf{q}_w^v)}(\mathbf{p}_w^i + \mathbf{C}_{(\mathbf{q}_w^i)}^T \mathbf{p}_i^c) + \mathbf{p}_w^v + \mathbf{n}_{z_{pose_p}} \\ \mathbf{q}_i^c \otimes \mathbf{q}_w^i \otimes \mathbf{q}_w^{v-1} \otimes \delta \mathbf{q}_{\mathbf{n}_{z_{pose_q}}} \end{bmatrix}. \quad (13)$$

Similarly to other measurements, the vision-based pose measurement also contains white Gaussian noise $\mathbf{n}_{z_{pose_p}}$ and $\mathbf{n}_{z_{pose_q}}$.

3) *Range*: The range sensor measures the distance between two antennas, where one is on the UAV and the other is in the field. The measurement model of this sensor also includes the relative position of the UAV-antenna with respect to the IMU and the position of the field-antenna in the world frame (see \mathbf{p}_i^u and \mathbf{p}_w^r in Eq. 10). The measurement model is:

$$z_{range} = \left\| \mathbf{p}_w^i + \mathbf{C}_{(\mathbf{q}_w^i)}^T \mathbf{p}_i^u - \mathbf{p}_w^r \right\|^2 + \mathbf{n}_{z_{range}} \quad (14)$$

with Gaussian noise $\mathbf{n}_{z_{range}}$.

D. Time delays

When fusing multiple sensors, one of the important challenges is the time delay between the measurements. In order to handle the delays, we, similarly to [7], introduce a time-sorted buffer for states and measurements. When an older measurement occurs out of order, it is applied to the closest state in a buffer and an updated state is propagated to the most current time stamp. This procedure is repeated for all old measurements to ensure that each measurement is applied to the most up-to-date state. For more details of this approach see [7].

IV. STATISTICAL MEASUREMENT VALIDATION

In real-world scenarios sensor measurements will often be corrupted with noise and sparse outliers due to faults in the sensors or changes in conditions (e.g. transitions from outdoor to indoor). In such cases, the underlying assumption of the EKF that the measurement residuals are Gaussian

distributed is violated, and updating the filter using such measurements can introduce large errors into the state and covariance estimates. Instead, it is desirable to automatically detect outliers and remove these measurements before they corrupt the filter. Using the probabilistic analysis described in [1], we use a hypothesis test to automatically identify and remove erroneous measurements in our multi-sensor fusion framework.

The EKF formulation includes the calculation of the measurement innovation \mathbf{r} and its covariance \mathbf{S} :

$$\mathbf{r} = \mathbf{z} - h(\hat{\mathbf{x}}) \quad (15)$$

$$\mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R} \quad (16)$$

where $h(\hat{\mathbf{x}})$ is the measurement function, \mathbf{P} is the current covariance of the state estimate and \mathbf{R} is the covariance of the measurement \mathbf{z} .

Since the residual vector \mathbf{r} is assumed to be distributed according to a multivariate Gaussian with covariance matrix \mathbf{S} , the normalized sum of squares of its values should be distributed according to the Chi-Squared (χ^2) distribution with as many degrees of freedom as there are measurements. We perform a one-sided hypothesis test to determine whether the residuals are likely to have come from this distribution:

$$\mathbf{r}^T \mathbf{S}^{-1} \mathbf{r} < \chi^2(0.95). \quad (17)$$

The threshold on the Chi-Squared test is set to 95%. Values above this threshold are labeled as outliers and not used to update the filter.

The Chi-Squared test provides a modular test of measurement validity. Moreover, the implementation of new sensor modules automatically includes the probabilistic validation.

V. INITIALIZATION

In order for the filter to function correctly, the initial state values have to be in the region of convergence of the EKF. In addition, the corresponding initial covariance of the state \mathbf{P} must be estimated accordingly to capture the initial uncertainty of the states.

Some of the states in the state vector can be easily initialized (e.g. setting \mathbf{b}_w and \mathbf{b}_a to zero, or using a GPS and visual measurement to determine \mathbf{p}_w^v and \mathbf{q}_w^v). However, the initial measurement of the world-relative range module position cannot be estimated based on the initial measurements and it may be tedious to measure for the user. Therefore, we facilitate the initialization process by enabling self-initialization of the range module position in the world frame.

A. Least Squares Solution

Since the initial position of the field module is not known and cannot be estimated from a single measurement, an initialization strategy is needed that takes into account multiple measurements over time to resolve the position ambiguity. In addition, due to the non-linearity of the system, the position of the module might be outside of the beacon of convergence if initialized arbitrarily. We propose a set of linear equations

which enables us to solve for the world-relative field range module position using the least squares method.

Let \mathbf{p}_w^u be the position of the UAV range module in the world frame:

$$\mathbf{p}_w^u = \mathbf{C}_{(\mathbf{q}_w^i)}^T \mathbf{p}_i^u + \mathbf{p}_w^i. \quad (18)$$

Given the fact that the range measurement can be modeled as:

$$z_{range} = \|\mathbf{p}_w^u - \mathbf{p}_w^r\|^2 \quad (19)$$

where the right side of the equation for each dimension results in terms of the form:

$$\mathbf{p}_{wx}^u{}^2 - 2\mathbf{p}_{wx}^u \mathbf{p}_{wx}^r + \mathbf{p}_{wx}^r{}^2, \quad (20)$$

the system can be linearized by grouping the quadratic terms into a distance measure:

$$d_{r2w}{}^2 = \mathbf{p}_{wx}^r{}^2 + \mathbf{p}_{wy}^r{}^2 + \mathbf{p}_{wz}^r{}^2. \quad (21)$$

Rearranging the terms in the range measurement model in a set of linear equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ leads to:

$$\begin{bmatrix} -2\mathbf{p}_{wx}^u(t) & -2\mathbf{p}_{wy}^u(t) & -2\mathbf{p}_{wz}^u(t) & 1 \\ -2\mathbf{p}_{wx}^u(t+1) & -2\mathbf{p}_{wy}^u(t+1) & -2\mathbf{p}_{wz}^u(t+1) & 1 \\ -2\mathbf{p}_{wx}^u(t+2) & -2\mathbf{p}_{wy}^u(t+2) & -2\mathbf{p}_{wz}^u(t+2) & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{p}_{wx}^r \\ \mathbf{p}_{wy}^r \\ \mathbf{p}_{wz}^r \\ d_{r2w} \end{bmatrix} = \begin{bmatrix} z_{range}^{(t)}{}^2 - d_{p2w}^{(t)}{}^2 \\ z_{range}^{(t+1)}{}^2 - d_{p2w}^{(t+1)}{}^2 \\ z_{range}^{(t+2)}{}^2 - d_{p2w}^{(t+2)}{}^2 \\ \vdots \end{bmatrix} \quad (22)$$

where $z^{(t)}$ represents the measurement obtained at the t -th time stamp, \mathbf{z}_x denotes the x component of the vector \mathbf{z} and d_{p2w} and d_{r2w} correspond to the distance between the IMU position in the world frame, and the field range module position in the world frame, respectively.

The covariance of the parameters can be estimated [8] as:

$$\Sigma_{\mathbf{x}} = \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2}{n - k} (\mathbf{A}^T \mathbf{A})^{-1} \quad (23)$$

where n and k are the number of rows and columns of the \mathbf{A} matrix, which is finally used to initialize the range block of the state covariance \mathbf{P} .

$$\mathbf{P}_{\mathbf{p}_w^r} = \Sigma_{\mathbf{p}_w^r}. \quad (24)$$

B. Measurement Inclusion Test

One crucial aspect needed to correctly initialize the range module position is the quality of the \mathbf{A} matrix in the least squares method. If the rank of the \mathbf{A} matrix is smaller than 4 the least squares solution will be ill conditioned and likely incorrect. Since the \mathbf{A} matrix consists of the measurements and state estimates over time, we can select which measurements and states should be included for the final solution. Therefore, to check the quality of the least squares solution, a measurement inclusion test is preformed.

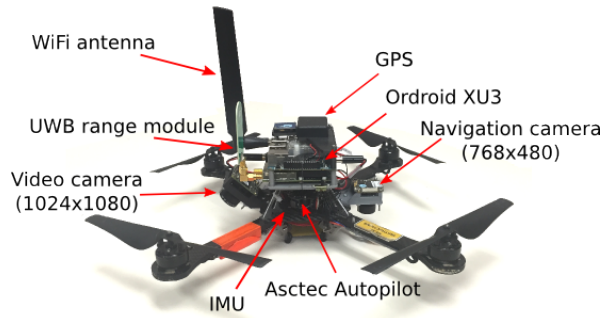


Fig. 3. The experimental platform (Asctec Hummingbird) with all sensors and the processing hardware.

This ensures that the motion of the vehicle was sufficient to make the position of the range module observable. In theory, more measurements increase the accuracy of the solution. However, in our implementation we keep the size of the \mathbf{A} matrix small (15 rows) for efficiency reasons. Given this size constraint, it is important to choose the best data points available to correctly initialize the range module position.

In order to check if an additional row of the \mathbf{A} matrix (i.e. additional measurement and state estimate) improves the solution of the least squares, we perform a matrix condition number analysis. The condition number of a matrix is defined as the ratio between the largest to the smallest singular value of the matrix. Upon each range measurement, the condition number of the current \mathbf{A} matrix is calculated and stored. At the next step, the condition number of the \mathbf{A} matrix extended by the most current linear equation is computed. If the condition number of the matrix decreased, the row is added to the matrix as it contributes to the final solution.

Since the computation of the condition number of the \mathbf{A} matrix upon each measurement is computationally expensive, a simple test is used to gate the measurements before evaluation: If the traveled distance between two measurements is below a threshold, it is discarded. This way we can avoid adding the measurements and computing the singular value decomposition while the vehicle is not moving.

C. Post-Initialization Test

In order to smoothly initialize the range module position and start including range measurements into the EKF, we must avoid abrupt changes in other parts of the state vector. This is achieved by a post-initialization test.

Once the initial state value and its covariance are estimated according to Eqs.22- 24, they are not immediately included in the state. Instead, a number of Chi-Squared hypothesis tests are performed to ensure that the range measurement will not cause an abrupt change in the other state estimates.

If the majority of the tested measurements fail the Chi-Squared tests, the range module position is reinitialized. Otherwise, future range measurements are added to the EKF state estimation.

VI. EVALUATION AND DISCUSSION

A. Experimental Setup

In the following experiments, we use an Asctec Hummingbird quadrotor equipped with an UWB ranging module (Timedomain P410), a downward looking camera (Matrix Vision, blueFox MLC-200, 768x480, 110 degree FOV), a GPS receiver (included in Asctec Hummingbird), and an additional flight computer (Odroid XU3 with Samsung Exynos5522, 2GHz) (Fig. 3). All processing, including the state estimation, is performed on-board. In addition, we use three Timedomain P410 UWB range modules placed arbitrarily in the environment. The range module measurement is available at 30Hz and its std. dev. is equal to 0.07m. As the vision-based pose sensor we employ a visual fiducial of a known size which is seen by the navigation camera. In order to estimate the pose of the vehicle from the image of the fiducial, we use the openly available ar-track-alvar ROS package¹. The vision-based pose measurement is acquired at 30Hz with a std. dev. of 0.02m and 0.06 radians. For all experiments we simulate GPS by using position information from a VICON motion capture system at 5 Hz corrupted by white Gaussian noise with std. dev. of 0.2m, which is the measured standard deviation of the Asctec GPS module. However, we are aware that this simulation does not reflect the full complexity of the true GPS signal.

B. Chi-Squared Test Evaluation

In the first experiment, we evaluated the importance of the inclusion of the Chi-Squared test to the filter. In order to simulate faulty measurements, we randomly added a small error to the measurement of the visual fiducial. We chose the error to be in the same order of magnitude as the sensor accuracy (i.e. 0.02m) which, intuitively, the proposed signal analysis should be able to detect as an outlier and reject. The visual pose measurement was obtained at 30Hz and the IMU propagation was performed at 100Hz. The error was added randomly to 5% of all measurements.

Fig. 4 shows an example of a Chi-Squared test experiment. The spikes in the top plot correspond to the residuals obtained from the erroneous measurements. As shown in the middle plot of Fig. 4, none of the erroneous measurements passed the test and therefore, did not corrupt the state estimates.

In order to evaluate the influence of the Chi-Squared test on the state estimation, the error of the vehicle position with and without performing the test is compared in Fig. 4-bottom. The ground truth data was obtained from the motion capture system. For the evaluation, we use the visual pose sensor as the only sensor in the system so that the errors are not influenced by any other source of measurements.

The position error of the system without the Chi-Squared test is noticeably larger, especially when the erroneous measurements were introduced (the larger errors correspond to the spikes in Fig. 4-top and middle). The root mean squared error (RMSE) for the filter with the Chi-Squared

¹http://wiki.ros.org/ar_track_alvar

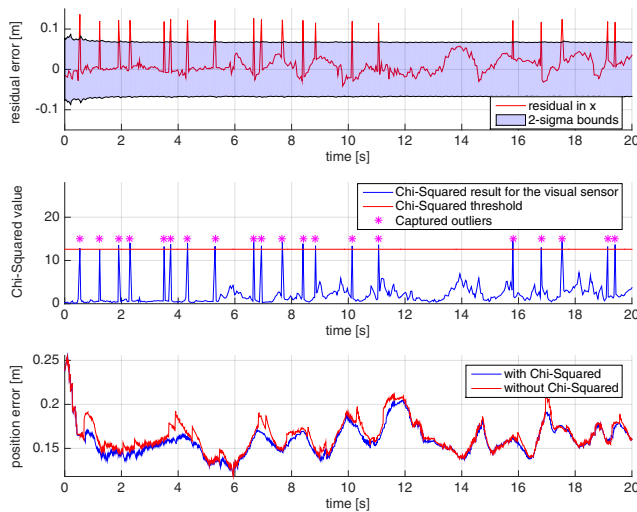


Fig. 4. Chi-Squared test experiment. Top: the residual error in the x dimension for the visual pose measurement surrounded with the 2-sigma bounds (since it corresponds to the 95% of the Gaussian distribution) based on the covariance matrix \mathbf{S} . Middle: The corresponding result of the Chi-Squared test. The Chi-Squared threshold is set to 95%. None of the erroneous measurements passed the test. Bottom: The position error with and without the Chi-Squared test computed based on the ground truth data from the motion capture system. During the experiment, only the visual pose measurements were available to the system.

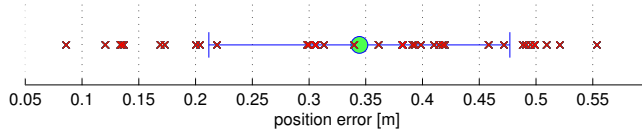


Fig. 5. Statistics over 40 runs for the range module initialization for the real robot experiment. The data points are depicted in red. The mean error is equal to $0.3444m \pm 0.1326m$.

test was equal to 0.1613, whereas the RMSE for the filter without the Chi-Squared test was equal to 0.1659. It is worth noting that the state drifts based on the IMU integration when the measurement is not included (since this is the only measurement in the filter). Despite this and the fact that the measurement was corrupted by a small error (average of 0.02m), the use of the Chi-Squared test has a noticeable influence on the state estimates. Based on these results, the Chi-Squared test was used for all of the sensors for the rest of the experiments.

C. Range Module Initialization and Convergence Experiment

To evaluate the quality of the initialization of the range modules, we performed multiple flights and initialized each of the three modules if sufficient information was gathered (see Sec. V-B). Afterwards, we compared their estimated positions to the ground truth data obtained through the motion capture system.

Fig. 5 shows the error statistics for 40 runs of range module position initialization that was performed on the UAV. The average distance to the range modules during the experiment was approximately 2m which results in an average initialization error of 17%. It is important to remember that the quality of the initialization is highly dependent on the

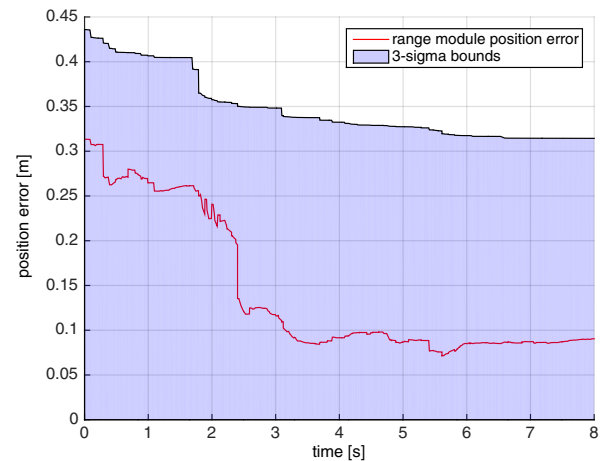


Fig. 6. An example of range module convergence. 3-sigma bounds was computed based on the average of the diagonal terms of the covariance of the x, y and z range module position.

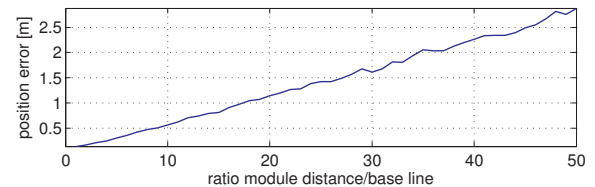


Fig. 7. Monte Carlo simulation for the position error of the range module position computed with the least squares method explained in Sec. V-A. The x-axis shows the ratio between the distance d_{r2w} and the baseline the UAV moved during acquiring the necessary measurements for the least squares.

motion presented to the system. In this set of experiments, the vehicle performed an elliptical motion with a varying radius of approximately 0.4m.

Despite the initial error of the range module position estimate, the system was able to converge. An example of a convergence process is shown in Fig. 6. The range module position was initialized with an initial error of 0.32m and it took approximately 3s to converge to the position that resulted in a 0.08m error. Taking into account that the range measurement std. dev. is equal to 0.07m (however, it may be larger depending on the environment), the final error of 0.08m is acceptable. The uncertainty of the range module position estimate is decreasing (see the blue field in Fig. 6) which is an expected behavior of the filter.

We performed a series of simulation experiments to evaluate the quality of the initialization with respect to the motion extended to the UAV. As the UAV moves, measurements

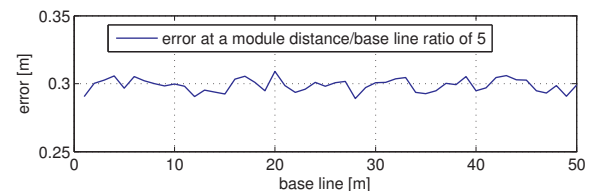


Fig. 8. Monte Carlo simulation for the position error of the range module position computed with the least squares method explained in Sec. V-A. The x-axis shows the baseline the UAV moved during acquiring the necessary measurements for the least squares. The error remains the same as long as the distance d_{r2w} is adapted such that the ratio of these two values is constant.

of the range modules are acquired at different positions to populate the least squares problem in Eq. 22. We define the action radius during this initialization as *baseline*. Naturally, the precision of the least squares solution is a function of the ratio between the distance of the range module to the UAV and the baseline. We simulated 2000 initializations with realistic noise values for the position estimates and the distance measurements. Fig. 7 shows the influence on the initialization accuracy when varying this ratio. Fig. 8 shows that the error is constant if the baseline is adapted to the range module distance (i.e. if the ratio is constant). The simulated accuracy is slightly below the one we observe on true data (Fig. 5).

D. Full System Evaluation

In the last experiment, we evaluated the behavior of the full system. The UAV performed a trajectory shown in Fig. 9-top. During its flight time, different sensor modalities were switched on and off to show the capabilities of the system. The quadrotor started using GPS, visual fiducial and 3 UWB-range sensors. The first phase of the flight consisted of circular motions in 3 dimensions to correctly initialize the range module states (green line in Fig. 9-middle and bottom). Once the modules were initialized, the UAV performed a square-like figure. During this time, the visual measurements were not acquired and the range measurements were available only at the beginning and the end of this part of the trajectory (see columns in Fig. 9-middle and bottom). In the final phase, the vehicle was again using all 3 sensor modalities to accurately land on the visual fiducial. This experiment corresponds to the real world application where a UAV has to take off from a landing site with a visual fiducial on it, fly away from it, and then fly even further where the range measurements are not available anymore. Finally, the UAV comes back to the site, acquiring measurements from all the sensors again.

The middle part of Fig. 9 shows the position error of the quadrotor when the system was using the Chi-Squared test. The general tendency of the error plot is exactly as expected - the more sensors were used, the more accurate was the state estimation of the vehicle. It is apparent, that the sensor suite switching happened seamlessly throughout the experiment. One can also observe initial peaks around the 13th and 19th seconds of the execution time which correspond to the moments when the visual fiducial was temporarily not visible in the camera image. The last phase of the flight, when all the sensors were used again, shows a very small and constant position error because the visual fiducial was always in the field of view of the camera.

In order to evaluate the effect of the Chi-Squared test in this scenario, we executed the filter on the collected data again without the test. The result of this run is depicted in Fig. 9-bottom. The position error in this experiment is much larger and contains spikes which would be fatal for a controller. The RMSE of the system with the Chi-Squared test is equal to 0.129m whereas without the test it is 0.193m. The spikes in the state estimates come mainly from the erroneous range measurements (for example, the large spike

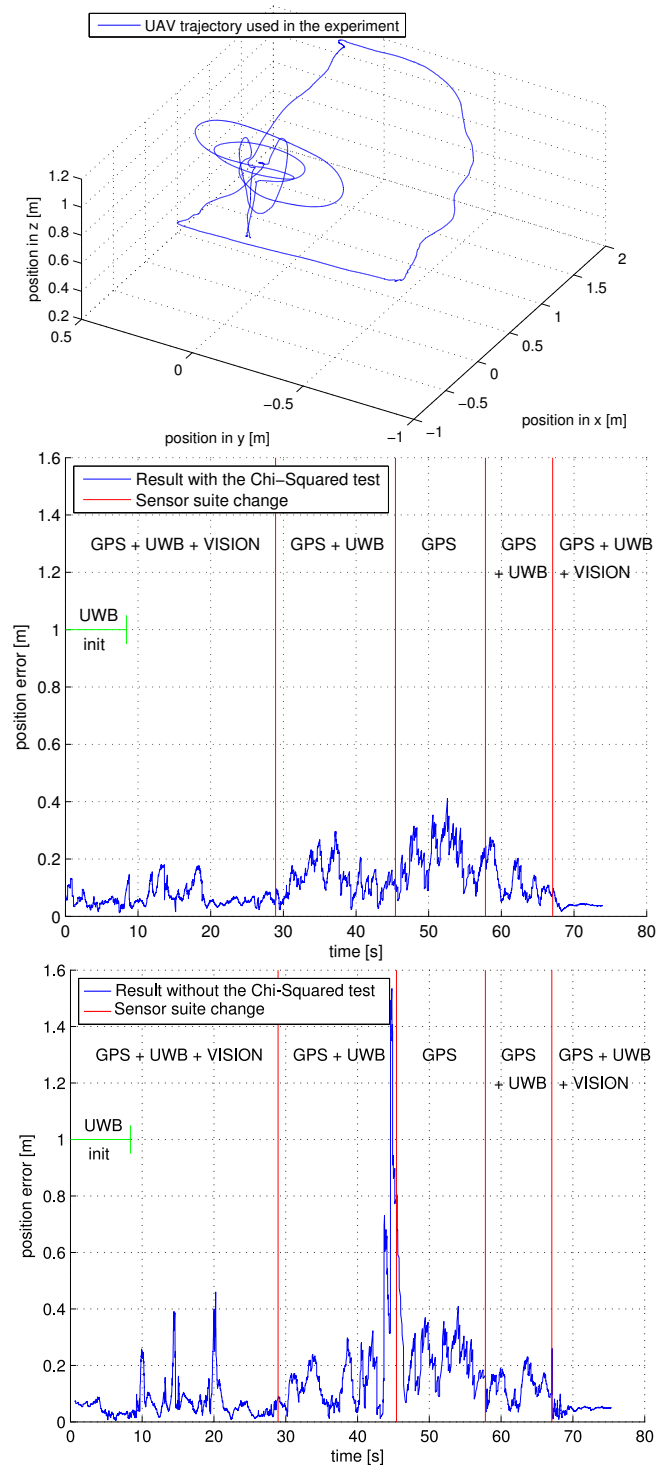


Fig. 9. Full system evaluation. Top: the trajectory executed by the UAV. Middle: the position error obtained by the system with the Chi-Squared test. Bottom: the position error obtained by the system without the Chi-Squared test.

at 45s) but also from the vision sensor when the UAV received a faulty measurement by seeing the fiducial from a large distance (at 10, 15 and 20s). It is worth noting that in the case when the Chi-Squared test was included, these visual measurements were not applied, hence the error increased but there were no spikes (see Fig. 9-middle). This experiment confirms our hypothesis that the Chi-Squared test is crucial

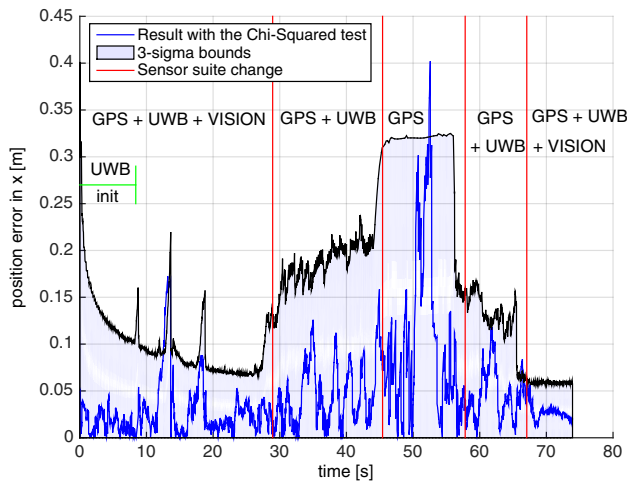


Fig. 10. The position error in the x dimension together with its uncertainty calculated based on the state covariance. The data comes from the same experiment as in Fig. 9.

for correct fusion of multiple sensors in the presence of erroneous measurements.

Fig. 10 shows the position error in the x dimension together with its uncertainty for the run with the Chi-Squared test. The very initial drop of the uncertainty shows the convergence rate of the filter. One can notice that the 3-sigma bound of the state generally reflects the error of the estimate. The uncertainty of the estimate is smoother when using the GPS only due to the fact that this is the only measurement with a constant covariance in our setup. All the other measurements arrive with varying covariance that is either specified by the manufacturer (the UWB range modules) or estimated through a set of experiments (the visual fiducial).

VII. CONCLUSIONS AND FUTURE WORK

We introduced a self-initializing and self-calibrating multi-sensor fusion approach that is able to probabilistically switch between sensors and uses a statistical measurement validation approach to not incorporate erroneous measurements in the state. In particular, we showed in a series of experiments that our system is able to fuse vision, GPS and range measurements with no need for external calibration or initialization. We evaluated different parts of the system such as the statistical measurement validation and the range module initialization separately to explain their influence on the system behavior. Finally, we performed a full system evaluation on a real UAV which showed the ability to seamlessly filter and switch between different sensors modalities during run time.

As next steps, we plan to further enhance our system for long-term UAV autonomy in outdoor missions to show its application in larger environments.

ACKNOWLEDGMENTS

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

- [1] BD Brumback and Mandyam D Srinath. A chi-square test for fault-detection in kalman filters. *Automatic Control, IEEE Transactions on*, 32(6):552–554, 1987.
- [2] Jason N Gross, Yuntao Gu, and Matthew B Rhudy. Robust uav relative navigation with dgps, ins, and peer-to-peer radio ranging. In *2014 Institute of Navigation GNSS+*. IEEE.
- [3] Jeroen Hol, Fred Dijkstra, Henk Luinge, and Thomas Schön. Tightly coupled uwb/imu pose estimation. In *Proceedings of IEEE International Conference on Ultra-Wideband*, page 688692. IEEE, 2009.
- [4] Eagle S Jones and Stefano Soatto. Visual-inertial navigation, mapping and localization: A scalable real-time causal approach. *The International Journal of Robotics Research*, 30(4):407–430, 2011.
- [5] Jonathan Kelly and Gaurav S Sukhatme. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. *The International Journal of Robotics Research*, 30(1):56–79, 2011.
- [6] Stefan Leutenegger and Roland Y Siegwart. A low-cost and fail-safe inertial navigation system for airplanes. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 612–618. IEEE, 2012.
- [7] Simon Lynen, Markus W Achtelik, Steven Weiss, Maria Chli, and Roland Siegwart. A robust and modular multi-sensor fusion approach applied to mav navigation. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 3923–3929. IEEE, 2013.
- [8] D.C. Montgomery, E.A. Peck, and G.G. Vining. *Introduction to Linear Regression Analysis*. Wiley Series in Probability and Statistics. Wiley, 2015. ISBN 9781119180173. URL <https://books.google.com/books?id=27kOCgAAQBAJ>.
- [9] Anastasios Mourikis, Stergios Roumeliotis, et al. A multi-state constraint kalman filter for vision-aided inertial navigation. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 3565–3572. IEEE, 2007.
- [10] Elias Mueggler, Ben Huber, and Davide Scaramuzza. Event-based, 6-dof pose tracking for high-speed maneuvers. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2761–2768. IEEE, 2014.
- [11] Mark W Mueller, Michael Hamer, and Raffaello D’Andrea. Fusing ultra-wideband range measurements with accelerometers and rate gyroscopes for quadcopter state estimation. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 1730–1736. IEEE, 2015.
- [12] Shaojie Shen, Nathan Michael, and Vijay Kumar. Autonomous multi-floor indoor navigation with a computationally constrained mav. In *Robotics and automation (ICRA), 2011 IEEE international conference on*, pages 20–25. IEEE, 2011.
- [13] Shaojie Shen, Yash Mulgaonkar, Nathan Michael, and Vijay Kumar. Vision-based state estimation and trajectory control towards high-speed flight with a quadrotor. In *Robotics: Science and Systems*. Citeseer, 2013.
- [14] Stephan Weiss, Markus W Achtelik, Simon Lynen, Margarita Chli, and Roland Siegwart. Real-time onboard visual-inertial state estimation and self-calibration of mavs in unknown environments. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 957–964. IEEE, 2012.
- [15] Stephan Weiss, Markus W. Achtelik, Simon Lynen, Michael C. Achtelik, Laurent Kneip, Margarita Chli, and Roland Siegwart. Monocular vision for long-term micro aerial vehicle state estimation: A compendium. *Journal of Field Robotics*, 30(5):803–831, 2013.