A VISUAL-INERTIAL NAVIGATION ALGORITHM FOR MICRO AERIAL VEHICLES USING THE UNSCENTED KALMAN FILTER

João Paulo de A. Barbosa* Raphael Ballet* Davi A. dos Santos* Stiven S. Dias†

*Praça Marechal Eduardo Gomes, 50 - Vila das Acácias - 12.228-900
Instituto Tecnológico de Aeronáutica - ITA
São José dos Campos, São Paulo, Brazil

†Avenida Doutor Altino Bondensan, 500 - Distrito de Eugênio de Melo - 12247-016
Embraer
São José dos Campos, São Paulo, Brazil

Email: barbosajoaopa@gmail.com; raphaelballet@hotmail.com; davists@ita.br; stiven.dias@embraer.com.br

Abstract— This work proposes a position, velocity and heading estimation algorithm based on the unscented Kalman filter (UKF) for a multirotor unmanned aerial vehicle (UAV) using low-cost components, including a gyro-stabilized two-axis gimbal platform containing a downward-facing camera, inertial sensors and an ultrasonic range sensor. In the filter formulation, the inertial sensors together with kinematics equations are used in the prediction step, while landmark vector measurements computed from the camera images and the ultrasonic sensor readings are applied in the update step. This method is evaluated via extensive Monte Carlo simulations, which show its effectiveness. Furthermore, this paper shows a performance comparison between the proposed method and a similar EKF-based method. The proposed method shows a slightly better performance, but has a higher computational burden.

Keywords— Unscented Kalman filter, visual-inertial navigation, sensor fusion, unmanned aerial vehicle.

Resumo— Este trabalho propõe um algoritmo de estimação de posiçãoo, velocidade e guinada, baseado no filtro de Kalman *unscented* (UKF), para um veículo aéreo não-tripulado (VANT) do tipo multirotor utilizando componentes de baixo custo, incluindo uma plataforma giro-estabilizada de dois eixos contendo uma câmera apontada para baixo, sensores inerciais e um sensor ultrassônico. Na formulação do filtro, os sensores inerciais, juntamente com as equações dinâmicas, são usados na etapa de predição, enquanto as medidas dos vetores dos *landmarks*, computados a partir das imagens da câmera, e do sensor ultrassônico são aplicados na etapa de atualização. Esse método é avaliado por simulações Monte Carlo, as quais demonstra sua eficiência. Além disso, esse artigo mostra uma comparação de performance entre o método proposto e um método baseado no EKF. O método proposto demonstra uma performance levemente melhor, mas com maior carga computacional.

Palavras-chave— Filtro de Kalman *unscented*, navegação inercial auxiliada, fusão sensorial, veículo aéreo não-tripulado.

1 Introduction

Micro Aerial Vehicles (MAV) have attracted great attention in recent years due to its wide range of possible applications in both military and civil areas. Special attention is given to autonomous UAV, which represents a multidisciplinary challenge. Some recent applications include terrain mapping for precision agriculture (Primicerio et al., 2012), border surveillance (Beard et al., 2006), building and power line inspection (Metni and Hamel, 2007), and search and rescue (Erdos et al., 2013). All of these applications relies on an accurate navigation algorithm.

Traditional navigation systems were mostly focused on outdoors operations, using the well-known GPS/INS sensor fusion scheme (Farrell and Barth, 2007; St-Pierre and Gingras, 2004). More recent researches have concerned with accurate and robust navigation in GPS-denied environment, such as indoor environments (Bachrach et al., 2010; Carrillo et al., 2012). However, both projects rely on accurate and usually expensive equipments and the authors do not provide a performance analysis of different and newer filtering methods to the chosen set of sensors.

Sensor fusion for navigation is mostly based on the framework of a stochastic optimal state estimator. The Kalman filter (KF)(Kalman, 1960) is the most popular of such estimator, but it is only suitable for linear-Gaussian systems. For nonlinear problems, there exists several adaptations of the original Kalman filter, such as the extended Kalman filter (EKF) (Gelb, 1974), the unscented Kalman filter (UKF) (Julier and Uhlmann, 2004), the cubature Kalman filter (CKF) (Arasaratnam and Haykin, 2009), and others. The EKF method is the most popular estimation filter for navigation applications (Scaramuzza et al., 2014; Bachrach et al., 2010), despite of its limitations in accuracy for highly nonlinear systems compared to UKF (St-Pierre and Gingras, 2004).

The present paper proposes an UKF-based navigation algorithm appropriate for micro aerial vehicles and compares its effectiveness with an EKF-based algorithm. Similar analysis is presented in St-Pierre and Gingras (2004). How-

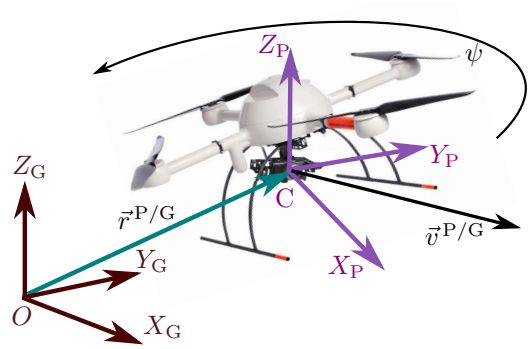


Figure 1: The navigation platform and the Cartesian Coordinate Systems (CCS).

ever, the present project focus on GPS-denied environment and uses low cost sensors, characteristic that also differs this work from those presented in Bachrach et al. (2010) and Carrillo et al. (2012).

The sensors platform consists of a rate-gyro, a three-axis accelerometer, a downward-facing camera and an ultrasonic range finder, all mounted on a two-axis gyro-stabilized gimbal, which maintains the sensors platform plane always parallel to the ground plane. Further, the vehicle is assumed to navigate through an environment with known visible landmarks, that are captured by the camera and fused within the estimation filter. The method is formulated for estimating the vehicle's position, velocity, heading and biases of the inertial sensors.

2 Problem Definition

2.1 Dynamic Model

Figure 1 illustrates a navigation platform and two Cartesian coordinate systems (CCS) for describing the problem. Ground $S_{\mathrm{G}} \triangleq \{X_{\mathrm{G}}, Y_{\mathrm{G}}, Z_{\mathrm{G}}\}$ is fixed to the ground at point O, with Z_{G} aligned with the local vertical. The platform $S_{\mathrm{P}} \triangleq \{X_{\mathrm{P}}, Y_{\mathrm{P}}, Z_{\mathrm{P}}\}$ is fixed on the platform, at point C, with Z_{P} parallel to Z_{G}.

The platform is composed with a two-axis gyro-stabilized gimbal with a three-axis accelerometer aligned with S_{P}, a single-axis rate-gyro, an ultrassonic range sensor, and a downward pointing camera all aligned with Z_{P} axis.

Consider the S_{G} representation of $\vec{r}^{\,\mathrm{P/G}}$ and $\vec{v}^{\,\mathrm{P/G}}$ denoted, respectively, by $\mathbf{r}_{\mathrm{G}}^{\mathrm{P/G}}$ and $\mathbf{v}_{\mathrm{G}}^{\mathrm{P/G}}$. The nonlinear set of equations that describes the dynamic $\mathbf{r}_{\mathrm{G}}^{\mathrm{P/G}}$ and $\mathbf{v}_{\mathrm{G}}^{\mathrm{P/G}}$ are given by

$$\dot{\mathbf{r}}_{\mathrm{G}}^{\mathrm{P/G}} = \mathbf{v}_{\mathrm{G}}^{\mathrm{P/G}} + {}^{r}\mathbf{w} \tag{1}$$

$$\dot{\mathbf{v}}_{\mathrm{G}}^{\mathrm{P/G}} = \left[\mathbf{D}^{\mathrm{P/G}}(\psi)\right]^{\mathrm{T}} (\check{\mathbf{a}}_P - \beta_a) + \mathbf{g}_{\mathrm{G}} + {}^{v}\mathbf{w} \tag{2}$$

$$\dot{\psi} = \check{\omega}_z - \beta_g + {}^{\psi}\mathbf{w} \tag{3}$$

$$\dot{\beta}_g = {}^{g}w \tag{4}$$

$$\dot{\beta}_a = {}^{a}\mathbf{w} \tag{5}$$

where $\{\mathbf{w}\} \in \mathbb{R}^{11}$ is defined as the state-noise vector

$$\mathbf{w} \triangleq \left[({}^{r}\mathbf{w})^{\mathrm{T}} \quad ({}^{v}\mathbf{w})^{\mathrm{T}} \quad {}^{\psi}w \quad {}^{g}w \quad ({}^{a}\mathbf{w})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{11}, \tag{6}$$

assumed to be a zero-mean white Gaussian noise with covariance $\mathbf{Q} \in \mathbb{R}^{11}$; and $\check{\mathbf{a}}$ and $\check{\omega}_z$ denote the measurements taken from the accelerometer and rate-gyro, respectively. The variables β_a and β_g represent the measurements biases, respectively; $\mathbf{g}_{\mathrm{G}} \triangleq \begin{bmatrix}0 & 0 & -g\end{bmatrix}^{\mathrm{T}}$ is the gravity acceleration vector represented in S_{G}, and $D^{\mathrm{P/G}}(\psi)$ is the attitude matrix of S_{G} with respect to S_{P}.

Now, Eq. (1)-(5) can be rewritten in state-space form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) + \mathbf{w} \tag{7}$$

where

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) \triangleq \begin{bmatrix} v_{\mathrm{G,x}}^{\mathrm{P/G}} \\ v_{\mathrm{G,y}}^{\mathrm{P/G}} \\ v_{\mathrm{G,z}}^{\mathrm{P/G}} \\ \mathrm{c}\psi\,(\check{a}_{\mathrm{P,x}} - \beta_{a,x}) - \mathrm{s}\psi\,(\check{a}_{\mathrm{P,y}} - \beta_{a,y}) \\ \mathrm{s}\psi\,(\check{a}_{\mathrm{P,x}} - \beta_{a,x}) + \mathrm{c}\psi\,(\check{a}_{\mathrm{P,y}} - \beta_{a,y}) \\ \check{a}_{\mathrm{P,z}} - \beta_{a,z} - g \\ \check{w}_z - \beta_g \\ \mathbf{0}_{4\times 1} \end{bmatrix} \tag{8}$$

with $\mathrm{c}\psi$ and $\mathrm{s}\psi$ representing $\cos\psi$ and $\sin\psi$, respectively, and the state vector is defined as

$$\mathbf{x} \triangleq \left[\left(\mathbf{r}_{\mathrm{G}}^{\mathrm{G/P}}\right)^{\mathrm{T}} \left(\mathbf{v}_{\mathrm{G}}^{\mathrm{G/P}}\right)^{\mathrm{T}} \ \psi \ \beta_g \ (\beta_a)^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{11}, \tag{9}$$

and the known input vector is given by

$$\mathbf{u} \triangleq \left[(\check{\mathbf{a}}_P)^{\mathrm{T}} \quad \check{\omega}_z\right]^{\mathrm{T}} \in \mathbb{R}^4. \tag{10}$$

2.2 Measurement Equation

Figure 2 shows the navigation system with a pin-hole camera model (Hartley and Zisserman, 2004) and a map of landmarks. We introduce a new CCS, the camera CCS defined as $S_{\mathrm{C}} \triangleq \{X_{\mathrm{C}}, Y_{\mathrm{C}}, Z_{\mathrm{C}}\}$ with origin also coincident with C. Therefore, $\vec{r}^{\,\mathrm{P/G}} = \vec{r}^{\,\mathrm{C/G}}$.

Each landmark is represented in S_{G} as a position vector $\vec{l}^{\,i}$, for $i = 1, 2, ..., m$; where m is the total number of known landmarks in the environment. Similarly, the landmarks can be represented in S_{C} with $\vec{s}^{\,i}$, for $i = 1, 2, ..., m$. Hence, from Figure 2 one could geometrically obtain $\vec{r}^{\,\mathrm{C/G}}$ for the i-th landmark as

$$\vec{r}^{\,\mathrm{C/G}} = \vec{l}^{\,i} - \vec{s}^{\,i} \tag{11}$$

Representing Eq. (11) in S_{G}, we obtain

$$\mathbf{r}_{\mathrm{G}}^{\mathrm{C/G}} = \mathbf{l}_{\mathrm{G}}^{i} - \left[\mathbf{D}^{\mathrm{C/P}}\mathbf{D}^{\mathrm{P/G}}(\psi)\right]^{\mathrm{T}} \mathbf{s}_{\mathrm{C}}^{i} \tag{12}$$

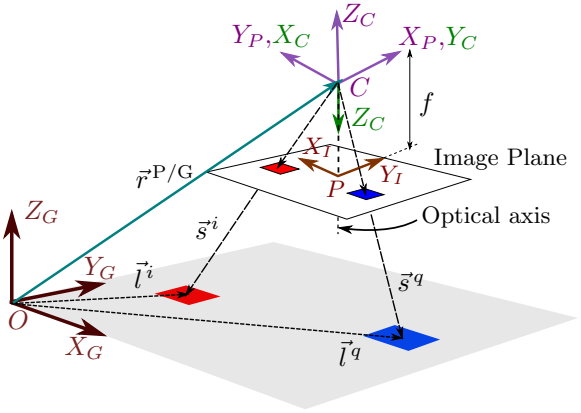


Figure 2: Navigation environment and a pinhole camera model.

Hence

$$\mathbf{s}_{\mathrm{C}}^{i} = \mathbf{D}^{\mathrm{C/P}}\mathbf{D}^{\mathrm{P/G}}\left(\psi\right)\left(\mathbf{l}_{\mathrm{G}}^{i} - \mathbf{r}_{\mathrm{G}}^{\mathrm{C/G}}\right) \tag{13}$$

where $\mathbf{D}^{\mathrm{P/G}}(\psi)$ consists of a rotation by ψ around the Z_{C} axis and the rotation matrix $\mathbf{D}^{\mathrm{C/P}}$ is obtained from a rotation of the frame S_{P} to frame S_{C}.

Furthermore, the general form of measurement model is given by:

$$\mathbf{y}_{k}^{i} = \mathbf{h}^{i}\left(\mathbf{x}_{k}\right) + \mathbf{v}_{k}^{i} \tag{14}$$

where $\{\mathbf{v}_{k}^{i}\} \in \mathbb{R}^{3}$ is the measurement noise vector of the i-th landmark defined as a zero-mean Gaussian white noise, with covariance matrix $\mathbf{R}_{k}^{i}$, and $\mathbf{h}^{i}\left(\mathbf{x}_{k}\right)$ is the measurement equation described in Eq. (13). Therefore,

$$\mathbf{h}^{i}\left(\mathbf{x}_{k}\right) \triangleq \begin{bmatrix} -s\psi\left(l_{x}^{i} - r_{\mathrm{G,x}}^{\mathrm{C/G}}\right) + c\psi\left(l_{y}^{i} - r_{\mathrm{G,y}}^{\mathrm{C/G}}\right) \\ c\psi\left(l_{x}^{i} - r_{\mathrm{G,x}}^{\mathrm{C/G}}\right) + s\psi\left(l_{y}^{i} - r_{\mathrm{G,y}}^{\mathrm{C/G}}\right) \\ -l_{z}^{i} + r_{\mathrm{G,z}}^{\mathrm{C/G}} \end{bmatrix} \tag{15}$$

represents the nonlinear measurement equation defined to an arbitrary i-th landmark. Note that the full measurement equation has a different size depending on the number of visible landmarks measured by the camera at a given sample time. Thus, for q measured landmarks, we obtain

$$\mathbf{y}_{k} = \mathbf{h}\left(\mathbf{x}_{k}\right) + \mathbf{v}_{k}, \tag{16}$$

where $\mathbf{y}_{k} \triangleq \begin{bmatrix}\mathbf{y}_{k}^{1} & \mathbf{y}_{k}^{2} & \dots & \mathbf{y}_{k}^{q}\end{bmatrix}^{\mathrm{T}}$, and the noise vector $\mathbf{v}_{k} \triangleq \begin{bmatrix}\mathbf{v}_{k}^{1} & \mathbf{v}_{k}^{2} & \dots & \mathbf{v}_{k}^{q}\end{bmatrix}^{\mathrm{T}}$, and

$$\mathbf{h}\left(\mathbf{x}_{k}\right) \triangleq \begin{bmatrix} \mathbf{h}^{1}\left(\mathbf{x}_{k}\right) \\ \mathbf{h}^{2}\left(\mathbf{x}_{k}\right) \\ \vdots \\ \mathbf{h}^{q}\left(\mathbf{x}_{k}\right) \end{bmatrix} \tag{17}$$

Problem statement: Given the measurements from the inertial sensors $\mathbf{u}$, a map of landmarks, and the nonlinear kinematic equations, the problem is to obtain the vectors $\mathbf{s}_{\mathrm{C}}^{i}$, for $i = 1, 2, ..., q$, from the image taken from the camera, and accurately estimate the state vector $\mathbf{x}$, thus obtaining a navigation solution for the UAV.

3 Problem Solution

3.1 Camera Model and Range Sensor

In order to compute the vector measurements $s_{\mathrm{C}}^{i} \triangleq \begin{bmatrix}x_{\mathrm{C}}^{i} & y_{\mathrm{C}}^{i} & z_{\mathrm{C}}^{i}\end{bmatrix}^{\mathrm{T}}$, $i = 1, 2, ..., q$, from a 2D image, a pinhole camera model is adopted (Hartley and Zisserman, 2004) (See Figure 2). Define the image coordinate system $S_{\mathrm{I}} \triangleq \{X_{\mathrm{I}}, Y_{\mathrm{I}}\}$, with origin at point P, coincident at the camera's principal point; and X_{I} and Y_{I} are parallel to X_{C} and Y_{C}, respectively.

Denote the projection of s_{C}^{i} onto S_{I} by $s_{\mathrm{I}}^{i} \triangleq \begin{bmatrix}x_{\mathrm{I}}^{i} & y_{\mathrm{I}}^{i}\end{bmatrix}^{\mathrm{T}}$. Thus, one can write

$$x_{\mathrm{C}}^{i} = \frac{z_{\mathrm{C}}^{i}}{f}x_{\mathrm{I}}^{i} \qquad y_{\mathrm{C}}^{i} = \frac{z_{\mathrm{C}}^{i}}{f}y_{\mathrm{I}}^{i}, \tag{18}$$

where x_{C}^{i}, y_{C}^{i} and z_{C}^{i} are the components of $\mathbf{s}_{\mathrm{C}}^{i}$ and f is the camera's focal length, considered constant.

3.2 Unscented Kalman Filter

The estimation algorithm chosen in this paper to resolve the problem stated in section 2.1 is the UKF. The filter was designed to better propagate the probability density of states through the dynamic model. In order to achieve this goal, the Unscented Transform (UT) is used to represent the true mean and covariance of the states by using a well defined set of sample points, called sigma-points (Julier and Uhlmann, 2004). Once defined these sigma-points, it is possible to propagate the probability density through the true nonlinear system and find out the posterior mean and covariance of the states. Wan and Merwe (2002) defines the sigma-points $\mathcal{X}_{i} \in \mathbb{R}^{n}$, for $i = 1, 2, ..., (2n + 1)$, and associated weights $\mathcal{W}_{i} \in \mathbb{R}^{(2n+1)}$ of a random variable $\mathbf{x} \in \mathbb{R}^{n}$, with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{x}$, where $n = 11$ is the state vector's dimension defined in the dynamic model.

Moreover, the dynamic model presented in Section 2 is a continuous-time stochastic process while the measurement model is discrete-time. These characteristics drove the solution to use a derivation of the UKF for continuous-discrete filtering problem. Särkkä (2007) proposes a method for continuous-discrete unscented Kalman filter (CDUKF). The difference to conventional UKF is present in the prediction step of the filter. Mean $\mathbf{m} \in \mathbb{R}^{n}$ and covariance $\mathbf{P} \in \mathbb{R}^{n \times n}$ prediction equations for the CDUKF are presented

by Särkkä (2007) as follow:

$$\begin{aligned}\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} &= \mathbf{f}(\mathcal{X})\mathcal{W}_m \\ \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \mathcal{X}\mathcal{W}\mathbf{f}^{\mathrm{T}}(\mathcal{X}) \\ &\quad + \mathbf{f}(\mathcal{X})\mathcal{W}\mathcal{X}^{\mathrm{T}} \\ &\quad + \mathbf{L}\mathbf{Q}\mathbf{L}^{\mathrm{T}}\end{aligned} \tag{19}$$

where $\mathbf{f}$ is the drift function, defined by Eq. (8) and $\mathbf{Q}$ is the covariance of state noise. $\mathcal{X} \in \mathbb{R}^{n\times(2n+1)}$ represents the matrix of sigma-points defined by the unscented transform in continuous-time application while $\mathcal{W}_m \in \mathbb{R}^{(2n+1)}$ and $\mathcal{W} \in \mathbb{R}^{(2n+1)\times(2n+1)}$ are associated weights used to calculate the posterior mean and covariance in matrix form.

Therefore, the prediction step of CDUKF consists of integrating the differential equations of Eq. (19), considering the integration limits as the conditions from time t_{k-1} until actual time t_k. This leads to prediction of mean $\mathbf{m}_k^-$ and covariance $\mathbf{P}_k^-$ of the states (Särkkä, 2007).

The update step of the CDUKF filter is similar to the conventional discrete UKF algorithm. Thus, with $\mathbf{m}_k^-$ and $\mathbf{P}_k^-$ predicted from previous step, it is possible to calculate the new sigma-points $\mathcal{X}_k^-$, propagate them through measurement equations $\mathbf{h}(\mathcal{X}_k^-)$, then estimate mean μ_k and covariance $\mathbf{S}_k$ of the measurement and cross-covariance $\mathbf{C}_k$ of the state and measurement, as described in Eq. (20) (Särkkä, 2007).

$$\begin{aligned}\mathcal{X}_k^- &= \begin{bmatrix}\mathbf{m}_k^- & \dots & \mathbf{m}_k^-\end{bmatrix} \\ &\quad + \sqrt{c}\begin{bmatrix}0 & \sqrt{\mathbf{P}_k^-} & -\sqrt{\mathbf{P}_k^-}\end{bmatrix} \\ \mathcal{Y}_k^- &= \mathbf{h}(\mathcal{X}_k^-) \\ \mu_k &= \mathcal{Y}_k^-\mathcal{W}_m \\ \mathbf{S}_k &= \mathcal{Y}_k^-\mathcal{W}[\mathcal{Y}_k^-]^{\mathrm{T}} + \mathbf{R}_k \\ \mathbf{C}_k &= \mathcal{X}_k^-\mathcal{W}[\mathcal{Y}_k^-]^{\mathrm{T}}\end{aligned} \tag{20}$$

where $\mathbf{h}$ is the measurement model defined in Eq. (17), R_k is the covariance of measurement noise and c is a UT parameter defined by $c = \alpha^2(n+k)$, and α and k are parameters used to determine sigma-points (Wan and Merwe, 2002). Finally, it is possible to calculate the filter gain $\mathbf{K}_k$ and the updated state mean $\mathbf{m}_k$ and covariance $\mathbf{P}_k$. Matrix equations for this step are defined as follow (Särkkä, 2007):

$$\begin{aligned}\mathbf{K}_k &= \mathbf{C}_k\mathbf{S}_k^{-1} \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k[\mathbf{y}_k - \mu_k] \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T\end{aligned} \tag{21}$$

Therefore, using dynamic and measure models defined in sections 2.1 and 2.2 respectively with the CDUKF-based method described in this section, it is possible to determine the problem states defined in Eq. (9).

4 Simulation and Results

4.1 Computational Simulation

To evaluate the proposed problem, a simulation platform is used. The model is simulated in Simulink with a Runge-Kutta method and sampling time T = 0.1 s. The Simulink 3D Animation simulates the environment containing visible landmarks and dynamics of the sensors. A viewpoint attached to the vehicle simulates the embedded camera and it is fixed to the vehicle translation and rotation about the Z_{P} axis. The simulated vehicle follows a predefined trajectory over the map of landmarks and the simulator returns the real position, the sensors output using Eq. (1)-(5), and the image taken from the simulated camera, that are used in the filtering algorithm. It is important to notice that the trajectory over the map of landmarks is generic, since it does not interfere in the estimation filter.

Simulation environment contains 25 landmarks that are displaced uniformly on a virtual square grid. The side distance of the grid is 0.5 m in both horizontal and vertical direction. Each landmark has its own unique color that is used in a simple processing algorithm in order to track the position of each landmark. The simulated camera has a diagonal field-of-view (FOV) of 0.7854 radian and a focal length $f = 1$ m. Furthermore, the output image has a pixel resolution of 200×300.

The state's noise covariance is tuned in $\mathbf{Q} = \mathrm{diag}\{\sigma_r^2, \sigma_v^2, \sigma_\psi^2, \sigma_g^2, \sigma_a^2\}$ with $\sigma_r^2 = 1.0\times10^{-6}\mathbf{I}_3$ m^2, $\sigma_v^2 = 1.0\times10^{-6}\mathbf{I}_3$ (m/s)^2, $\sigma_\psi^2 = 1.0\times10^{-6}$ rad^2, $\sigma_g^2 = 1.0\times10^{-6}$ (rad/s)^2 and $\sigma_a^2 = 1.0\times10^{-4}\mathbf{I}_3$ (m/s^2)^2. Likewise, the covariance of the measurement noise is tuned in $\mathbf{R}_k^i = 1.0\times10^{-4}\mathbf{I}_3$ m^2. In addition, to test the filter's robustness, it is introduced in the ultrasonic range sensor measurement a white Gaussian noise with variance $\sigma_u^2 = 1.0\times10^{-4}$ m^2, as well as in the measurements of the camera x_{I}^i and y_{I}^i, with variances of $\sigma_{\mathrm{I,x}}^2 = \sigma_{\mathrm{I,y}}^2 = 1.0\times10^{-4}$ m^2, simulating the image processing algorithm errors. The initial conditions are also random variables, with $\mathbf{x}_0 \sim \mathcal{N}\left(\overline{\mathbf{x}}, \overline{\mathbf{P}}\right)$, with $\overline{\mathbf{x}} = \begin{bmatrix}0 & 0 & 1.5 & \mathbf{0}_{1\times8}\end{bmatrix}^{\mathrm{T}}$ and $\overline{\mathbf{P}} = \mathrm{diag}\{\sigma_{0,r}^2, \sigma_{0,v}^2, \sigma_{0,\psi}^2, \sigma_{0,g}^2, \sigma_{0,a}^2\}$, where $\sigma_{0,r}^2 = \mathrm{diag}\{1, 1, 1.0\times10^{-2}\}$ m^2, $\sigma_{0,v}^2 = 1.0\times10^{-4}\mathbf{I}_3$ (m/s)^2, $\sigma_{0,\psi}^2 = (0.25\pi)^2$ rad^2, $\sigma_{0,g}^2 = 1.0\times10^{-6}$ (rad/s)^2 and $\sigma_a^2 = 1.0\times10^{-6}\mathbf{I}_3$ (m/s^2)^2.

In order to define the sigma-points for Unscented Transform, the scaling parameters were defined as $\alpha = 1$, $\beta = 2$ and $k = 0$. These values generated most representative set of sigma-points for the project proposed (Wan and Merwe, 2002). Moreover, in prediction step, as described in Section 3.2, Equations 19 have to be numerically integrated. Thus, as one of the purposes of this work is to compare the obtained result of an UKF-based algorithm with a similar EKF-based one proposed

by Ballet et al. (2016), the integration method chosen were the same for both filter, the Runge-Kutta 4.

4.2 Results

The UKF filter presented in this paper is evaluated via extensive Monte Carlo (MC) simulations in order to demonstrate its performance. For each MC simulation the number of landmarks used in update phase is changed, varying from 1 to 4 landmarks. Moreover, each MC simulation is performed by 100 estimation runs.

The same problem were also solved using EKF (Ballet et al., 2016). A comparison between these two different approaches is relevant to identify the differences and find the most appropriate method for the estimation problem presented. Table 1 shows the results from both, UKF and EKF, methods.

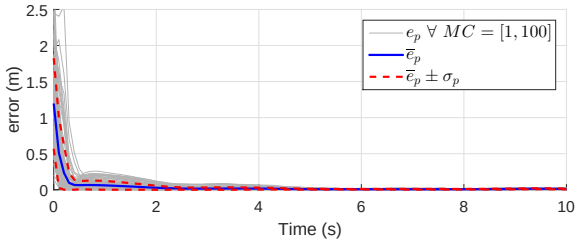
Table 1: Monte Carlo simulations for different values of landmarks used (q).

	q	ε_p	σ_p	ε_v	σ_v	ε_ψ	σ_ψ
UKF	1	8.4	15.7	5.6	5.5	12.3	69.3
	2	5.6	6.5	7.0	23.2	5.2	8.6
	3	5.3	15.3	8.0	78.2	3.8	5.7
	4	4.5	3.4	6.7	22.9	3.3	1.3
EKF	1	9.4	26.2	6.4	9.8	15.7	132.3
	2	5.7	4.6	6.2	11.5	8.1	20.4
	3	5.2	7.7	6.9	26.4	6.2	19.1
	4	4.7	3.1	6.5	15.3	5.3	9.7

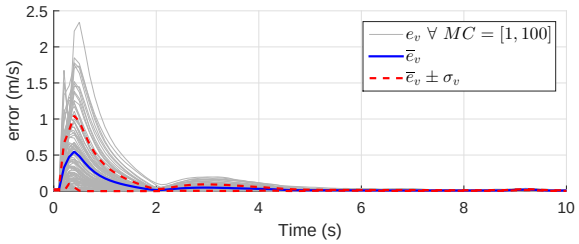
Therefore, when comparing the results presented in Table 1 it is possible to notice that, in terms of states mean estimation error ε_p and standard deviation σ_p from a single MC simulation (100 runs), the UKF-based algorithm presents a better performance for heading and position, but worse for velocity estimation. The reason for this difference is related to a faster position estimation convergence using the UKF model, in which leads to a wider variation in velocity.

Figure 3 shows the temporal response of the position, velocity and heading estimation error for the proposed UKF-based method. Monte Carlo simulations were executed using 4 landmarks. As can be observed, the filter has higher errors between 0 and 2 seconds of simulation. This occurs mostly due to the large initial conditions errors in position and heading. It explains the error spike in Figure 3b at approximately 1 second, where the position convergence of the filter produces a virtual restoring velocity, as the measurement model (Eq. (15)) does not provide information about the vehicle's velocity. Moreover, it can be seen that the filter converges in approximately 2 seconds.

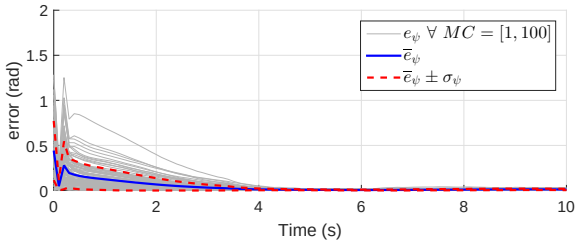
Even though the method presented in this paper presents slightly better estimations (see Table 1), some other aspects should be taken into account. Another important characteristic, that matters for embedding the filter in a real UAV with limited computational resources, is the time for processing the filter operations. Table 2 presents the simulation time of each filtering method and the ratio between them for different number of landmarks used.



(a) Position estimation error.



(b) Velocity estimation error.



(c) Heading estimation Error

Figure 3: Monte Carlo simulations using $q = 4$ landmarks for the UKF method.

Table 2: Average time of one simulation for different values of landmarks used (q).

	Simulation Time [s]		Time Proportion
q	UKF	EKF	UKF/EKF
1	0.5725	0.0539	10.62
2	0.5870	0.0640	9.17
3	0.5982	0.0695	8.61
4	0.6104	0.0748	8.16

As can be seen in Table 2, simulation time for the EKF-based filter is around ten times shorter than the UKF one. In this case, both use a Runge-Kutta 4 method for integrating the dynamic equation in the prediction step, however, the UKF filter has to perform this step for all sigma points (23 sigma-points for the problem presented) in order to propagate the probability density of the prediction step. This is the reason why CDUKF is way slower than the Continuous-Discrete Ex-

tended Kalman Filter (CDEKF).

Therefore, the EKF method seems to be more suitable than UKF for real-time applications in slower and low-cost processors. Although the state error of the UKF during all the estimation period may be slightly smaller, mainly for heading, after the convergence this difference is negligible, meanwhile, the computational effort contrast drives the EKF-based filter to be, in overall, a more suitable solution for the estimation problem studied in this project. Similar conclusions can be found in St-Pierre and Gingras (2004).

5 Conclusion

The present paper has proposed an unscented Kalman filter formulation for accurate and robust estimation of position, velocity, and heading, as well as the inertial sensors' bias, of a micro aerial vehicle using a downward-facing camera, inertial and ultrasonic sensors. Thus, we claim that it is suitable for operation in GPS-denied environments and can be performed using low-cost equipments in a real experiment application.

Results show that the method can successfully estimate the desired navigation. It is sufficiently accurate for most applications and shows a slight better performance compared to a similar EKF-based method. On the other hand, the UKF-based method requires a significant higher computational effort, which makes it less attractive than the EKF in practical applications.

References

Arasaratnam, I. and Haykin, S. (2009). Cubature kalman filters, *Automatic Control, IEEE Transactions on* **54**(6): 1254–1269.

Bachrach, A., De Winter, A., He, R., Hemann, G., Prentice, S. and Roy, N. (2010). RANGE - Robust autonomous navigation in GPS-denied environments, *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 1096–1097.

Ballet, R., Santos, D. A. and Zoran, S. (2016). Position, Velocity And Heading Estimation For Unmanned Aerial Vehicles Using Camera And Inertial Sensors, *Congresso Brasileiro de Engenharia Mecânica (CONEM)*.

Beard, R. W., McLain, T. W., Nelson, D. B., Kingston, D. and Johanson, D. (2006). Decentralized cooperative aerial surveillance using fixed-wing miniature UAVs, *Proceedings of the IEEE* **94**(7): 1306–1323.

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 and Robotic Systems: Theory and Applications* **65**(1-4): 373–387.

Erdos, D., Erdos, A. and Watkins, S. E. (2013). An experimental UAV system for search and rescue challenge, *IEEE Aerospace and Electronic Systems Magazine* **28**: 32–37.

Farrell, J. A. and Barth, M. (2007). *The Global Positioning System & Inertial Navigation*, 2nd edn, Wiley-Interscience.

Gelb, A. (1974). *Applied Optimal Estimation*, Vol. 64, The MIT Press.

Hartley, R. and Zisserman, A. (2004). *Multiple view geometry in computer vision*, Vol. 1.

Julier, S. and Uhlmann, J. (2004). Unscented Filtering and Non Linear Estimation, *Proceedings of the IEEE* **92**(3): 401–422.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems.

Metni, N. and Hamel, T. (2007). A UAV for bridge inspection: Visual servoing control law with orientation limits, *Automation in Construction* **17**(1): 3–10.

Primicerio, J., Di Gennaro, S. F., Fiorillo, E., Genesio, L., Lugato, E., Matese, A. and Vaccari, F. P. (2012). A flexible unmanned aerial vehicle for precision agriculture, *Precision Agriculture* **13**(4): 517–523.

Särkkä, S. (2007). On unscented Kalman filtering for state estimation of continuous-time nonlinear systems, *IEEE Transactions on Automatic Control* **52**(9): 1631–1641.

Scaramuzza, D., Achtelik, M. C., Doitsidis, L., Friedrich, F., Kosmatopoulos, E., Martinelli, A., Achtelik, M. W. et al. (2014). Vision-controlled micro flying robots: From system design to autonomous navigation and mapping in GPS-denied environments, *IEEE Robotics and Automation Magazine* **21**(3): 26–40.

St-Pierre, M. and Gingras, D. (2004). Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated navigation information system, *IEEE Intelligent Vehicles Symposium, 2004* pp. 831–835.

Wan, E. A. and Merwe, R. v. d. (2002). The unscented Kalman filter for nonlinear estimation, *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000* pp. 153–158.