24th COBEM - 2017

24th ABCM International Congress of Mechanical Engineering
December 3-8, 2017, Curitiba, PR, Brazil

COBEM-2017-1006
COMPARISON OF FILTERING METHODS FOR STEREO VISUAL-INERTIAL NAVIGATION OF MULTIROTOR AERIAL VEHICLES

Beatriz Arruda Asfora
Davi Antônio dos Santos
Instituto Tecnológico da Aeronáutica, Praça Marechal Gomes 50, São José dos Campos, 12228-900 SP, Brasil
beatriz.asfora@gmail.com
davists@ita.br

Abstract. *Vision-aided inertial systems are thought to be a good fit for the micro multirotor aerial vehicle navigation issue, not only due to the camera cost, lightweight and low power consumption, but also for its ability to provide rich information about the environment. The present paper focus on the evaluation of three non-linear Kalman-based filters applied to the estimation problem of the position, linear velocity and attitude of an aerial vehicle, along with the biases of its inertial sensors. A non-linear dynamic model is proposed, allowing sensor fusion to be performed in a prediction-update framework, using measurements from a strapdown inertial measurement unit and a stereo visual system in a GPS-denied environment with four known landmarks. The extended Kalman filter, unscented Kalman filter and ensemble Kalman filter are compared regarding accuracy, computational burden and robustness with respect to initial conditions, in a simulated scenario using the same tuning parameters. Results from Monte Carlo simulations show that regarding accuracy and precision, the unscented Kalman filter proved to be equivalent or superior to the other two in all configurations that were tested. As expected, the extended Kalman filter remains as the computationally lighter approach, but not the most robust one: in this matter, the still not fully explored ensemble Kalman filter excelled.*

Keywords: *Visual-inertial navigation, multirotor-aerial vehicles, sensor fusion, Kalman filter*

1. INTRODUCTION

Unmanned aviation can be traced down to the very beginning of human aviation, for pilotless models were commonly used as testing devices before attempting flight in man-carrying vehicles since the early development of aircrafts (Newcome, 2004). Unmanned aerial vehicles (UAVs) evolved to surpass their initial role and today they are often associated with cutting edge technology. They present in diverse sizes and autonomy levels according to the application they are intended for, being used not only in the fields of Aerospace and Military, but in a wide range of applications such as structure inspection and maintenance (Metni and Hamel, 2007; Nikolic *et al.*, 2013), search and rescue operations (Doherty and Rudol, 2007), precision agriculture (Zhang and Kovacs, 2012; Alsalam *et al.*, 2017), aerial surveillance and reconnaissance (Wallar *et al.*, 2015), monitoring of wildfire (Phan and Liu, 2008), livestock (Xu *et al.*, 2015) or crowds (Mohammed *et al.*, 2014), delivery (Scott and Scott, 2017; Murray and Chu, 2015) and recreational activities.

Particularly, there has been an increasing interest in micro multirotor aerial vehicles (MAVs), due to their high maneuverability in restricted spaces and capability of providing information on demand, where and when it is needed, allied with relatively short time fabrication and affordability. These advantages however are coupled with their fast dynamics and small size, imposing challenges for their navigation and control (Lozano, 2013). An accurate and reliable navigation system is of paramount importance to achieve a high degree of autonomy in an MAV and make it capable to cope with critical missions. As clearly stated by Corke *et al.* (2007), the navigation system aims at providing a non-ambiguous and robust representation of robotic motion, with respect to some reference coordinate system and based on measurements acquired from different sensors.

Inertial sensors such as gyroscopes and accelerometers are a good fit for navigation issues because they are independent of visibility conditions or external information, providing a self-contained navigation technique. Inertial navigation has traditionally been applied in a variety of platforms, such as aircrafts, submarines, spacecrafts and ships, but with the advances in Micro-Electronic Mechanical Systems (MEMS) during the 1990s, it has been possible to use them also in small platforms such as MAVs. Inertial navigation however integrates the errors and biases of the measurements, increasing navigation error without bound and producing time-diverging estimates. Experiments conducted by Woodman (2007) with an Inertial Measurement Unit (IMU) present in an off-the-shelf motion tracking device showed that the average error in position can reach 150 m after only 60 s of operation if the IMU is used by itself.

In order to obviate the error characteristics of pure inertial navigation, Global Positioning System (GPS) is commonly used. Still, GPS may be quite unreliable or even unavailable when operating in places with poor lines-of-sight to GPS satellites, such as interior buildings, caves, mines, forests, urban canyons, and, in general, low altitude environments (Blösch *et al.*, 2010), all of these being scenarios of interest for operations with MAVs. Visual inertial navigation is a strong alternative for the limitations posed by the GPS-INS approach, and since cameras are lightweight, cost-effective sensors, compact, and with low power consumption, they are well-suited for MAVs, as seen in Carrillo *et al.* (2012) and Kun (2015). Moreover, they provide rich information on the environment, being also useful for obstacle deviation, collision avoidance, and mapping. As stated by Forster *et al.* (2014), MAVs are expected to play an important role in critical areas such as disaster management and environment conservation, and for that GPS-INS information only is simply not enough.

The present work is concerned with the navigation problem of an MAV, which consists of estimating the vehicle's 3D position, velocity, attitude and sensors biases. The proposed method relies on sensor fusion, using vector measurements extracted from two downward facing cameras and a strap-down IMU, composed of triaxial accelerometers and rate-gyros. Three suboptimal filters in the minimum square error sense are investigated for solving the sensor fusion task: the extended Kalman filter (EKF), the unscented Kalman filter (UKF) and the ensemble Kalman filter (EnKF). The EKF (Gelb, 1974) is a traditional estimation method and the UKF (Wan and Van Der Merwe, 2000) has been previously designed for many navigation problems with interesting results as in Marina *et al.* (2012) and Andersen and Taylor (2007). The EnKF (Evensen, 2003) on the other hand has been used mainly in weather forecast applications and its evaluation within the signal processing area is relatively new, have being discussed by Gillijns *et al.* (2006) and more recently by Roth *et al.* (2017); we claim there is yet much to investigate about its implementation on navigation systems.

The paper is organized as follows: Section 2 presents the system modeling and futher description of the aforementioned filters. The computational procedure is presented in Section 3 as well as the simulation results, and Section 4 finalizes the paper by drawing some conclusions.

2. SYSTEM MODELING AND PROPOSED SOLUTIONS

2.1 General Notation and Assumptions

The notation employed in this work was adapted from Santos and Gonçalves (2016) and essential definitions will be summarized now for clarification. 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 Direction Cosine Matrix (DCM) represented as $\mathbf{D}^{\mathrm{A/B}}$. Thus, the attitude matrix $\mathbf{D}^{\mathrm{A/B}}$ is such that $\mathbf{r}_{\mathrm{A}} = \mathbf{D}^{\mathrm{A/B}}\mathbf{r}_{\mathrm{B}}$.

Six CCS are defined for this problem: *Ground*, the inertial global reference: $\mathcal{S}_{\mathrm{G}} = \{\mathbf{x}_{\mathrm{G}}, \mathbf{y}_{\mathrm{G}}, \mathbf{z}_{\mathrm{G}}\}$; *Body*, MAV's Gravitational Center: $\mathcal{S}_{\mathrm{B}} = \{\mathbf{x}_{\mathrm{B}}, \mathbf{y}_{\mathrm{B}}, \mathbf{z}_{\mathrm{B}}\}$; *Camera 1*, fixed on its optical center C1: $\mathcal{S}_{\mathrm{C1}} = \{\mathbf{x}_{\mathrm{C1}}, \mathbf{y}_{\mathrm{C1}}, \mathbf{z}_{\mathrm{C1}}\}$; Camera 2 (fixed on its optical center C2): $\mathcal{S}_{\mathrm{C2}} = \{\mathbf{x}_{\mathrm{C2}}, \mathbf{y}_{\mathrm{C2}}, \mathbf{z}_{\mathrm{C2}}\}$; Image 1 (center I of image plane from C1): $\mathcal{S}_{\mathrm{I1}} = \{\mathbf{x}_{\mathrm{I1}}, \mathbf{y}_{\mathrm{I1}}\}$ Image 2 (center I of image plane from C2): $\mathcal{S}_{\mathrm{I2}} = \{\mathbf{x}_{\mathrm{I2}}, \mathbf{y}_{\mathrm{I2}}\}$.

The model assumes there are four landmarks with known position with respect to $\mathcal{S}_{\mathrm{G}}$, which are always inside both cameras' field of view. Furthermore, the position ($\vec{\boldsymbol{r}}^{\,\mathrm{C1/B}}$ and $\vec{r}^{\,\mathrm{C2/B}}$) and attitude ($\boldsymbol{D}^{C1/B}$ and $\boldsymbol{D}^{C2/B}$) of the cameras w.r.t. $\mathcal{S}_{\mathrm{B}}$ are known due to their mounting on the vehicle's frame.

It is well known that the DCM is unambiguous and does not present any singularities (Bar-Shalom *et al.*, 2004), but requires a high number of states to be estimated, which is undesirable for the issue considered here. For this work, the DCM is therefore parametrized in Euler angles through elementary rotations as shown in Lee *et al.* (1967). Since the vehicle in question is attitude limited, it does not reach the singularity range inherent to this representation, similarly to the vehicles considered in Bar-Itzhack and Idan (1987). The Euler angles 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 of rotation.

2.2 Model Description

The problem combines a non-linear model and information provided by a triad of accelerometers, triad of rate-gyros and images from two downward facing cameras attached to the MAV plaftorm. The measurements acquired by the triad of accelerometers and rate-gyros will be considered inputs of the system, while the information extracted from the images will be the measurements per se.

The state equation is shown in Eq. (1), with state and input vectors defined respectively in Eq. (2) and (3).

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

$$\mathbf{x} \triangleq \left[\left(\mathbf{r}_{\mathrm{G}}^{\mathrm{B/G}}\right)^{\mathrm{T}} \ \left(\mathbf{v}_{\mathrm{G}}^{\mathrm{B/G}}\right)^{\mathrm{T}} \ \left(\boldsymbol{\alpha}^{\mathrm{B/G}}\right)^{\mathrm{T}} \ ({}^{a}\boldsymbol{\beta})^{\mathrm{T}} \ ({}^{g}\boldsymbol{\beta})^{T} \right]^{\mathrm{T}} \in \mathbb{R}^{15}, \tag{2}$$

$$\mathbf{u} \triangleq \left[\check{\mathbf{a}}_{\mathrm{B}}^{\mathrm{T}} \ \check{\boldsymbol{\omega}}_{\mathrm{B}}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{6}, \tag{3}$$

where $\mathbf{r}_{\mathrm{G}}^{\mathrm{B/G}} \in \mathbb{R}^3$ and $\mathbf{v}_{\mathrm{G}}^{\mathrm{B/G}} \in \mathbb{R}^3$ are the respective representations of the MAV's position and linear velocity w.r.t. $\mathcal{S}_{\mathrm{G}}$; $\boldsymbol{\alpha}$ is the MAV's attitude also w.r.t $\mathcal{S}_{\mathrm{G}}$, parametrized by Euler angles Sequence 123 , $\boldsymbol{\alpha} = [\phi\ \theta\ \psi]^T$; ${}^{a}\boldsymbol{\beta} \in \mathbb{R}^3$ is the bias of the accelerometer and ${}^{g}\boldsymbol{\beta} \in \mathbb{R}^3$ is the bias of the rate-gyro; $\check{\mathbf{a}}_{\mathrm{B}}$ and $\check{\boldsymbol{\omega}}_{\mathrm{B}}$ represent respectively the accelerometers and gyros measurements.

In order to define the state and noise functions, one must consider the modeling of the sensors measurements, as shown in Eq. (4-5), and attitude kinematics in Euler angles as shown in Eq. (6):

$$\check{\mathbf{a}}_{\mathrm{B}} = \mathbf{D}(\boldsymbol{\alpha})^{\mathrm{B/G}}\left(\dot{\mathbf{v}}_{\mathrm{G}}^{\mathrm{B/G}} - \mathbf{g}_{\mathrm{G}}\right) + {}^{a}\boldsymbol{\beta} + {}^{a}\mathbf{w}, \tag{4}$$

$$\check{\boldsymbol{\omega}}_{\mathrm{B}} = \boldsymbol{\omega}_{\mathrm{B}}^{\mathrm{B/G}} + {}^{g}\boldsymbol{\beta} + {}^{g}\mathbf{w}, \tag{5}$$

$$\dot{\boldsymbol{\alpha}} = \mathcal{A}(\boldsymbol{\alpha})\boldsymbol{\omega}_{\mathrm{B}}^{\mathrm{B/G}}, \tag{6}$$

where $\mathbf{D}(\boldsymbol{\alpha})^{\mathrm{B/G}} \in SO(3)$ represents the attitude of the platform w.r.t. $\mathcal{S}_{\mathrm{G}}$, parametrized in Euler angles 123; $\dot{\mathbf{v}}_{\mathrm{G}}^{\mathrm{B/G}} \in \mathbb{R}^3$ is the $\mathcal{S}_{\mathrm{G}}$ representation of the body acceleration relative to the ground, $\mathbf{g}_{\mathrm{G}} \triangleq [0\ 0\ -g]^{\mathrm{T}}$ is the $\mathcal{S}_{\mathrm{G}}$ representation of the gravitational acceleration with magnitude g; ${}^{a}\boldsymbol{\beta} \in \mathbb{R}^3$ is the accelerometer bias as defined in the state vector and ${}^{a}\mathbf{w} \in \mathbb{R}^3$ is a realization of a zero-mean white Gaussian noise with covariance ${}^{a}\mathbf{Q} \in \mathbb{R}^{3\times3}$. Following the adopted notation, the true angular velocity of the vehicle is represented by $\boldsymbol{\omega}_{\mathrm{B}}^{\mathrm{B/G}}$, and similarly to the accelerometer, ${}^{g}\mathbf{w} \in \mathbb{R}^3$ is a realization of a zero-mean white Gaussian noise with covariance ${}^{g}\mathbf{Q} \in \mathbb{R}^{3\times3}$ related to the gyros' measurements. The attitude kinematics applies the Euler 123 kinematics matrix (Hughes, 2012).

The sensors biases are modeled as continuous-time Wiener processes, with ${}^{\beta a}\boldsymbol{w} \in \mathbb{R}^3$ and ${}^{\beta g}\boldsymbol{w} \in \mathbb{R}^3$ being realizations of zero-mean white Gaussian noises with covariances ${}^{\beta a}\mathbf{Q} \in \mathbb{R}^{3\times3}$ and ${}^{\beta g}\mathbf{Q} \in \mathbb{R}^{3\times3}$:

$${}^{a}\dot{\boldsymbol{\beta}} = {}^{\beta a}\boldsymbol{w}, \tag{7}$$

$${}^{g}\dot{\boldsymbol{\beta}} = {}^{\beta g}\boldsymbol{w}. \tag{8}$$

Considering Eq. (2-8), one can define the noise vector, along with the non-linear functions $\mathbf{f}(\mathbf{x}, \mathbf{u})$ and $\mathbf{G}(\mathbf{x})$:

$$\mathbf{w} \triangleq \left[{}^{a}\mathbf{w}^{\mathrm{T}}\ {}^{g}\mathbf{w}^{\mathrm{T}}\ {}^{\beta a}\boldsymbol{w}^{\mathrm{T}}\ {}^{\beta g}\boldsymbol{w}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{12}, \tag{9}$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) \triangleq \begin{bmatrix} \mathbf{v}_{\mathrm{G}}^{\mathrm{B/G}} \\ (\mathbf{D}(\alpha)^{\mathrm{B/G}})^{\mathrm{T}}(\check{\mathbf{a}}_{\mathrm{B}} - {}^{a}\boldsymbol{\beta}) + \mathbf{g}_{\mathrm{G}} \\ \mathcal{A}(\boldsymbol{\alpha})(\check{\boldsymbol{\omega}}_{\mathrm{B}} - {}^{g}\boldsymbol{\beta}) \\ \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times1} \end{bmatrix} \in \mathbb{R}^{15}, \tag{10}$$

$$\mathbf{G}(\mathbf{x}) \triangleq \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ -(\mathbf{D}(\alpha)^{\mathrm{B/G}})^{\mathrm{T}} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & -\mathcal{A}(\boldsymbol{\alpha}) & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \end{bmatrix} \in \mathbb{R}^{15\times12}. \tag{11}$$

The two cameras are considered identical and modeled as pinhole (Forsyth and Ponce, 2011), attached to the vehicle's structure as represented in Fig. 1. Geometrically through vector addition, one can notice that for each camera and landmark *i*, the following holds:

$$\vec{\boldsymbol{r}}^{\,\mathrm{B/G}} + \vec{\boldsymbol{r}}^{\,\mathrm{C1/B}} + \vec{\boldsymbol{s}}_1^{\,i} = \vec{\boldsymbol{p}}^{\,i}, \tag{12}$$

$$\vec{\boldsymbol{r}}^{\,\mathrm{B/G}} + \vec{\boldsymbol{r}}^{\,\mathrm{C2/B}} + \vec{\boldsymbol{s}}_2^{\,i} = \vec{\boldsymbol{p}}^{\,i}, \tag{13}$$

where $\vec{p}^{\,i}$ is the position of each landmark i w.r.t. $\mathcal{S}_{\mathrm{G}}$ and $\vec{s}_1^{\,i}$, $\vec{s}_2^{\,i}$ are the true vector from each camera to the considered landmark. The non-linear measurement function $\mathbf{h}^i(\mathbf{x})$ is the representation of Eq. (12-13) in $\mathcal{S}_{\mathrm{I1}}$ and $\mathcal{S}_{\mathrm{I2}}$ respectively:

$$\mathbf{h}^i(\mathbf{x}) \triangleq \begin{bmatrix} \mathbf{s}_{\mathrm{I1}}^i & \mathbf{s}_{\mathrm{I2}}^i \end{bmatrix}^T \in \mathbb{R}^4, \tag{14}$$

where:

$$\mathbf{s}_{\mathrm{I1}}^i = \frac{f}{\mathbf{e}_3^{\mathrm{T}}\mathbf{s}_{\mathrm{C1}}^i} \begin{bmatrix} \mathbf{e}_1^{\mathrm{T}}\mathbf{s}_{\mathrm{C1}}^i \\ \mathbf{e}_2^{\mathrm{T}}\mathbf{s}_{\mathrm{C1}}^i \end{bmatrix} \in \mathbb{R}^2, \tag{15}$$

$$\mathbf{s}_{\mathrm{I2}}^i = \frac{f}{\mathbf{e}_3^{\mathrm{T}}\mathbf{s}_{\mathrm{C2}}^i} \begin{bmatrix} \mathbf{e}_1^{\mathrm{T}}\mathbf{s}_{\mathrm{C2}}^i \\ \mathbf{e}_2^{\mathrm{T}}\mathbf{s}_{\mathrm{C2}}^i \end{bmatrix} \in \mathbb{R}^2, \tag{16}$$

where f is the focal length of the cameras, $e_1 = [1\ 0\ 0]^T$, $e_2 = [0\ 1\ 0]^T$, $e_3 = [0\ 0\ 1]^T$,

$$\mathbf{s}_{\mathrm{C1}}^i = \mathbf{D}^{\mathrm{C1/B}}\mathbf{D}(\alpha)^{\mathrm{B/G}} \left(\mathbf{p}_{\mathrm{G}}^i - \mathbf{r}_{\mathrm{G}}^{\mathrm{B/G}} - \left(\mathbf{D}(\alpha)^{\mathrm{B/G}}\right)^{\mathrm{T}} \mathbf{r}_{\mathrm{B}}^{\mathrm{C1/B}} \right), \tag{17}$$

$$\mathbf{s}_{\mathrm{C2}}^i = \mathbf{D}^{\mathrm{C2/B}}\mathbf{D}(\alpha)^{\mathrm{B/G}} \left(\mathbf{p}_{\mathrm{G}}^i - \mathbf{r}_{\mathrm{G}}^{\mathrm{B/G}} - \left(\mathbf{D}(\alpha)^{\mathrm{B/G}}\right)^{\mathrm{T}} \mathbf{r}_{\mathrm{B}}^{\mathrm{C2/B}} \right). \tag{18}$$

The measurement equation can be now defined as:

$$\mathbf{y}^i = \mathbf{h}^i(\mathbf{x}) + \mathbf{n}^i,\ i = 1, ..., 4, \tag{19}$$

where for each landmark i, $\mathbf{n}^i \in \mathbb{R}^4$ is a noise term, assumed to be a realization of a zero-mean white Gaussian process, with covariance $\mathbf{R}^i \in \mathbb{R}^{4\times4}$.

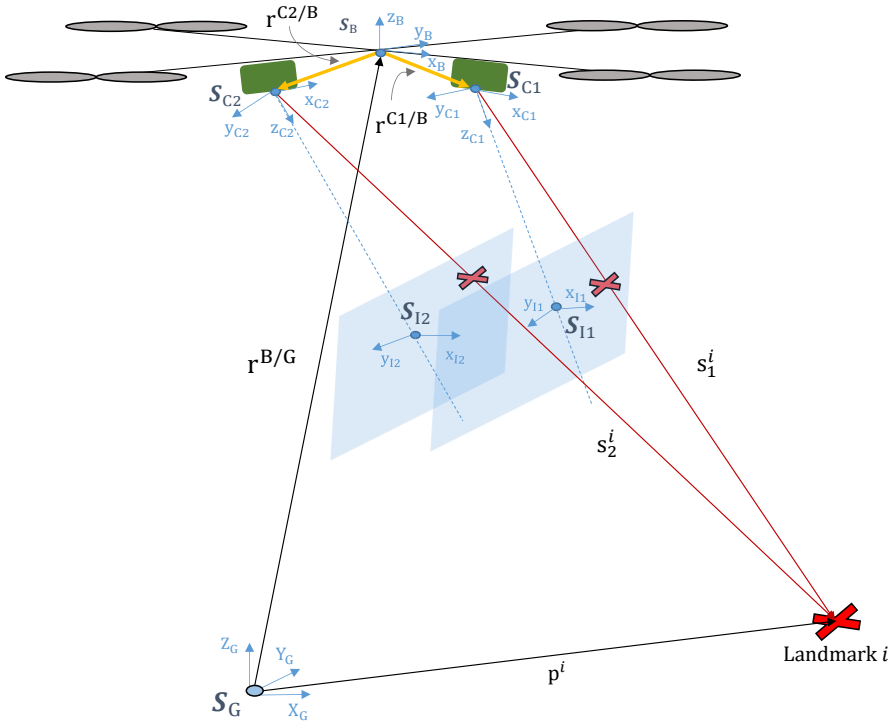

Figure 1. Schematic of the vectorial computation for each landmark.

2.3 Proposed Solutions

The filtering problem consists of inferring the value of the states of a dynamic system at each instant based on indirect, inaccurate or uncertain observations (Bar-Shalom *et al.*, 2004). An optimal solution in the minimum squared error sense for linear and Gaussian systems was proposed by Kalman *et al.* (1960), in an prediction-update scheme. Since then, its framework been incorporated to many of the Kalman-based filters that aim to provide an approximately optimal solution for the nonlinear and non-Gaussian systems present in the physical world. In general, the nonlinear filtering problem consists in finding the conditional probability density of the state, given the observations up to current time (Sorenson and Stubeberud, 1968). All three filters proposed for this work apply the underlying Kalman framework, differing mostly in how the states or their covariance is propagate through the model.

In the Kalman filter scheme, the initial conditions ($\boldsymbol{X}_1$), state and measurements noise covariance matrices ($\boldsymbol{Q}$ and $\boldsymbol{R}$), hereinafter called tuning parameters of the filters, are in accordance with the model presented in Eq. (1) and (19). They are as follows, for each instant k and landmark i: $\boldsymbol{X}_1 \sim (\bar{\boldsymbol{x}}, \bar{\boldsymbol{P}})$, $\boldsymbol{W}_k \sim (\boldsymbol{0}, \boldsymbol{Q})$ and $\boldsymbol{N}_k^i \sim (\boldsymbol{0}, \boldsymbol{R})$.

Therefore the inertial-vision navigation problem goal is to obtain an approximation for the minimum mean squared error (MMSE) estimate of the states ($\mathbf{x}$), using the measurements ($\mathbf{y}^i$, for $i = 1, ..., 4$ landmarks), the inertial measurements ($\mathbf{u}$), the state equation (Eq. 1) and the measurement equation (Eq. 19), employing three different filtering algorithms: extended Kalman filter (EKF), unscented Kalman Filter (UKF) and Ensemble Kalman Filter (EnKF).

As it is well-known, the benchmark for nonlinear filtering is the EKF, which uses a first-order Taylor series expansion of both the state and measurement models. The EKF is detailed in many papers and books (Gelb, 1974; Jazwinski, 2007; Grewal, 2011), being an important basis for comparison, although it may present poor performance in highly non-Gaussian systems. The UKF is based on a deterministic sampling approach: for *n* states, a number of *2n+1* sigma-points are generated and propagated through the true nonlinear model, providing statistic data used for the calculation of the weighted mean and covariance of the system output. With similar computational complexities, the main advantage of the UKF over the EKF is its nonlinear estimation, which means there is no need to calculate the Jacobians of the system. For linear systems, the accuracy of the UKF would be comparable to the third-order approximation of the Taylor series expansion, presenting at least a second-order improvement for non-Gaussian data (Julier and Uhlmann, 1997; Wan and Van Der Merwe, 2000). Albeit it is a not a complex algorithm for implementation, the use of UKF may result in an ill-conditioned problem, when the covariance matrix, in theory positive-definite, becomes negative. This can be mitigated through filtering parameters tuning and some enhancements to the traditional UKF algorithm (Chen, 2003).

Summarized as "a Monte Carlo based implementation of the Kalman filter" by Roth *et al.* (2017), the EnKF propagates an ensemble of *N* state realizations through the nonlinear model, instead of estimating the state covariance. Since the ensemble is composed by random samples and the error is also propagated to avoid underestimating the system uncertainty (Burgers *et al.*, 1998), they tend to reflect the distribution of the true state, even if it is not strictly Gaussian. The EnKF was first proposed by Evensen (1994), targeted to the highly uncertain and non-linear systems present in weather forecast applications, with thousands or millions of states and measurements. Its computational advantage surface when $N < n$, as it is the case for such applications. With the increasing need for the processing of more data and estimation of larger states, such as in problems of simultaneous localization and mapping with numerous landmarks or extended target tracking, the EnKF appears as a relevant and yet not fully explored topic of research, as pointed out by Roth *et al.* (2017).

3. COMPUTATIONAL PROCEDURE AND RESULTS

The truth model is first simulated from arbitrary initial conditions using the MATLAB/Simulink environment. The sensors' measurements are acquired using the truth state subjected to bias and white Gaussian noise according to the parameters of the accelerometer and rate-gyro present in the commercial motion tracking Xsens Mtx device (XsensTechnologies, 2010), a newer version of the one used in Woodman (2007). The information extraction from the cameras is performed according to Eq. (17-18) and also subjected to white Gaussian noise to mimic the USB Web Cam Microsoft LifeCam Cinema. The vehicle is modeled to illustrate a quadrotor, but it can be easily extrapolated to other aerial vehicles. Its basic motion consists of an oscillatory flight around 20 m of altitude in a 5 m radius, where all states vary with time and there are four known landmarks available, as previously mentioned.

The data acquired from simulation of the sensors and camera measurements is processed using the EKF, UKF and EnKF schemes. By performing extensive Monte Carlo simulations using the same tuning parameters and simulated data for all three in order to provide better comparison, the non-linear state estimation problem can be evaluated regarding accuracy, computational performance and robustness with respect to initial conditions. The filters' tuning parameters are shown in Table 1, where *diag* stands for the main diagonal of a diagonal matrix.

To evaluate the accuracy of each filter, we use the average error and root mean square deviation at each instant, as defined respectively in Eq.(20 -21) below, where *MC* is the number of Monte Carlo simulations:

$$\tilde{\boldsymbol{x}}_{avg} \triangleq \frac{1}{MC} \sum_{j=1}^{MC} |\boldsymbol{x}_j - \hat{\boldsymbol{x}}_{\boldsymbol{j}}| \tag{20}$$

$$RMSD \triangleq \sqrt{\frac{1}{MC}\sum_{j=1}^{MC}(\boldsymbol{x}_j - \hat{\boldsymbol{x}}_j)^2} \tag{21}$$

Since all the states (position, velocity, attitude in Euler angles, accelerometer and gyro bias) are three-dimensional vectors of components 1-2-3, we can define their norms as shown in Eq. (22-23), for better visualization.

$$\epsilon \triangleq \sqrt{\tilde{x}_{avg,1}^2 + \tilde{x}_{avg,2}^2 + \tilde{x}_{avg,3}^2} \tag{22}$$

$$\sigma \triangleq \sqrt{RMSD_1^2 + RMSD_2^2 + RMSD_3^2} \tag{23}$$

Table 1. Tuning parameters chosen for EKF, UKF and EnKF

Parameter	**Value**
$\bar{\boldsymbol{x}} \in \mathbb{R}^{15}$	$[0\ 0\ 15\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$
$\bar{\boldsymbol{P}} \in \mathbb{R}^{15\times15}$	$\text{diag}[4.10^{-3}\ 4.10^{-3}\ 8.10^{-3}\ 8.10^{-3}\ 8.10^{-3}\ 8.10^{-3}\ 4.10^{-3}\ 4.10^{-3}\ 4.10^{-3}\ 10^{-3}\ 10^{-3}\ 10^{-4}\ 10^{-4}\ 10^{-4}\ 10^{-4}]^T$
$\boldsymbol{Q} \in \mathbb{R}^{12\times12}$	$\text{diag}[154.10^{-7}\ 154.10^{-7}\ 154.10^{-7}\ 76.10^{-8}\ 76.10^{-8}\ 76.10^{-8}\ 10^{-8}\ 10^{-8}\ 10^{-8}\ 33.10^{-8}\ 33.10^{-8}\ 33.10^{-8}]^T$
$\boldsymbol{R}^i \in \mathbb{R}^{4\times4}, i = 1,2,3,4.$	$\text{diag}[36.10^{-6}\ 36.10^{-6}\ 36.10^{-6}\ 36.10^{-6}]^T$

When comparing the error from the estimate and the truth state in 100 realizations, the EKF presents the smallest average error (ϵ) and standard deviation (σ) in the beginning, but on the long term it is outperformed by the UKF and in some cases by the EnKF, as can be seen in Figures 2 and 3. The EnKF represents a trade-off between accuracy and ensemble size (N), showing in general an acceptable error for $80 \leqslant N \leqslant 100$, but a much higher deviation in all cases - an order above the other two filters (as shown on Table 2).

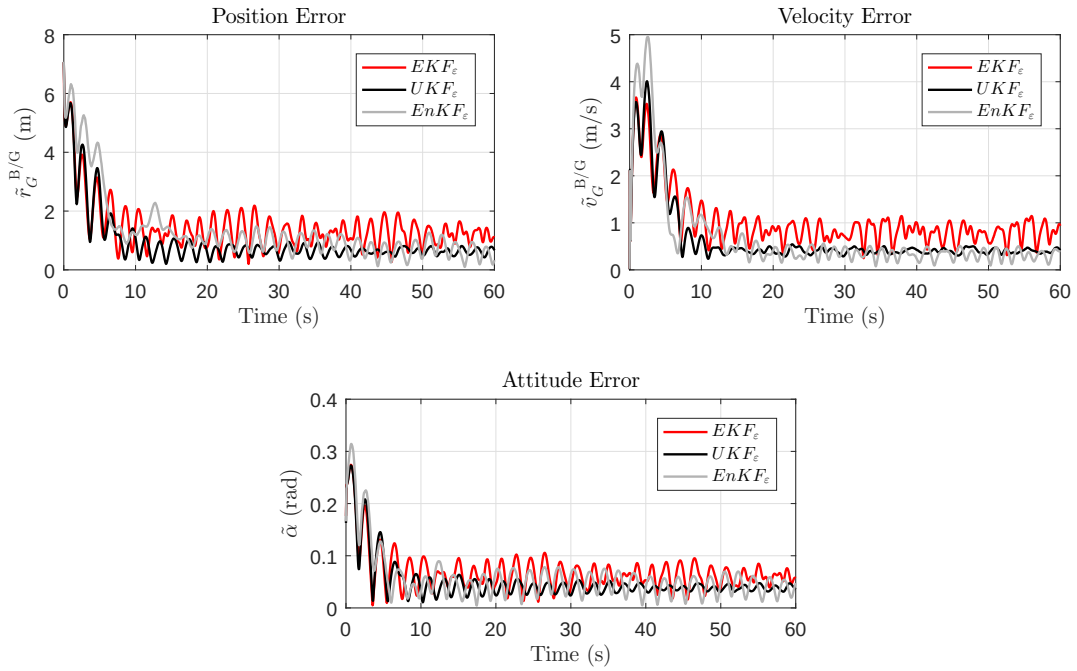

Figure 2. Error performance in position, velocity and attitude estimation using EKF, UKF and EnKF.

In terms of computational burden, as expected the EKF is by far the fastest filter since it does not have to deal with matrix square roots and numerical unscented/ensemble integration of differential equations. Using the EKF as reference, the EnKF with an ensemble size of 100 members performed approximately 3% slower than the UKF.

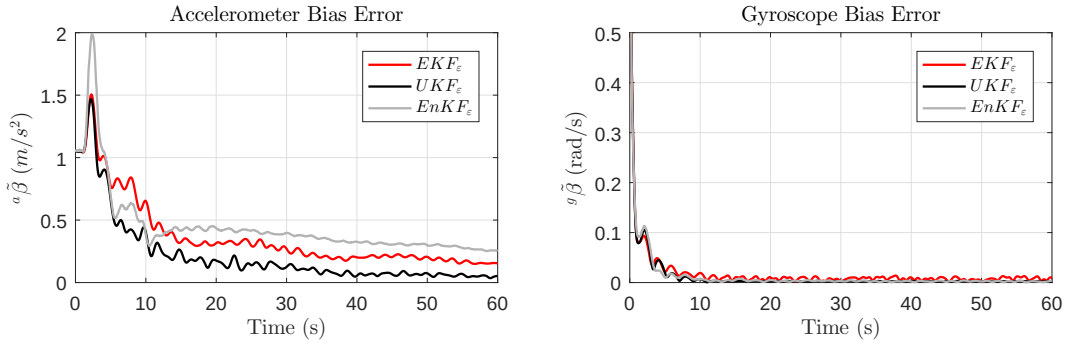

Figure 3. Error performance in bias estimation using EKF, UKF and EnKF

Table 2. Averaged state estimate error and standard deviation after stabilization of the filters

	EKF		**UKF**		**EnFK (N=100)**	
	ϵ	σ	ϵ	σ	ϵ	σ
3D Position [m]	1.2008	0.0785	0.6356	0.0484	0.6743	1.2043
Velocity [m/s]	0.7882	0.0419	0.4011	0.0266	0.3243	0.4098
Attitude [rad]	0.0589	0.0035	0.0385	0.0021	0.0410	0.0527
Accelerometer Bias [m/s^2]	0.1822	0.0269	0.0593	0.0105	0.2880	0.4675
Gyroscope Bias [rad/s]	0.0075	0.0009	0.0025	0.0005	0.0024	0.0069

When analyzing the robustness of each filter, the tuning parameters showed in Table 1 were changed until the point where one or more filters would diverge or present high error in the estimates. In this ad-hoc scenario, the EKF presented higher sensitivity for uncertainty in the initial states and their variance, diverging roughly in 12% of the Monte Carlo realizations, whilst the EnKF and UKF converged in all of them for the same initial conditions. The main difficulty found with the UKF when varying the tuning parameters, was the creation of an ill-conditioned problem as described in Section 2, but since this can be solved numerically, it was not considered as a robustness issue. Furthering the analysis, the ensemble size of the EnKF was also varied, showing that when using an ensemble size of less than 40 it would only diverge on position and attitude estimate.

4. CONCLUSIONS

This paper addressed the vision-aided inertial navigation issue in micro-multirotor aerial vehicles using a simulated environment, in order to investigate three Kalman-based filtering algorithms of interest regarding accuracy, computational burden and robustness with respect to initial conditions. We proposed a non-linear model with fifteen states, comprising an MAVs' three-dimensional position, linear velocity, attitude, as well as the bias of its inertial sensors, a triad of accelerometers and rate-gyros. Based on measurements acquired from the strap-down IMU and stereo camera system, we were able to successfully estimate all fifteen states using the extended Kalman filter, the unscented Kalman filter and the ensemble Kalman filter.

Monte Carlo simulations presented interesting results regarding the performance of the aforementioned filters for the problem proposed. In matters of accuracy, the UKF showed a smaller average error and dispersion in most of the states, reinstating its usefulness in navigation problems and its slight improvement from the EKF in non linear approximations. The traditional EKF remained as the computationally lighter approach amongst all three for the number of states and complexity of the problem, suggesting it might be a better choice for embedded systems with lower computational power. The EnKF however proved to be simpler to implement and more robust with respect to initial conditions than both of the previously discussed filters, which can be crucial in real world navigation problems, where there is a high state uncertainty and disturbances possibly unaccounted for. Moreover, realizing it would be computationally advantageous to propagate an ensemble of sampled points than numerically integrating a high number of differential equations, the EnKF might be a compelling tool for more complex issues, in which the ensemble size is outnumbered by the states.

Further investigation must be conducted to supplement the findings of this work, mainly regarding the application of the ensemble Kalman filter in navigation problems. Initial results were in some aspects auspicious, but also unfolded some challenges in comparison to the current renowned methods, such as the UKF and EKF. Future work is expected to be performed in this area, also concerning experimental evaluation in order to corroborate and extend the results found in simulation.

5. ACKNOWLEDGEMENTS

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

6. REFERENCES

Alsalam, B.H.Y., Morton, K., Campbell, D. and Gonzalez, F., 2017. "Autonomous uav with vision based on-board decision making for remote sensing and precision agriculture". In *Aerospace Conference, 2017 IEEE*. IEEE, pp. 1–12.

Andersen, E.D. and Taylor, C.N., 2007. "Improving mav pose estimation using visual information". In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, pp. 3745–3750.

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.

Blösch, 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*. IEEE, pp. 21–28.

Burgers, G., Jan van Leeuwen, P. and Evensen, G., 1998. "Analysis scheme in the ensemble kalman filter". *Monthly weather review*, Vol. 126, No. 6, pp. 1719–1724.

Carrillo, L.R.G., López, A.E.D., Lozano, R. and Pégard, C., 2012. "Combining stereo vision and inertial navigation system for a quad-rotor uav". *Journal of Intelligent & Robotic Systems*, Vol. 65, No. 1-4, pp. 373–387.

Chen, Z., 2003. "Bayesian filtering: From kalman filters to particle filters, and beyond". *Statistics*, Vol. 182, No. 1, pp. 1–69.

Corke, P., Lobo, J. and Dias, J., 2007. "An introduction to inertial and visual sensing".

Doherty, P. and Rudol, P., 2007. "A uav search and rescue scenario with human body detection and geolocalization". In *Australian Conference on Artificial Intelligence*. Springer, Vol. 4830, pp. 1–13.

Evensen, G., 1994. "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics". *Journal of Geophysical Research: Oceans*, Vol. 99, No. C5, pp. 10143–10162.

Evensen, G., 2003. "The ensemble kalman filter: Theoretical formulation and practical implementation". *Ocean dynamics*, Vol. 53, No. 4, pp. 343–367.

Forster, C., Pizzoli, M. and Scaramuzza, D., 2014. "Svo: Fast semi-direct monocular visual odometry". In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, pp. 15–22.

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

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*. IEEE, pp. 6–pp.

Grewal, M.S., 2011. "Kalman filtering". In *International Encyclopedia of Statistical Science*, Springer, pp. 705–708.

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

Jazwinski, A.H., 2007. *Stochastic processes and filtering theory*. Courier Corporation.

Julier, S.J. and Uhlmann, J.K., 1997. "A new extension of the kalman filter to nonlinear systems". In *Int. symp. aerospace/defense sensing, simul. and controls*. Orlando, FL, Vol. 3, pp. 182–193.

Kalman, R.E. *et al.*, 1960. "A new approach to linear filtering and prediction problems". *Journal of basic Engineering*, Vol. 82, No. 1, pp. 35–45.

Kun, L., 2015. *NAVIGATION AND AUTONOMOUS CONTROL OF MAVS IN GPS-DENIED ENVIRONMENTS*. Ph.D. thesis, National University of Singapore.

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.

Marina, H.G.D., Pereda, F.J., Giron-Sierra, J.M. and Espinosa, F., 2012. "Uav attitude estimation using unscented kalman filter and triad". *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 11, pp. 4465–4474.

Metni, N. and Hamel, T., 2007. "A uav for bridge inspection: Visual servoing control law with orientation limits". *Automation in construction*, Vol. 17, No. 1, pp. 3–10.

Mohammed, F., Idries, A., Mohamed, N., Al-Jaroodi, J. and Jawhar, I., 2014. "Uavs for smart cities: Opportunities and challenges". In *Unmanned Aircraft Systems (ICUAS), 2014 International Conference on*. IEEE, pp. 267–273.

Murray, C.C. and Chu, A.G., 2015. "The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery". *Transportation Research Part C: Emerging Technologies*, Vol. 54, pp. 86–109.

Newcome, L.R., 2004. *Unmanned aviation: a brief history of unmanned aerial vehicles*. Aiaa.

Nikolic, J., Burri, M., Rehder, J., Leutenegger, S., Huerzeler, C. and Siegwart, R., 2013. "A uav system for inspection of industrial facilities". In *Aerospace Conference, 2013 IEEE*. IEEE, pp. 1–8.

Phan, C. and Liu, H.H., 2008. "A cooperative uav/ugv platform for wildfire detection and fighting". In *System Simulation*

and Scientific Computing, 2008. ICSC 2008. Asia Simulation Conference-7th International Conference on. IEEE, pp. 494–498.

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.A. and Gonçalves, P.F., 2016. "Attitude determination of multirotor aerial vehicles using camera vector measurements". *Journal of Intelligent & Robotic Systems*, Vol. 86, No. 1, pp. 139–149.

Scott, J. and Scott, C., 2017. "Drone delivery models for healthcare". In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Sorenson, H.W. and Stubeberud, A.R., 1968. "Non-linear filtering by approximation of the a posteriori density". *International Journal of Control*, Vol. 8, No. 1, pp. 33–51.

Wallar, A., Plaku, E. and Sofge, D.A., 2015. "Reactive motion planning for unmanned aerial surveillance of risk-sensitive areas". *IEEE Transactions on Automation Science and Engineering*, Vol. 12, No. 3, pp. 969–980.

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*. Ieee, pp. 153–158.

Woodman, O.J., 2007. "An introduction to inertial navigation". Technical report, University of Cambridge, Computer Laboratory.

XsensTechnologies, 2010. "Mti and mtx user manual and technical documentation".

Xu, J., Solmaz, G., Rahmatizadeh, R., Turgut, D. and Bölöni, L., 2015. "Animal monitoring with unmanned aerial vehicle-aided wireless sensor networks". In *Local Computer Networks (LCN), 2015 IEEE 40th Conference on*. IEEE, pp. 125–132.

Zhang, C. and Kovacs, J.M., 2012. "The application of small unmanned aerial systems for precision agriculture: a review". *Precision agriculture*, Vol. 13, No. 6, pp. 693–712.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.