CILAMCE 2017
Florianópolis - Santa Catarina - Brazil

COMPARATIVE EVALUATION OF FILTERING METHODS IN AN EXPERIMENTAL VISUAL-INERTIAL ATTITUDE DETERMINATION SETUP

Beatriz A. Asfora

Éder A. Moura

Davi A. Santos

beatriz.asfora@gmail.com

edermoura@gmail.com

davists@ita.br

Instituto Tecnológico de Aeronáutica

Praça Marechal Eduardo Gomes, 50. 12228-900 SP, São José dos Campos, Brasil

Abstract. *The small size and fast dynamics of Micro-Multirotor Aerial Vehicles impose challenges regarding their navigation and control. Visual-Inertial schemes are a good fit for tackling the navigation issue, and although Monte Carlo Simulations provide valuable information on sensor fusion methods, it is fundamental to have a controlled real-world comparative basis for practical implementation of the algorithms. This work is concerned with the experimental comparison of the Extended Kalman Filter, Unscented Kalman Filter and Ensemble Kalman Filter. In the setup presented, they each rely on the measurements of a three-axis rate-gyro and an upward-facing camera to estimate the platform attitude at each instant based on the known positions of landmarks. The platform is a Quanser 3-DOF Hover kit, with high accuracy encoders that provide an alternative measure of attitude, taken as the truth state. The performance of each filter is evaluated in a similar manner as it would be in computational simulations: they can be tested in a standardized but real environment, providing a complementary analysis, so there can be a mindful choice of the more suitable approach for each goal.*

Keywords: *Multirotor aerial vehicle, attitude determination, visual-inertial navigation, experimental analysis*

1 INTRODUCTION

Pilot-less or unmanned aerial vehicles (UAVs) became recognized systems during military projects in World War I, but it was not until the 1990s that this area experimented its research boom, with the advances in micro-mechanical systems and microelectronic components, as story is told by Fahlstrom and Gleason (2012) and Mueller (2009). The use of such aircrafts is widely seen in the Aeronautics & Defense industry, in applications including local reconnaissance, border surveillance, search and rescue operations and real-estate aerial photography, the latter also being a demand of the agricultural field as cited by Colomina and Molina (2014).

Multi-rotors vehicles present high maneuverability and allow for vertical landing and take-off, which makes them very suitable for flight in areas of difficult access. Moreover, when dealing with micro multi-rotor vehicles (MAVs), missions can be performed virtually anywhere, from urban canyons to the inside of buildings. Quad-rotors are particularly interesting for their capability of stable hovering and precise flight, accomplished through balancing the propulsion forces of the four rotors (Lozano, 2013). Accompanied by their many advantages, however, come the challenges imposed by the small size and fast dynamics of these vehicles. In order to accomplish critical missions, it's desirable for the MAV to achieve a high degree of autonomy, and for that an accurate and reliable navigation system is of paramount importance. The navigation system aims at providing a non-ambiguous and robust representation of robotic motion, as defined by Corke et al. (2007), with respect to some reference coordinate system and based on measurements acquired from different sensors.

As it became clear in recent years, vision-aided inertial navigation systems are a good fit for implementation in MAVs in GPS-denied environments, and data from the inertial measurement unit (IMU) and the visual system can be combined through several sensor fusion techniques. Aspects of vision-aided inertial navigation of UAVs using a single camera have been addressed in many works, such as Graovac (2004), that investigated the separate contributions of each system in the overall accuracy of the navigation problem, and Webb et al. (2007), who explored the robustness issues using a MAV simulation model developed at NASA. Blosch et al. (2010) went further and presented an efficient vision-based navigation and guidance scheme for an UAV that is capable of successfully hovering and following set-points in an unknown and unstructured indoor environment, using an IMU and a downward looking monocular camera.

Whilst Monte Carlo Simulations provide valuable information on each filtering method, it is fundamental to have a controlled real-world comparative basis for practical implementation of the algorithm, so there can be a mindful choice of the more suitable approach for each goal. The present work is therefore concerned with the comparative evaluation of three non-linear filters based on Kalman et al. (1960) scheme in an experimental platform, namely a Quanser 3-DOF Hover. The specific issue comprises the on-board attitude determination in a experimental setup of a quad-rotor platform in a GPS-denied environment with known landmarks.

The chosen filters are the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF) and Ensemble Kalman Filter (EnKF), all suboptimal in the minimum square error sense but with varying characteristics. While the EKF (Gelb, 1974) is considered the traditional benchmark for non-linear dynamics, the UKF (Wan and Van Der Merwe, 2000) has proven to be very effective on navigation issues and the EnKF was only recently been rediscovered by the signal processing field (Roth et al., 2017), as it is capable to deal with a very large number of variables and huge amount of data, which seems to be the natural progression of today's technologies.

The paper is organized as follows: Section 2 presents the system modeling in a space state approach for implementation within the aforementioned filters, which are further described as proposed solutions in Section 3. The experimental setup is presented in Section 4, and Section 5 discusses the implementation and obtained results. Section 6 finalizes the paper by drawing some conclusions.

2 SYSTEM MODELING

This section provides the equations applied in the modeling of the quadrotor platform and its sensors. First, the adopted notation is summarized, followed by the description of the state and measurement equations of the system, together with the encoder measurements that were considered the truth state for comparison purposes.

2.1 General Notation

The notation used for this problem was adapted from Santos and Goncalves (2016). Essential definitions will be briefly referenced now for straightforward understanding of the problem modeling. The cartesian coordinate systems (CCS) are denoted by a set of three orthogonal versors as $\mathcal{S}_{\mathrm{R}} = \{\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{R}}, \mathbf{z}_{\mathrm{R}}\}$. The projection of a geometric vector onto a CCS, namely the algebraic vector, is also denoted with an uppercase letter in the subscript ($\mathbf{r}_{\mathrm{A}}$). Measurement of a vector will be denoted by $\check{\mathbf{r}}_{\mathrm{R}}$ and its estimate by $\hat{\mathbf{r}}_{\mathrm{R}}$. Finally, the Attitude of a CCS w.r.t. other CCS, in other words, the three-dimensional orientation of $\mathcal{S}_{\mathrm{A}}$ w.r.t. $\mathcal{S}_{\mathrm{B}}$ is the attitude matrix represented as $\mathbf{D}^{\mathrm{A/B}}$. Thus, the attitude matrix $\mathbf{D}^{\mathrm{A/B}}$ is such that $\mathbf{r}_{\mathrm{A}} = \mathbf{D}^{\mathrm{A/B}}\mathbf{r}_{\mathrm{B}}$.

For the purpose of this work, six CCS are defined to comply with the experimental arrange: *Ground*, the ground reference upon which the landmarks positions are known: $\mathcal{S}_{\mathrm{G}} = \{\mathbf{x}_{\mathrm{G}}, \mathbf{y}_{\mathrm{G}}, \mathbf{z}_{\mathrm{G}}\}$; *Reference*, fixed at the center of rotation of the platform, aligned with the initial encoder orientation $\mathcal{S}_{\mathrm{R}} = \{\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{R}}, \mathbf{z}_{\mathrm{R}}\}$; *Encoder*, according to the built-in encoders of the platform: $\mathcal{S}_{\mathrm{ENC}} = \{\mathbf{x}_{\mathrm{ENC}}, \mathbf{y}_{\mathrm{ENC}}, \mathbf{z}_{\mathrm{ENC}}\}$; *Body*, MAV's Gravitational Center: $\mathcal{S}_{\mathrm{B}} = \{\mathbf{x}_{\mathrm{B}}, \mathbf{y}_{\mathrm{B}}, \mathbf{z}_{\mathrm{B}}\}$; *Camera*, fixed on its optical center: $\mathcal{S}_{\mathrm{C}} = \{\mathbf{x}_{\mathrm{C}}, \mathbf{y}_{\mathrm{C}}, \mathbf{z}_{\mathrm{C}}\}$; *Gyroscope*, according to the positioning of the Inertial Measurement Unit used: $\mathcal{S}_{\mathrm{Gy}} = \left\{\mathbf{x}_{\mathrm{Gy}}, \mathbf{y}_{\mathrm{Gy}}, \mathbf{z}_{\mathrm{Gy}}\right\}$. As an usual practice, at instant zero, $\mathcal{S}_{\mathrm{B}}$ is defined as to be aligned with $\mathcal{S}_{\mathrm{R}}$.

2.2 Model Description

We wish to compute the MAV's three-dimensional orientation with respect to a frame or reference. The system model combines a non-linear model and real-world information provided by a triad of rate-gyros and images from one upward camera attached to the Quanser platform. Whilst measurements of the triad of rate-gyros are available at each instant and will be considered inputs of the system, the information extracted from the images have a much higher sample time and will be used as the measurements per se.

The model assumes there are eleven landmarks with known position with respect to $\mathcal{S}_{\mathrm{G}}$. Furthermore, the position ($\vec{\boldsymbol{r}}^{\,\mathrm{C/B}}$) and attitude ($\boldsymbol{D}^{C/B}$) of the camera, rate-gyro ($\boldsymbol{D}^{Gy/B}$) and encoders ($\boldsymbol{D}^{Enc/B}$) w.r.t. $\mathcal{S}_{\mathrm{B}}$ are known.

The state and measurement equations are shown below and derived in the sequel:

$$\dot{\mathbf{x}} = \mathbf{f}\left(\mathbf{x}, \mathbf{u}\right) + \mathbf{G}\left(\mathbf{x}\right)\mathbf{w} \tag{1}$$

$$\mathbf{y}^i = \mathbf{h}^i(\mathbf{x}) + \mathbf{n}^i,\ i = 1, ..., 11. \tag{2}$$

where $\mathbf{f}(\mathbf{x}, \mathbf{u})$ and $\mathbf{G}(\mathbf{x})$ are non-linear functions and $\mathbf{w}$ is a zero-mean white Gaussian noise with covariance $\mathbf{Q}$. For each landmark i, $\mathbf{h}^i \in \mathbb{R}^3$ is non linear, $\mathbf{n}^i \in \mathbb{R}^3$ is a noise term, assumed to be zero-mean white Gaussian process, with known covariance $\mathbf{R}^i \in \mathbb{R}^{3\times3}$.

State Equation

The Attitude Estimation problem consists of estimating the three-dimensional orientation of $\mathcal{S}_\mathrm{B}$ with respect to $\mathcal{S}_\mathrm{G}$. The Direction Cosine Matrix (DCM) that relates both CCS is unambiguous and does not present any singularities, but requires a high number of states to be estimated, which becomes undesirable for practical implementations. It is usual to parametrize the DCM, and for this work the chosen parameters were the Euler angles (Shuster and Markley, 2003), which can be related to the DCM through elemental rotations as shown in Lee et al. (1967). As it was later stated by Bar-Itzhack and Idan (1987), while the use of Euler angles can be inconvenient due to possible singularities, in many cases the vehicle is attitude limited, which means it will not reach the singularity range, as it happens with passenger and cargo planes, helicopters and multi-rotor vehicles such as the one discussed here. Although this representation is not the most computationally advantageous, it is both well known and more intuitive, since the control commands are usually direct functions of the vehicle Euler angles. Ergo, for the purposes of this work, the Euler angles are considered satisfactory. They are defined according to the right-hand rule: when positive, they represent a clockwise rotation considering the positive direction of each axis of the CCS in the 123 sequence.

Let $\boldsymbol{\alpha}$ denote the three Euler angles, namely *roll*, *pitch* and *yaw* respectively. So $\boldsymbol{\alpha} = [\phi\ \theta\ \psi]^T$, and one can express $\boldsymbol{D}(\boldsymbol{\alpha})$ as:

$$\boldsymbol{D}(\boldsymbol{\alpha}) = \begin{bmatrix} \mathrm{c}\psi\mathrm{c}\theta & \mathrm{c}\psi\mathrm{s}\theta\mathrm{s}\phi + \mathrm{s}\psi\mathrm{c}\phi & -\mathrm{c}\psi\mathrm{s}\theta\mathrm{c}\phi + \mathrm{s}\psi\mathrm{s}\phi \\ -\mathrm{s}\psi\mathrm{c}\theta & -\mathrm{s}\psi\mathrm{s}\theta\mathrm{s}\phi + \mathrm{c}\psi\mathrm{c}\phi & \mathrm{s}\psi\mathrm{s}\theta\mathrm{c}\phi + \mathrm{c}\psi\mathrm{s}\phi \\ \mathrm{s}\theta & -\mathrm{c}\theta\mathrm{s}\phi & \mathrm{c}\theta\mathrm{c}\phi \end{bmatrix} \in \mathrm{SO}(3), \tag{3}$$

where c stands for cosine and s for sine, and:

$$\phi = -tan^{-1}\left(\frac{D_{32}}{D_{33}}\right), 0^o \leqslant \phi < 2\pi, \tag{4}$$

$$\theta = sin^{-1}(D_{31}), -\frac{\pi}{2} < \theta < \frac{\pi}{2}, \tag{5}$$

$$\psi = -tan^{-1}\left(\frac{D_{21}}{D_{11}}\right), 0^o \leqslant \psi < 2\pi. \tag{6}$$

The gyroscope measurement can be modeled as a combination of the true body angular velocity w.r.t. $\mathcal{S}_\mathrm{G}$, with the sensor's bias and noise:

$$\breve{\boldsymbol{\omega}}_{\mathrm{Gy}} = \boldsymbol{\omega}_{\mathrm{Gy}}^{\mathrm{B/G}} + {}^g\boldsymbol{\beta} + {}^g\mathbf{w}, \tag{7}$$

where $\boldsymbol{\omega}_{\mathrm{Gy}}^{\mathrm{B/G}}$ is the true angular velocity of the body measured in $\mathcal{S}_{\mathrm{Gy}}$, ${}^g\boldsymbol{\beta} \in \mathbb{R}^3$ is the bias term assumed to be constant and ${}^g\mathbf{w} \in \mathbb{R}^3$ is a zero-mean white Gaussian noise with covariance

${}^{g}\mathbf{Q} \in \mathbb{R}^{3\times 3}$. Due to the mounting on platform of the inertial measurement unit, $D^{B/Gy}$ is known and the angular velocity can be represented in $\mathcal{S}_{\mathrm{B}}$.

Equation (1) then becomes:

$$\dot{\boldsymbol{\alpha}} = \mathcal{A}(\boldsymbol{\alpha})\boldsymbol{\omega}_{\mathrm{B}}^{\mathrm{B/G}} = \mathcal{A}(\boldsymbol{\alpha})\breve{\boldsymbol{\omega}}_{\mathrm{B}}^{\mathrm{B/G}} - \mathcal{A}(\boldsymbol{\alpha})^{g}\mathbf{w}, \tag{8}$$

with

$$\mathcal{A}(\boldsymbol{\alpha}) \triangleq \begin{bmatrix} \mathrm{c}\psi/\mathrm{c}\theta & -\mathrm{s}\psi/\mathrm{c}\theta & 0 \\ \mathrm{s}\psi & \mathrm{c}\psi & 0 \\ -\mathrm{c}\psi\mathrm{s}\theta/\mathrm{c}\theta & \mathrm{s}\psi\mathrm{s}\theta/\mathrm{c}\theta & 1 \end{bmatrix}. \tag{9}$$

Measurement Equation

The camera is modeled as pinhole (Forsyth and Ponce, 2011), attached to the Quanser structure at a known distance $\vec{\boldsymbol{r}}^{\,\mathrm{C/B}}$ of $\mathcal{S}_{\mathrm{B}}$, as shown in Figure 1. Geometrically through vector addition, one can notice that:

$$\vec{\boldsymbol{r}}^{\,\mathrm{B/G}} + \vec{\boldsymbol{r}}^{\,\mathrm{C/B}} + \vec{\boldsymbol{s}}^{\,\mathrm{i}} = \vec{\boldsymbol{p}}^{\,\mathrm{i}}, \tag{10}$$

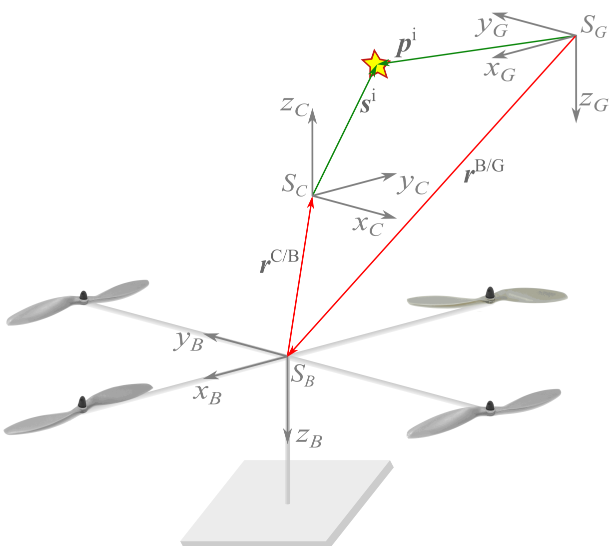


Figure 1: Vectorial computation for each landmark.

The $\mathcal{S}_{\mathrm{G}}$ representation of $\vec{\boldsymbol{s}}^{\,\mathrm{i}}$ is given by:

$$\boldsymbol{s}_{\mathrm{G}}^{i} = \boldsymbol{p}_{\mathrm{G}}^{i} - \boldsymbol{r}_{\mathrm{G}}^{\mathrm{B/G}} - \left(\boldsymbol{D}^{\mathrm{B/G}}(\boldsymbol{\alpha})\right)^{\mathrm{T}} \boldsymbol{r}_{\mathrm{B}}^{\mathrm{C/B}}, \tag{11}$$

where for $i = 1, ..., 11$; $\boldsymbol{p}_{\mathrm{G}}^{i}$ and $\boldsymbol{s}_{\mathrm{G}}^{i}$ are respectively the representations of $\boldsymbol{p}^{i}$ and $\boldsymbol{s}^{i}$ in $\mathcal{S}_{\mathrm{G}}$. The rotation matrix $\boldsymbol{D}^{\mathrm{B/G}}(\boldsymbol{\alpha})$ is given by Equation (3), and $\boldsymbol{r}_{\mathrm{B}}^{\mathrm{C/B}}$ is the $\mathcal{S}_{\mathrm{B}}$ representation of the camera position relative to the body (known).

The $\mathcal{S}_\mathrm{B}$ representation of $\boldsymbol{s}^i$ can be defined using the rotation matrix w.r.t. the ground:

$$\boldsymbol{s}^i_\mathrm{B} = \boldsymbol{D}^{\mathrm{B/G}}(\boldsymbol{\alpha})\boldsymbol{s}^i_\mathrm{G}. \tag{12}$$

Since the camera is attached to the body's structure, the attitude matrix of $\mathcal{S}_\mathrm{C}$ w.r.t. $\mathcal{S}_\mathrm{B}$ is known and can be expressed by $\boldsymbol{D}^{\mathrm{C/B}}$. Thus the representation of $\boldsymbol{s}^i$ with regard to the CCS of the camera can be expressed from $\boldsymbol{s}^i_\mathrm{B}$:

$$\boldsymbol{s}^i_\mathrm{C} = \boldsymbol{D}^{\mathrm{C/B}}\ \boldsymbol{s}^i_\mathrm{B}. \tag{13}$$

These past equations can be combined to form $\boldsymbol{h}^i$ shown in Eq. 2, for each landmark i:

$$\boldsymbol{h}^i\left(\boldsymbol{x}\right) \triangleq \boldsymbol{D}^{\mathrm{C/B}}\boldsymbol{D}^{\mathrm{B/G}}(\boldsymbol{\alpha})\left(\boldsymbol{p}^i_\mathrm{G} - \boldsymbol{r}^{\mathrm{B/G}}_\mathrm{G} - \left(\boldsymbol{D}^{\mathrm{B/G}}(\boldsymbol{\alpha})\right)^\mathrm{T}\boldsymbol{r}^{\mathrm{C/B}}_\mathrm{B}\right) \in \mathbb{R}^3 \tag{14}$$

Truth State

In simulation analysis, one can always compare the state estimates with the known truth state. Thats is not the case for practical implementation of navigation algorithms, where one must consider another reliable measurement as the basis for comparison. In this work, the truth state is assumed to be acquired from the built-in high accuracy incremental encoders of the Quanser Platform.

As proven by Hughes (2012), successive rotations can be concatenated to yield the rotation matrix between two specific CCS. Since we have the attitude of $\mathcal{S}_\mathrm{Enc}$ w.r.t. $\mathcal{S}_\mathrm{B}$ due to the physical platform and the later is usually aligned with $\mathcal{S}_\mathrm{R}$ at instant zero, the body's attitude of interest can be found using the attitude $\boldsymbol{D}^\mathrm{Enc}_\mathrm{R}$ at each given instant concatenated with $\boldsymbol{D}^\mathrm{Enc}_\mathrm{B}$. To avoid errors due to misalignments of the platform known to happen in real world applications, the reference system $\mathcal{S}_\mathrm{R}$ is chosen to be aligned with $\mathcal{S}_\mathrm{Enc}$ at time zero of the experiment. In this way, all three CCS are aligned at instant zero and errors can be minimized.

3 PROPOSED SOLUTIONS

The inertial-vision navigation problem goal is to obtain an approximation for the minimum mean squared error (MMSE) estimate of the state ($\mathbf{x}$), using the vector measurements ($\mathbf{y}^i$, for $i = 1, ..., 11$ landmarks), the inertial measurements ($\mathbf{u}$), the state equation (Eq. 1) and the measurement equation (Eq. 2).

Three suboptimal filters in the MMSE sense are proposed as solutions for the sensor fusion task: the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF) and Ensemble Kalman Filter (EnKF). They are all variations of the classic Kalman filter (Kalman et al., 1960) for use with non-linear systems, and they all apply the general Forecast and Update scheme shown below. The notation $k+1|k$ stands for prediction of $k+1$ given measurements up until instant k, where x, P and y symbols state, covariance and system output, respectively:

Prediction: Given $\hat{\boldsymbol{x}}_{k|k}$ and $\boldsymbol{P}_{k|k}$:

$$\hat{\boldsymbol{x}}_{k+1|k} \triangleq \mathbb{E}\left[\boldsymbol{x}_{k+1}|\boldsymbol{y}_{1:k}\right], \tag{15}$$

$$\boldsymbol{P}_{k+1|k} \triangleq \mathbb{E}\left[\left(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k}\right)^\mathrm{T}|\boldsymbol{y}_{1:k}\right], \tag{16}$$

$$\hat{\boldsymbol{y}}_{k+1|k} \triangleq \mathbb{E}\left[\boldsymbol{y}_{k+1}|\boldsymbol{y}_{1:k}\right], \tag{17}$$

$$\boldsymbol{P}^{Y}_{k+1|k} \triangleq \mathbb{E}\left[\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}\right)\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}\right)^{\mathrm{T}}|\boldsymbol{y}_{1:k}\right], \tag{18}$$

$$\boldsymbol{P}^{XY}_{k+1|k} \triangleq \mathbb{E}\left[\left(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}\right)^{\mathrm{T}}|\boldsymbol{y}_{1:k}\right], \tag{19}$$

Update: Given measurements $\boldsymbol{y}_{k+1}$,

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1}\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}\right), \tag{20}$$

$$\boldsymbol{P}_{k+1|k+1} = \boldsymbol{P}_{k+1|k} - \boldsymbol{P}^{XY}_{k+1|k}\left(\boldsymbol{P}^{Y}_{k+1|k}\right)^{-1}\left(\boldsymbol{P}^{XY}_{k+1|k}\right)^{T}, \tag{21}$$

where

$$\boldsymbol{K}_{k+1} \triangleq \boldsymbol{P}^{XY}_{k+1|k}\left(\boldsymbol{P}^{Y}_{k+1|k}\right)^{-1}. \tag{22}$$

The Extended Kalman Filter (Gelb, 1974) is a traditional estimation method for non-linear systems, known since 1970. It comprises a series expansion of the non-linear dynamics and measurement equations, e.g. linearization by Taylor Series approximation. It then propagates the approximated mean and covariance of the linearized system. This method is considered a benchmark for non-linear systems, but has the main disadvantage of requiring the calculation of the system's Jacobians. They are often computed analytically and updated at each iteration.

The Unscented Kalman Filter (Wan and Van Der Merwe, 2000) relies on the unscented transformation, e.g. a discrete approximation of the continuous probability density function (pdf) of the state variable: it chooses deterministic sigma-points from the random state pdf and attributes weights for each point, which are then propagated trough the non-linear model of the system to result in the measurement sigma-points. These new sigma-points are used for computation of the forecast mean and covariance of the system output, without the need for linearization of the system. It is a fairly simple algorithm to be implemented, and its computational complexity is comparable to that of the EKF.

The Ensemble Kalman Filter, as described in Evensen (2003), belongs to the category of particle filters and has been used mainly in weather forecast applications, where often the models are of high order and extremely nonlinear, the initial states are highly uncertain, and there is a large number of measurements available (Gillijns et al., 2006). The EKF has only recently been brought to the signal processing area as seen in Roth et al. (2017), yet it reflects the necessity of handling an always increasing amount of data that is required from automated systems nowadays. The starting point of the EnKF is to choose a set of sample points, or in other words, an ensemble of state estimates. This ensemble of points is then propagated using the non-linear equations of the modeled system and make it possible to compute the mean and covariance of the system output. Since the sample points are chosen at random, they tend to reflect the distribution of the true state.

4 EXPERIMENTAL SETUP

The setup consists of a Quanser 3-DOF Hover kit (Quanser, 2017) with strapdown three-axis rate-gyros as well as an upward-facing camera. Over the camera, there is a panel containing a number of Aruco tags, which represent the landmarks. The sensor measurements and the camera images are acquired and recorded into a MyRIO board (Instruments, 2017) while the encoders of the hover are read in a MATLAB/Simulink (Bradski, 2017) environment. After the experiment, all this data is previously processed and synchronized using a MATLAB script, which generates the standard inputs for the attitude determination algorithms described in Section 3. The framework of the experiment can be seen in Fig. (2).

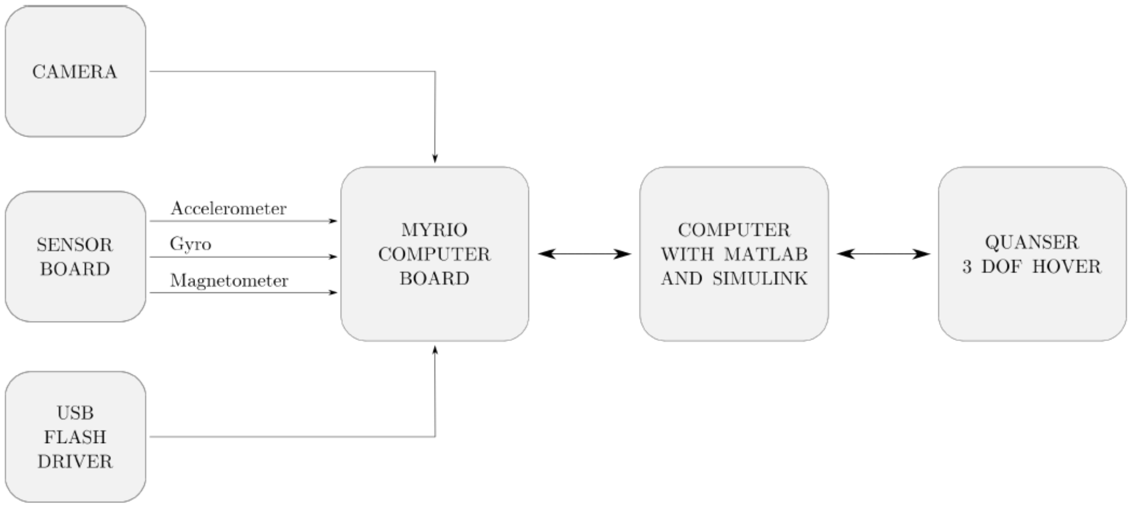


Figure 2: Experiment framework.

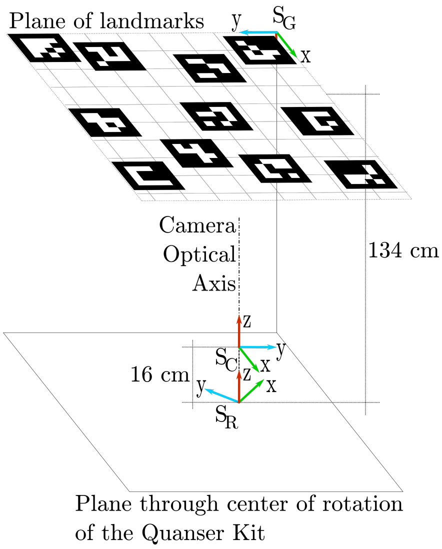
The experimental process consists at first to align the Quanser equipment to a delimited and previously established position relative to the plane of landmarks. Figure (3b) shows the physical setup of the experiment and Figure (3a) shows two frames of reference: S_r representing the fixed reference frame positioned at the center of rotation of the Quanser and S_c, representing a frame solidary to the camera.

After positioning, the MyRio board and the computer are setup to start recording data and a MATLAB application controls the 3 DOF Hover to perform a pre-programmed movement in order to simulate a MAV in flight.

A post-processing stage synchronizes and standardizes the data from the sensors, from the images and from the encoders in a structured file. The sensors and Aruco 3D relative positions are passed on to the algorithms of visual-inertial attitude determination, and those are compared to the attitude computed from the encoders' measurements.

The Quanser 3 DOF Hover (Quanser, 2017) is a flight dynamics and control research kit that can be operated with Matlab and Simulink, version R2015a. The sensor board is composed by a digital accelerometer, model ADXL345; a digital gyro model ITG-3200 and a digital compass model HMC5883L. In this experiment, only the rate-gyro was used.

The camera is an USB Web Cam Microsoft LifeCam Cinema, with resolution up to HD and a frame rate of 30 frames per second. MyRio is a computer board developed by National Instruments with a dual core ARM processor and a FPGA chip. This structure is responsible to record the data from the sensors and camera images in a folder at USB flash drive and it is programmed through LabView.

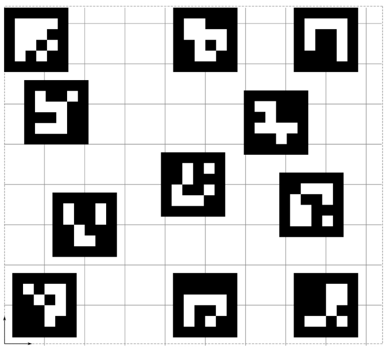


(a) Experimental setup schematic.

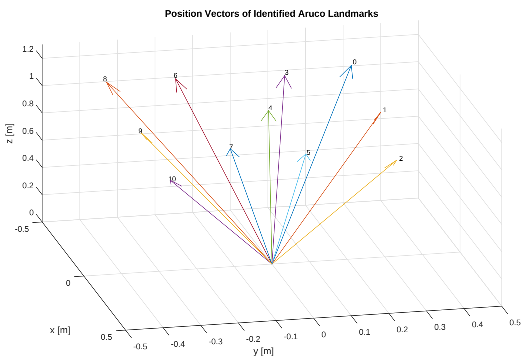
(b) Actual experimental arrangement.

Figure 3: Schematic and actual experiment layout with landmarks.

The camera images allow to identify landmarks used to improve the attitude estimation. This work utilizes the Aruco Tags (Garrido-Jurado et al., 2014, 2016) as landmarks, which allow the determination of the center of each tag with respect to the camera CCS (S_C). The landmarks' plane is shown in Fig (4).



(a) Distribution of tags in a paper of 90x100 cm.



(b) Representation of the position vector from camera CCS S_C to each Aruco landmark.

Figure 4: Images of plane of Aruco tags, used as landmarks

The image process stage was developed based on Opencv (Bradski and Itseez, 2017), an open source multi-platform and multi-language computer vision library. The Aruco is among the bundle of functions to general purpose image processing and computer vision available at Opencv, and was developed to virtual reality applications, allowing a precise pose determination between the camera and the tags.

5 IMPLEMENTATION AND RESULTS

The attitude determination algorithms discussed in Section 3 were tested using data recorded during approximately 120s for each experiment set. The sensors were recorded at 50 samples per second, while the images were sampled at 10 frames per seconds and the encoders at 10 samples per second. The images and sensors data were processed off-line, with the aid of a MATLAB script. This allowed the synchronization of the material collected through MyRio board (images and inertial measurements) and through the computer (encoder measurements), as well as any data treatment required, such as unit conversions.

During the experiment it was noticed a slight curvature of the Aruco Tags plane, and the measurements acquired from the camera during the first 20 seconds were used to correct the known height or the landmarks w.r.t. $\mathcal{S}_{\mathrm{G}}$, previously assumed to be zero. Although doing an initial reconnaissance of the environment was not predicted in the experiment scope, this step showed that the first images can be used for compensating small mistakes in landmark position data. Furthermore, as the experiment progressed and more data was recorded, it became clear that using the initial camera measurements as the basis for calculation of the landmarks positioning w.r.t. $\mathcal{S}_{\mathrm{R}}$ provided better results than using offline measurements. For example, any misalignment of the supposed orientation of the camera w.r.t. the platform would yield an undesirable offset between the attitude measured by the encoders and the attitude estimated using the algorithm. In addition, any modification of the Quanser position w.r.t. $\mathcal{S}_{\mathrm{G}}$ would require new measurements the vector $\vec{\boldsymbol{r}}^{\,\mathrm{B/G}}$, which is assumed to be known since the experiment purpose is specific attitude determination. By using the initial images of the camera as a calibration tool for the calculation of the landmarks and body position w.r.t. to the reference system, all these real world issues were overcome.

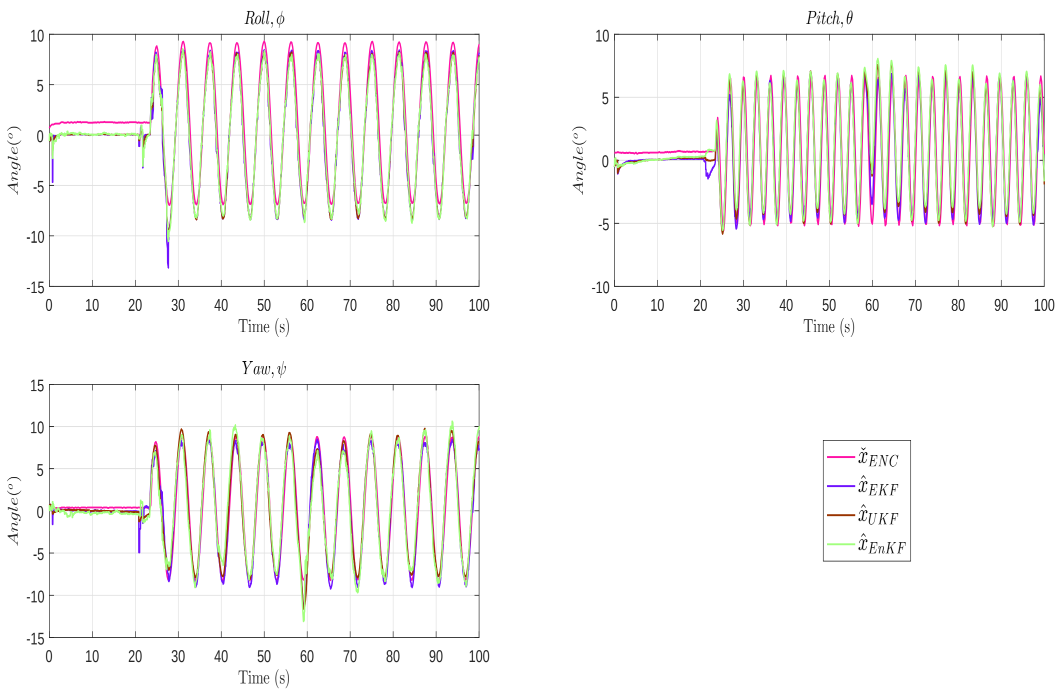


Figure 5: Attitude estimation on sinusoidal rotation movement

The attitude determination was performed as if the data was arriving at that moment, i.e., at any given instant, the prediction stage was accomplished using the system model and the rate-gyro input while the update stage would only be carried out if image data became available. The algorithms also dealt with missing landmarks, since not all of them were visible at all times due to the camera angle or even some failure in image capturing.

All three algorithms were tuned equally for better comparison and the results can be seen in Figure 5 below for an sinusoidal rotation. The body is kept at hovering state for the first 20 seconds to allow for the initial calibration, and after another 20 seconds of movement, the estimation error norm is found as 1.11 degrees using the EKF, 1.08 degrees for the UKF and 1.12 degrees using the EnKF.

6 CONCLUSIONS

This paper described the dynamic modeling of a quad-rotor vehicle consistent with the experimental platform Quanser 3D Hover and a number of solutions for the attitude estimation problem were presented. These solutions were implemented in an experiment setup, devised to allow performance investigation of different sensor fusion algorithms in a real, but restrained environment. The results from this experiment were twofold.

First, it was possible to analyze the performance of different attitude determination algorithms in an physical structure, accounting for all the non-linearities, misalignments and noise intrinsically present in a real world application. For the sinusoidal movement studied in this work, the extended Kalman filter, the unscented Kalman filter and the ensemble Kalman filter depicted similar results regarding the estimation error.

Furthermore, the experiment is also a step towards the development of a reliable test platform for future investigations. These may include throughly examination of the aforementioned filters' performance as well as other methods, in this setup or even in more complex ones, with more sensors and more states for instance. An intended future work is the estimation of not only the platform attitude but its position and sensors' biases using a stereo camera system. Also, in mirroring real time applications and dealing with practical difficulties, we hope the algorithms can be better analyzed and improved for future contributions within the area of increasingly embedded and autonomous aerial systems.

ACKNOWLEDGMENTS

The authors appreciate the support of CAPES, CNPq, FAPESP and ITA.

REFERENCES

Bar-Itzhack, I. and Idan, M. (1987). Recursive attitude determination from vector observations: Euler angle estimation. *AIAA Journal of Guidance Control and Navigation*.

Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2004). *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons.

Blosch, M., Weiss, S., Scaramuzza, D., and Siegwart, R. (2010). Vision based mav navigation in unknown and unstructured environments. In *Robotics and automation (ICRA), 2010 IEEE international conference on*, pages 21–28. IEEE.

Bradski, G. (2017). Mathwors - makers of matlab and simulink. https://www.mathworks.com/.

Bradski, G. and Itseez (2017). Opencv - open source computer vision library. http://opencv.org.

Colomina, I. and Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92:79–97.

Corke, P., Lobo, J., and Dias, J. (2007). An introduction to inertial and visual sensing. *The International Journal of Robotics Research*, 26:519–535.

Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367.

Fahlstrom, P. and Gleason, T. (2012). *Introduction to UAV systems*. John Wiley & Sons.

Forsyth, D. and Ponce, J. (2011). *Computer vision: a modern approach*. Upper Saddle River, NJ; London: Prentice Hall.

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., and Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292.

Garrido-Jurado, S., Munoz-Salinas, R., Madrid-Cuevas, F. J., and Medina-Carnicer, R. (2016). Generation of fiducial marker dictionaries using mixed integer linear programming. *Pattern Recognition*, 51:481–491.

Gelb, A. (1974). *Applied optimal estimation*. MIT press.

Gillijns, S., Mendoza, O. B., Chandrasekar, J., De Moor, B., Bernstein, D., and Ridley, A. (2006). What is the ensemble kalman filter and how well does it work? In *American Control Conference, 2006*, pages 6–pp. IEEE.

Graovac, S. (2004). Principles of fusion of inertial navigation and dynamic vision. *Journal of Field Robotics*, 21(1):13–22.

Hughes, P. C. (2012). *Spacecraft attitude dynamics*. Courier Corporation.

Instruments, N. (2017). myrio - student embedded device. http://www.ni.com/pt-br/shop/engineering-education/portable-student-devices/myrio-student-embedded-device/what-is-myrio.html.

Kalman, R. E. et al. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45.

Lee, H., Meyer, G., and Wehrend, W. R. (1967). A method for expanding a direction cosine matrix into an euler sequence of rotations.

Lozano, R. (2013). *Unmanned aerial vehicles: Embedded control*. John Wiley & Sons.

Mueller, T. J. (2009). On the birth of micro air vehicles. *International Journal of Micro Air Vehicles*, 1(1):1–12.

Quanser (2017). 3 dof hover. http://www.quanser.com/products/3dof_hover.

Roth, M., Hendeby, G., Fritsche, C., and Gustafsson, F. (2017). The ensemble kalman filter: A signal processing perspective. *arXiv preprint arXiv:1702.08061*.

Santos, D. and Goncalves, P. (2016). Attitude determination of multirotor aerial vehicles using camera vector measurements. *Journal of Intelligent & Robotic Systems*, 86(1):139–149.

Shuster, M. D. and Markley, F. L. (2003). Generalization of the euler angles. *Journal of the Astronautical Sciences*, 51(2):123–132.

Wan, E. A. and Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. Ieee.

Webb, T. P., Prazenica, R. J., Kurdila, A. J., and Lind, R. (2007). Vision-based state estimation for autonomous micro air vehicles. *Journal of Guidance Control and Dynamics*, 30(3):816–826.