High-frequency MAV state estimation using low-cost inertial and optical flow measurement units

Angel Santamaria-Navarro, Joan Solà and Juan Andrade-Cetto

Abstract—This paper develops a simple and low-cost method for 3D, high-rate vehicle state estimation, specially designed for free-flying Micro Aerial Vehicles (MAVs). We fuse observations from inertial measurement units and the recently appeared low-cost optical flow smart cameras. These smart cameras integrate a sonar altimeter, a triaxial gyrometer and an optical flow sensor, and directly provide metric ego-motion information in the form of body velocities and altitude. Compared to state-of-the-art visual-inertial odometry methods, we are able to drastically reduce the computational load in the main processor unit, and obtain an accurate estimation of the vehicle state at a high update rate of 100Hz. We thus extend the current use of these smart cameras from hovering purposes to odometry estimation. In order to propose a simple algorithmic solution, we investigate the performances of two Kalman filters, in the extended and error-state flavors, alongside a large number of algorithm variations, using simulations and real experiments with precise ground-truth. We observe that the marginal performance gain attained with these algorithm improvements does not pay for the effort of implementing them. We conclude that a classical EKF in its simplest form is sufficient for providing motion estimates that coherently exploit the available measurements.

I. INTRODUCTION

Micro Aerial Vehicles (MAVs), and in particular quadrotor systems, have substantially gained popularity in the research community in recent years, motivated by their significant increase in maneuverability, together with a decrease in weight and cost. In order to perform high level tasks with MAVs, such as autonomous exploration or scene interaction, robust state estimation at a high update rate is crucial for robot navigation and agile flight control.

Low-cost, low complexity systems for rich and high-dynamics MAV state estimation are rare. A first family of localization systems for MAVs relies on external infrastructure (GPS, RF beacons, visual tags, IR cameras), e.g. [1], which may not always be practical. When using low-cost setups, these systems usually lack precision, dynamics, or both. A second family of methods embark all hardware and software for self-localization, thus not relying on any external setup. These include any kind of SLAM or odometry system drawn from other robotics fields [2], [3].

This work has been partially funded by the EU project ARCAS FP7-287617, and by the Spanish Ministry of Economy and Competiveness projects PAU+ DPI-2011-27510 and ROBINSTRUCT TIN2014-58178-R.

The authors are with the Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens i Artigas 4-6, Barcelona 08028, Spain, asantamaria, jsola, cetto@iri.upc.edu.

The paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors. The material includes a video of the state estimation presented in the paper.
their output is the result of dead-reckoning, subject to drift. These modes can be observed by a higher level task, such as a visual servioing ([14], [15]), with a lower update rate.

To the knowledge of the authors, this work is the first example of the usage of such a low-cost visual-range-inertial sensor setup for 6DoF motion estimation, using also low complexity algorithms without sophisticated features. This sensor setup has the advantage of being simple, lightweight, low power and low cost, and is already included, with minor variations, in several commercial multi-copters as their basic instrumentation, being typically used as a means for automatic hovering. As such, a setup like this is already available to a huge amount of teams worldwide, both in research and industry. Many of these teams may not have the energy, the time or the interest of implementing complex, CPU-demanding motion estimation algorithms, yet they are not satisfied with a pure hovering solution.

When compared to other visual-intertial approaches, the optical-flow nature of the PX4, with a very narrow field of view of only $64 \times 64$ pixels or $1.6^\circ$, represents an important limitation, in the sense that visual features are only exploited locally, and there is no notion of features being tracked over long periods of time. The number and length of feature tracks are key to the high precision attained by the aforementioned sophisticated methods, and in consequence, we cannot expect performances of our setup comparable to other visual-odometry systems. But importantly, the number and length of these feature tracks are responsible for their algorithmic complexity and CPU consumption.

Fortunately for us, the high filter update rate made possible by the PX4 contributes to increase the validity of linearizations, minimizing the accumulation of error and thus enabling simpler methods for an equivalent performance. In this scenario, the main aim of this paper is to show that, given a sensor setup with such capabilities, we are able to derive motion estimates that are useful in the mid term (a few minutes, i.e., from a significant fraction to a full battery recharge), without the need to implement complicated algorithms.

Indeed, we explore a large number of algorithm improvements and variations defended in the literature, and show that their impact on system performance is marginal. These variations are shown in Tab. I, and are properly defined later in the text. They are summarized hereafter. First, we implement extended (EKF) and error-state (ESKF) Kalman filters [16]. Second, we express the orientation errors of ESKF both in local (LE) and global (GE) frames [17]. Note that in EKF the orientation error is additive and this distinction is irrelevant. Third, we compare different schemes for rotational motion integration of the quaternion [18], including forward zeroth-order (QOF), backward zeroth-order (QOB), and first-order (Q1). Fourth, we compare different integration forms and approximations of the system’s transition matrix ($F_1, F_2, F_3$) [5].

Other contributions of the paper follow. First, the observation model of the PX4 sensor, and its analytical derivation to get, together with the IMU, all the observable modes.

### TABLE I

<table>
<thead>
<tr>
<th>Filter type</th>
<th>Quat. error</th>
<th>Quat. integration</th>
<th>Trans. Mat. Trunc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EKF</td>
<td>–</td>
<td>QOF, QOB, Q1</td>
<td>$F_2$</td>
</tr>
<tr>
<td>ESKF</td>
<td>GE, LE</td>
<td>QOF, QOB, Q1</td>
<td>$F_1, F_2, F_3$</td>
</tr>
</tbody>
</table>

Also, we propose algorithmic solutions to correct some limitations of the PX4 that make it produce, under particular circumstances, aberrant measurements.

The remainder of this article is structured as follows. In the next section we provide a brief description of the system setup, with its sensor models and state parametrizations. The system kinematic equations are described in Sec. III, together with a brief note on observability. Sec. IV shows the EKF and ESKF filters. Simulation validations come in V, and real robot experiment results in Sec. VI. Finally, conclusions are given in Sec. VII.

## II. SYSTEM SETUP

We consider a quadrotor equipped with a MicroStrain 3dm-gx3-25 IMU and a PX4 optical flow sensor, moving with respect to a global coordinate frame, assumed inertial. Our goal is to track the vehicle state, i.e., the position, orientation and velocities of the platform, using the inertial measurements, and to correct these estimates with the PX4 readings, observing in turn the IMU biases for their proper compensation. The PX4 integrates a monocular camera looking downwards, an ultrasound range sensor in the same direction, and a triaxial gyrometer. It processes these data and delivers horizontal velocities $v_{of}(x,y)$ in local (i.e., body) frame, plus altitude $p_{of}(z)$ with respect to a flat, horizontal ground. The PX4 can also provide the raw optical flow data, which can be used in critical situations to recover from aberrant PX4 readings, as we explain in section VI-A. In order for the PX4 to perform correctly, scene lighting must be adequate, and the ground is required to have good visual texture and sufficient ultrasound reflectivity.

The IMU and the PX4 optical flow sensors are rigidly attached together as shown in Fig. 1. Their respective frames are aligned with each other and with the frame of the MAV via careful mechanics, without further calibration.

### A. State vectors: notation and conventions

We are interested in estimating a true state $x_t$ defined by

$$x_t = [p_t, v_t, q_t, a_{bt}^t, \omega_{bt}^t]^\top,$$

where $\{p, v, q\}$ are position, velocity, and orientation quaternion$^1$, all expressed in inertial frame, and $\{a_{bt}^t, \omega_{bt}^t\}$ accelerometer and gyrometer biases. In EKF, we estimate $x_t$ directly. In ESKF, we speak of true-, nominal- and error-state values, the true-state $x_t$ being expressed as a

$^1$We use the Hamilton convention for quaternions, which defines $ij = -ji = k$. This is the most widely used convention (e.g. Eigen, Ceres, ROS, and most literature), but it is in contrast with quite a few remarkable works on visual-inertial odometry (e.g. [5], [6], [18]) using the JPL convention, with $ji = -ij = k$. Because of the sign change, the Hamilton quaternion is right-handed, while the JPL is left-handed [18].
suitable composition (linear sum or quaternion product) of the nominal-state \( x \), which is only propagated in time, and the error-state \( \delta x \), which is estimated by the filter. They are related as follows,

\[
\begin{align*}
    x_t &= x \oplus \delta x \\
    x &= [p v q a_b \omega_b]^	op \\
    \delta x &= [\delta p \delta v \delta \theta \delta a_b \delta \omega_b]^	op .
\end{align*}
\]

Finally, we note \( R\{x\} \) and \( q\{x\} \) the rotation matrix and quaternion equivalents to the generic orientation \( x \). Likewise, we write \( R \triangleq R\{q\} \) and \( R_t \triangleq R\{q_t\} \).

### B. Sensor Models

The IMU is composed of a 3-axis accelerometer and a 3-axis gyrometer. The accelerometer measures acceleration \( a \) and gravity \( g \) together, expressed in body frame. This measurement is affected by additive bias \( a_{bd} \) and noise \( a_n \). The gyrometer measures angular rates in body frame, likewise affected by additive bias and noise. The sensed values are thus,

\[
\begin{align*}
    a_S &= R_t\{q_t\}(a - g) + a_{bd} + a_n \quad (7a) \\
    \omega_S &= \omega_t + \omega_{bd} + \omega_n . \quad (7b)
\end{align*}
\]

The PX4 optical flow sensor outputs the altitude from the ground, and a 2D linear velocity vector in the body XY plane, all affected by additive Gaussian noise. This can be expressed in compact form as

\[
\begin{bmatrix}
p_{of}(z) \\
v_{of}(x) \\
v_{of}(y)
\end{bmatrix} = S \begin{bmatrix}
p_t \\
R_t\{q_t\}v_t
\end{bmatrix} + n_{of} ,
\]

where \( S = [0_{3\times2} \ I_{3\times3} \ 0_{3\times1}] \) is a 3×6 matrix that selects the respective rows. The measurement noise \( n_{of} \) is assumed to be Gaussian with zero mean and covariances matrix \( N_{of} \). The planar velocity measurement \( v_{of} \) is computed directly by the PX4 from its own optical flow \( \Phi_{of} \), angular rates \( \omega_{of} \), and altitude \( h_{of} \), measurements according to

\[
v_{of} = -\frac{\Phi_{of} \cdot h_{of}}{\Delta t} + S' \left( \omega_{of} \times [0 \ 0 \ h_{of}] \right)^	op ,
\]

where \( f \) is the focal length of the PX4’s camera, expressed in pixels, and \( S' = [I_{2\times2} \ 0_{2\times1}] \) is also a selection matrix.

### III. System Kinematics for EKF and ESKF

#### A. Continuous time

As it is common practice in the literature of IMU navigation (e.g. [4]), we express the true dynamics of our system with the following set of equations,

\[
\begin{align*}
    \dot{p}_t &= v_t \quad (10a) \\
    \dot{v}_t &= R_s(a_S - a_{bd} - a_n) + g \quad (10b) \\
    q_t &= \frac{1}{2} q_t \times [0 \ \omega_S - \omega_{bd} - \omega_n]^	op \quad (10c) \\
    \dot{a}_{bd} &= a_w \quad (10d) \\
    \dot{\omega}_{bd} &= \omega_w , \quad (10e)
\end{align*}
\]

where true acceleration \( a_t \) and angular rates \( \omega_t \) have been substituted by those isolated from the IMU model (7). The biases have been considered non-constant, and in such case two small-energy random walks with white Gaussian generator noises \( a_w \) and \( \omega_w \) are used as dynamic models.

The system above, after integration to discrete time (see below), is estimated by an EKF. For the ESKF, we isolate the nominal-state kinematics by simply removing all error, noise and perturbation terms. The error-state kinematics are obtained by considering the composition of each system equation, solving for the error-state, rearranging some cross-products and simplifying all second order terms, yielding,

\[
\begin{align*}
    \dot{\delta p} &= \delta v \\
    \dot{\delta v} &= -[R(a_S - a_b)] \times \delta \theta - R \delta a_b - Ra_n \quad (11b) \\
    \dot{\delta \theta} &= -R \delta \omega_b - R \omega_n \quad (11c) \\
    \dot{\delta a}_b &= a_w \quad (11d) \\
    \dot{\delta \omega}_b &= \omega_w , \quad (11e)
\end{align*}
\]

which, remarkably, is a time-variant linear system. This system is valid for the globally-defined angular error, (GE, 5a). For a local definition, (LE, 5b), we just need to substitute equations (11b,11c) with,

\[
\begin{align*}
    \dot{\delta v} &= -R[a_S - a_b] \times \delta \theta - R \delta a_b - Ra_n \quad (11f) \\
    \dot{\delta \theta} &= -[\omega_S - \omega_b] \times \delta \theta - \delta \omega_b - \omega_n . \quad (11g)
\end{align*}
\]

#### B. Computing the integral to discrete time

The integral of continuous differential equations of the type \( \dot{x} = f(x,u) \) from times \((k-1)\Delta t \) to \( k\Delta t \) can be approximated in a number of ways. We want to evaluate the impact of these approximations on final performance, and thus we perform these integrals with different degrees of accuracy. For all the variables except the quaternion, we integrate the linearized system, \( \dot{x} = Ax + Bu \), with \( A = \partial f/\partial x \), \( B = \partial f/\partial u \), into the discrete-time \( x_k \approx Fx_{k-1} + Bu \Delta t \), and we truncate the Taylor series expressing the transition matrix \( F = e^{A \Delta t} \) at different orders, \( N \in \{1, 2, 3\} \), obtaining the different approximations \( F_N \) in Tab. I,

\[
F_N \triangleq \sum_{n=0}^{N} \frac{1}{n!} A^n \Delta t^n .
\]

For the quaternion, we contemplate the Q0F, Q0B, and Q1 integrators [18] of the differential equation (10c),

\[
\begin{align*}
\text{Q0F: } q_k &= q_{k-1} \otimes q(\omega_{k-1} \Delta t) \\
\text{Q0B: } q_k &= q_{k-1} \otimes q(\omega_k \Delta t) \\
\text{Q1: } q_k &= q_{k-1} \otimes q(\bar{\omega} \Delta t) + \frac{\Delta t^2}{24} \left[ 0 \right]
\end{align*}
\]

where \( \omega_k = \omega_{S,k} - \omega_{bt,k} \) and \( \omega_{bt,k} \approx \omega_{bt,k-1} \) and \( \bar{\omega} \triangleq (\omega_{k-1} + \omega_k)/2 \).

C. Discrete time

For convenience, we start with the integration of the error state. For space reasons, we show here only the case with first-order Taylor truncation of the transition matrix (F1), zeroth-order backward quaternion integration (Q0B), and global angular error definition (GE), to integrate the error-state equations (11),

\[
\begin{align*}
\delta p &\leftarrow \delta p + \Delta t \delta v \\
\delta v &\leftarrow \delta v - [R(aS - a_b) \times \Delta t \delta \theta - R \Delta t \delta a_b + v_i] \\
\delta \theta &\leftarrow \delta \theta - R \Delta t \delta \omega_b + \theta_i \\
\delta a_b &\leftarrow \delta a_b + a_i \\
\delta \omega_b &\leftarrow \delta \omega_b + \omega_i,
\end{align*}
\]

where “\( \leftarrow \)” stands for “gets updated with”, i.e., \( x \leftarrow f(x, \cdot) \) is equivalent to \( x_k = f(x_{k-1}, \cdot) \). Here, \( v_i, \theta_i, a_i, \) and \( \omega_i \) are random impulses applied to the velocity, orientation and bias estimates, modeled with Gaussian processes. Their mean is zero, and their covariance matrices are obtained by integrating the variances of the IMU measurement noises, \( a_n, \omega_n, \) and the IMU bias random walks, \( a_w, \omega_w, \) over the time step \( \Delta t \).

\[
\begin{align*}
V_i &= \sigma_{a_n}^2 \Delta t^2 I_3 \left[ m^2/s^2 \right] \\
\Theta_i &= \sigma_{\omega_n}^2 \Delta t^2 I_3 \left[ rad^2 \right] \\
A_i &= \sigma_{a_w}^2 \Delta t I_3 \left[ m/s^2 \right] \\
\Omega_i &= \sigma_{\omega_w}^2 \Delta t I_3 \left[ rad^2/s^2 \right],
\end{align*}
\]

D. Observability, and marginally-observable modes

The observability analysis of the system needs the evaluation of the rank and continuous symmetries of the observability matrix defined from the Lie derivatives [19]. Following this work, we detect three continuous symmetries,

\[
\begin{align*}
\omega_x &= [1, 0, 0, 0, \ldots, 0] \\
\omega_y &= [0, 1, 0, 0, \ldots, 0] \\
\omega_z &= [-p_x, p_y, 0, -v_y, v_x, 0, -q_x, q_y, q_z, q_w, 0, \ldots]
\end{align*}
\]

IV. EXTENDED AND ERROR-STATE KALMAN FILTERS

A. Extended Kalman filter

The development of an EKF from equations (16) and (8) is standard procedure and we do not include it here for space reasons. The transition matrix \( F_2 \) is the Jacobian of (16).

B. Error-state Kalman filter

a) Nominal-state propagation: The nominal state is propagated according to (16) with all perturbation impulses \{\( v_i, \theta_i, a_i, \omega_i \)\} set to zero.

b) Error-state KF prediction: The error-state system (14) can be posed as

\[
\delta x \leftarrow F_N \delta x + 1
\]

where \( i \sim \mathcal{N}(0, Q) \) is a vector of Gaussian random impulses with zero mean and covariances matrix \( Q \).

The ESKF time-update is applied only to the covariance,

\[
P \leftarrow F_N P F_N^T + Q
\]

which may accept some variations depending on the algorithm options, especially GE vs. LE.
c) Error-state KF correction: Our PX4 sensor output depends on the system state as shown in (8). This is written in function form as,

\[ v_{of} = h(x_t) + n_{of}. \]  

Based on this, the ESKF update is written,

\[
\begin{align*}
K &= PH^\top (HPH^\top + N_{of})^{-1} \\
\delta_x &= K(v_{of} - h(x)) \\
P &\leftarrow P - K(HPH^\top + N_{of})K^\top \\
x &\leftarrow x + \delta_x,
\end{align*}
\]

where \( H = \partial h(x_t)/\partial x \) is the observation matrix, derived from (8) and (3–5a) for GE (a slightly different expression is obtained for LE using with (5b)),

\[
H = S \begin{bmatrix}
I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 6} \\
0_{3 \times 3} & R_1^\top & R_1^\top & 0_{3 \times 6}
\end{bmatrix},
\]

and \( \odot \) is the same composition operator explained in Sec. II-A, paying special attention to the options in (5).

V. VALIDATION WITH SIMULATIONS

In order to study the performances and limitations of the proposed setup, we first present experiments with synthetic data under realistic flight conditions. Quadrotors are typically equipped with four aligned coplanar propellers. Motion control is achieved by altering the rotation speed of these propellers, thereby changing its torque load and thrust lift characteristics. With this actuation technique, a quadrotor becomes an under-actuated vehicle with only 4 DOF. Translational accelerations are achieved by tilting the rotor plane in the desired direction. We generate a ground truth trajectory that accounts for these motion constraints. We subsequently generate corrupted PX4 and IMU measurements, with noise and bias characteristics similar to those of the real units (see next section for noise values). Both sensors deliver data synchronously at 100Hz.

Fig. 2 shows the results with simulated data using an ESKF with GE, Q0B and F1. The 60-s trajectory encompasses an initial phase of hovering (phase A); a forward movement along the \( x \) axis (B); a backward movement (C) to the initial position; a second hovering phase (A); and finally left (D) and then right (E) movements along the \( y \) axis. The true trajectory finishes precisely at the starting position. X-Y-Z magnitudes are colored R-G-B, and \( \pm 3\sigma \) estimated error bounds are also plotted.

The \( xy \) position is not observable, and its estimated error increases with time (top frame in Fig. 2). However, the drift is small, a few centimeters after the whole one-minute flight. Altitude and velocity estimates converge quickly thanks to the PX4 measurements. Notice the transient increase of the position and velocity errors (phases B and D). This is due to the uncertainty in \( yaw \) (blue track of third plot in Fig 2, not observable) which produces position errors perpendicular to the displacement vector. These errors decrease as the MAV returns to the origin (C and E).

One of the most interesting aspects is the marginal observability of some of the biases. On one hand, an acceleration bias becomes observable when the quadrotor orientation produces a non-null projection of the gravity vector over its axis. This makes the accelerometer’s \( z \)-bias readily observable, but \( x \)- and \( y \)-biases require tilting the rotor plane. This is observable in Fig. 2, plot ‘Acc. Bias’. This is a critical issue due to the quadrotor’s restricted dynamics, seeking always horizontality and thus making these observations difficult. To further investigate this effect, the impact of a rotation over \( y \) (resp. \( x \)) on the observability of \( a_y(x) \) (resp. \( a_y(y) \)) is illustrated in Fig. 3. A quadrotor can easily tilt to 20°, but holding this tilt for a period of seconds represents a sustained high acceleration that is usually not desirable. Otherwise, this bias should be estimated beforehand and set up in the filter.

On the other hand, The gyroscope’s \( pitch \) and \( roll \) biases are well observable thanks to the gravity vector. The \( yaw \) bias \( \omega_y(z) \) (blue in ‘Ang. Vel’ plot) becomes observable only after a very long period, by taking small profit of transient inclinations that bring the gravity vector away from the local \( z \) axis. To achieve a better performance in \( yaw \)-bias, it is advisable to use extra heading measurements such as a compass or other stable external references.
TABLE II
KF TUNING PARAMETERS (STD; σ SYMBOLS OMITTED)

<table>
<thead>
<tr>
<th>Sensor noises</th>
<th>Initial std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_n$ m/s²</td>
<td>0.4</td>
</tr>
<tr>
<td>$\omega_n$ rad/s</td>
<td>0.005</td>
</tr>
<tr>
<td>$z_{of}$ m</td>
<td>0.05</td>
</tr>
<tr>
<td>$v_{of}$ m/s</td>
<td>0.1</td>
</tr>
<tr>
<td>$\delta z$ m</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta \phi$, $\delta \theta$ rad</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta a_b$ m/s²</td>
<td>0.02</td>
</tr>
<tr>
<td>$\delta \omega_b$ rad/s</td>
<td>0.004</td>
</tr>
</tbody>
</table>

VI. REAL MAV EXPERIMENTS

We present experiments with an ASCTEC PELICAN quadrotor. The IMU and the optical flow sensors deliver data asynchronously at 100Hz, and we associate each PX4 measurement with the closest IMU measurement. The sensors are attached below the platform (Fig. 1) using silicon damping links to reduce motor vibrations.

Tab. II shows all non-null sigma values for the filter tuning. All filter variants use the same parameters. For the IMU noises, we followed the datasheet to obtain $\sigma_{a_n} = 5.3 mm/s^2$ and $\sigma_{\omega_n} = 3.6 m rad/s$. However, empirical IMU data has revealed that the mechanical damping is not sufficient, and propeller vibrations inject much higher noises to the acceleration measurements. For this reason, the value of $\sigma_{a_n}$ is significantly increased as shown in the table. We assume null bias random walks, as our flights are short enough. Initial uncertainties are mostly null, except for altitude $\sigma_{z_{of}}$, initial roll $\sigma_{\delta \phi}$ and pitch $\sigma_{\delta \theta}$, accelerometer bias $\sigma_{\delta a_b}$, and gyrometer roll and pitch bias $\sigma_{\delta \omega_b}$.

A. Improving PX4 usability

The PX4 sensor has two key limitations that need to be addressed algorithmically in order to improve its robustness and usability. The first one is its inability to measure altitudes below 30 cm. For this, we divide each experiment in 4 phases depending on the robot state: landed (A), taking-off (B), flying (C) or landing (D). Thus during A, B and D intervals, the PX4 output is not reliable. We address this problem by reading status data of the robot to acknowledge whether it is in flying mode or not (e.g., motors on/off). Before takeoff (phase A), we overwrite the optical flow measurements by assuming that the MAV is on zero position, with zero velocity, and we set a small observation covariance. Thus, the bias uncertainties on the $z$ acceleration and $x$ and $y$ angular velocities are reduced. If the robot is flying below the minimum altitude of 0.3m (phases B and D), detected with PX4 readings, we set high PX4 covariances because the measures are not trustable. Hence the filter is propagating the nominal state with the IMU dynamical model, and practically does not correct with the PX4 sensor. During regular flight (phase C) the observation covariances are set to those in Tab. II, thus allowing the PX4 to correct the estimations properly.

A last feature of our algorithm deals with the second PX4 data integrity issue. In some occasions, tilts over 20 degrees produce wrong sonar echoes, deriving in aberrant PX4 outputs. We detect altitude outliers using the modified Thompson Tau method [20]. Once an outlier is detected, the filter correction can be accomplished by using the optical flow measurement as a function of the IMU’s gyrometer measurement $\omega_S$ and the state variables,

$$
\Phi_{of} = S'(\omega_S \times [0 \ 0 \ 1]^{\top} - R_{\{q\}^{\top}v} \ p(z)) \Delta t \cdot f + n_{\Phi}
$$

with $n_{\Phi}$ a one-pixel Gaussian noise. For simplicity, we however re-used the observation model (8) by computing the velocity $v_{of}$ with (9) substituting $h_{of} \leftarrow p(z)$ and $\omega_{of} \leftarrow \omega_S$. In this case, to compensate for the correlation between the state and the measurement, we increased the sensor covariances $N_{of}$ slightly. Fig. 4 shows the reconstructed $v_{of}$ in front of typical outliers.

B. Benchmarking all filter variants against ground truth

We present the results in an indoor testbed equipped with an eight-camera Optitrack motion capture system [21], with a flying area of some $3 \times 5 m$. Our aim is at comparing the outcome of our different filter variants against ground truth. Fig. 5(a) shows the overhead plot of a 100m sample trajectory (using an ESKF with GE, Q1, and $F_2$). Inspired by the methodologies in [7], [22], [23], we evaluate the odometry RMS errors as a function of the trajectory length, over 9 different trajectories. For each segment, the starting ground truth and estimation poses are aligned, and the error is evaluated at the end of the segment. Fig. 5(b) shows the negligible difference existing between a simple EKF and the aforementioned ESKF. Considering our flying zone, the statistic evaluation of these errors is only relevant for segment lengths up to around 10m (less than one turn to the arena), because after that the trajectory is rolled over itself and the nonlinear composition of angular and linear errors produces optimistic error evaluations. Below these 10m, odometry errors fit with the expected law of the square root of the length $d$ of the evaluated segments, as shown in Fig. 6(a). This good fit indicates that the yaw error contribution to position error is small. The yaw error, Fig. 6(b), does not suffer from this nonlinear composition and fits the $\sqrt{d}$ better. At an average of $0.2^\circ$ over the 10m, the induced linear error
This paper presents a simple method for rich and high-frequency vehicle state estimation fusing low-cost inertial and smart optical flow measurement units for a MAV, using low complexity algorithms without sophisticated features.

C. Outdoor arena experiments

The outdoor scenario corresponds to a GPS-denied zone in which ground truth from external positioning systems is not available. Instead, we perform manual flights around some fixed obstacles, taking care to take-off and land on the same base point.

Fig. 7 shows the calibrated outdoor flying arena for real robot testings. The quadrotor takes off and lands at the same point in the middle of the field (base).

Fig. 8. Trajectory estimations for 15 different flights (2 minutes, 70 m each). A sample trajectory is shown in blue. The orange ellipse corresponds to the 95% confidence region for the landing point.

TABLE III

<table>
<thead>
<tr>
<th>Seg. int.</th>
<th>GE</th>
<th>LE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q0F</td>
<td>Q0B</td>
<td>Q1</td>
</tr>
<tr>
<td>Q1</td>
<td>Q0F</td>
<td>Q0B</td>
</tr>
<tr>
<td>EKF F_2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EKF F_1</td>
<td>0.1986</td>
<td>0.2007</td>
</tr>
<tr>
<td>ESKF F_2</td>
<td>0.1987</td>
<td>0.2007</td>
</tr>
<tr>
<td>ESKF F_3</td>
<td>0.1987</td>
<td>0.2007</td>
</tr>
</tbody>
</table>

would be of the order of only 10m $\sim (0.2^\circ) = 3.5$ cm, a small fraction of the 30cm total position drift.

To evaluate all 21 possible filter variants (shown previously in Tab. I), we identified 100 straight segments of 2m long over 9 paths. The position RMSEs at the end of the segments are shown in Tab. III. Notice that the performance of all filter variants is essentially equivalent.

VII. Conclusions

This paper presents a simple method for rich and high-frequency vehicle state estimation fusing low-cost inertial and smart optical flow measurement units for a MAV, using low complexity algorithms without sophisticated features.
A large amount of variants for nonlinear KF have been benchmarked, revealing that, for such a basic setup, with high input data rates, the plain EKF filter performs equivalently to other, more sophisticated options.

A new observation model is derived for this sensor combination, and the observability limitations of the system are pointed out. Namely, it is difficult to recover the IMU’s \( xy \) accelerometer and \( yaw \) gyrometer biases because the quadrotor dynamics prevent from severe orientation changes.

Methods have also been proposed to overcome situations of poor PX4 data integrity, mainly due to low altitude or high tilt angles where the sonar data fails. These improvements have revealed crucial for improving the robustness of such a low-cost setup to acceptable levels.

We can think of three avenues for further research. First, a calibration warm-up period can be specially designed for quadrotors to compensate for the marginal observability. Second, some other validation experiments can be done using an IMU with lower cost than the Microstrain 3dm-gx3-25. Finally, the non-observable modes can be made observable by adding other sensor readings; for example, compass readings can easily avoid the \( yaw \) drift.

VIII. ACKNOWLEDGEMENTS

The authors would like to thank the unmanned systems unit from Eurecat Technology Center for giving us access to their indoor flying arena and facilities.

REFERENCES


