CONEM 2016
CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
21 - 25 AGOSTO DE 2016
FORTALEZA - CEARÁ

POSITION, VELOCITY AND HEADING ESTIMATION FOR UNMANNED AERIAL VEHICLES USING CAMERA AND INERTIAL SENSORS

Raphael Ballet, raphaelballet@hotmail.com[1]
Davi Antônio dos Santos, davists@ita.br[1]
Sjanic Zoran, Zoran.Sjanic@saabgroup.com[2]

[1]Instituto Tecnológico de Aeronáutica, Praça Marechal Eduardo Gomes, 50 - Vila das Acácias - CEP 12.228-900 – São José dos Campos – SP – Brasil
[2]Linköping University, Linköping University SE-581 83 - LINKÖPING - Sweden

Abstract: *Precise and robust navigation of Unmanned Aerial Vehicles (UAV) in GNSS-denied environments is a relevant research topic, especially for autonomous systems. Most of the proposed solutions to this problem rely on expensive and high-precision equipment, thus limiting the accessibility of this technology for research. This work proposes a position, velocity and heading estimation algorithm for multirotor UAV using low-cost components, such as a gyro-stabilized platform containing a downward-facing onboard camera, inertial sensors and an ultrasonic range sensor, using a two-axis gimbal. The estimation algorithm is composed of three steps. At first, the vehicle must detect and identify visible landmarks by processing the images. Second, the algorithm computes the vectors from the camera center to each of the visible landmarks. Finally, measurements from the inertial sensors, ultrasonic range sensor and the landmark's vector measurements are fused with an Extended Kalman Filter (EKF) to estimate the vehicle's position, velocity and heading as well as accelerometer and rate-gyro biases. A map of landmarks is assumed available and the vehicle is supposed to operate indoors. The method is evaluated via extensive Monte Carlo simulations, which show its effectiveness and some of its properties.*
Keywords: *Extended Kalman filter, Visual-inertial navigation, Sensor fusion, Unmanned Aerial Vehicle.*

1 INTRODUCTION

Unmanned Aerial Vehicles (UAV) 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 surveillance (Hoffmann *et al.*, 2008), and search and rescue (Erdos *et al.*, 2013).

The traditional navigation systems are was mostly focused on outdoors operations, using the well-known GPS/INS sensor fusion scheme (Erdos *et al.*, 2013; Wendel *et al.*, 2006). More recent researches in this subject have concerned with precise and robust navigation in GNSS-denied environment, such as indoor environments (Gini and Marchi, 2002; Carrillo *et al.*, 2012). In Gini and Marchi (2002), the authors use a single camera on the top of a mobile robot to track its position based on tracked objects on the floor. In Trawny *et al.* (2007), the authors fuse measurements from a gimbaled inertial sensors with a camera and using an EKF to estimate the position and velocity of an aircraft with validation on real data. The authors rely on mapped landmarks, such as craters or other visual features. Moreover, in Carrillo *et al.* (2012) a stereo visual odometry and inertial measurements are fused by a Kalman Filter (KF), in order to a multirotor vehicle autonomously navigate through an indoor environment.

Sensor fusion for navigation is mostly based on the framework of a stochastic optimal state estimator. The Kalman filter (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 to the original Kalman filter, such as the extended Kalman filter (EKF) (Gelb, 1974), unscented Kalman filter (UKF) (Zhang *et al.*, 2015), cubature Kalman filter (CKF) (Arasaratnam and Haykin, 2009), and the ensemble Kalman filter (EnKF) (Evensen, 2003).

The present paper proposes a EKF-based navigation algorithm appropriate for low-cost micro aerial vehicles. It uses measurements taken from a rate-gyro, an accelerometer, a downward-facing camera, and an ultrasonic range sensor, all mounted on a two-axis gyro-stabilized gimbal. The method is formulated for estimating the vehicle's position, velocity, heading, and biases of the inertial sensors. The vehicle navigates through an environment with known visible landmarks,

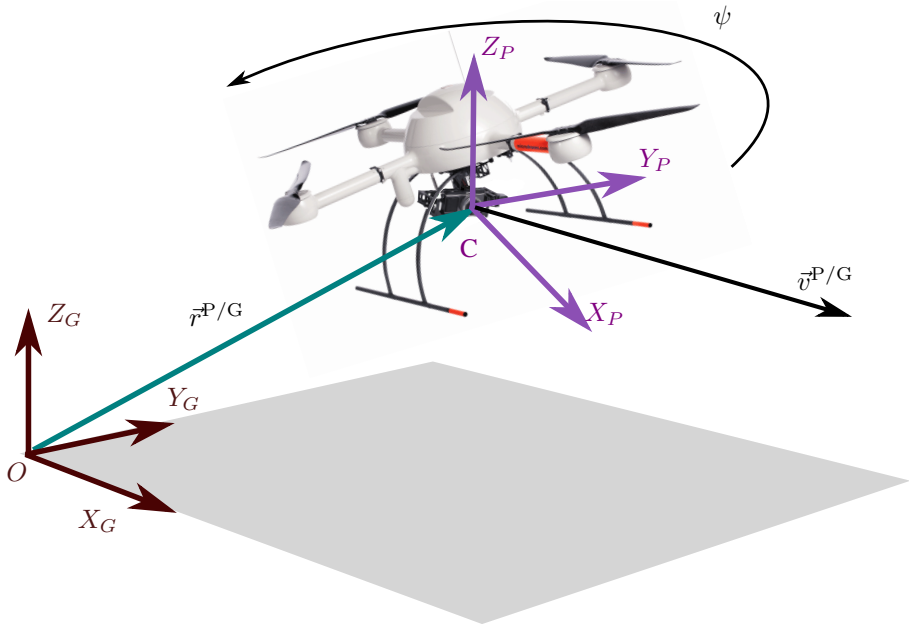


Figure 1: The navigation platform and the Cartesian Coordinate Systems (CCS).

that are captured by the camera, thus providing a position, velocity and heading estimation of the UAV's platform. This estimation is fused with the inertial sensors measurements in an Extended Kalman Filter algorithm, thereby obtaining a preciser estimate of the desired states, as well as accelerometer and rate-gyro biases.

The remaining text is organized in the following manner. Section 2 defines the navigation problem. Section 3 presents a solution for the indicated problem and devise the used EKF algorithm. Section 4 evaluates the proposed method by Monte Carlo simulations, where it gives considerations regarding the relationship between the precision of the estimator and the number of landmarks used. Finally, Section 5 presents the final conclusions.

2 PROBLEM DEFINITION

This section defines a vision-inertial navigation problem suitable for UAVs. To begin with, Section 2.1 presents the kinematics equations. Section 2.2 specifies the measurement equations. Finally, Section 3.1 provides a camera model for determining a three-dimensional point from a two-dimentional image and an ultrassonic range finder.

2.1 Dynamic Model

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

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

The present work aims at providing a precise and robust solution to the position $\vec{r}^{\mathrm{P/G}}$, velocity $\vec{r}^{\mathrm{P/G}}$ and heading angle ψ of an UAV.

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

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

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

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

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

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

where $\mathbf{w}$ is defined as the state-noise vector (Eq. (11)), $\breve{\mathbf{a}}$ and $\breve{\omega}_z$ denote the measurements taken from the accelerometer and rate-gyro, respectively. The variables β_a and β_g represent the biases of the measurements, respectively; $\mathbf{g}_G \triangleq \begin{bmatrix}0 & 0 & -g\end{bmatrix}^{\mathrm{T}}$ is the gravity acceleration vector represented in S_{G}, and

$$\mathbf{D}^{\mathrm{P/G}}(\psi) = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

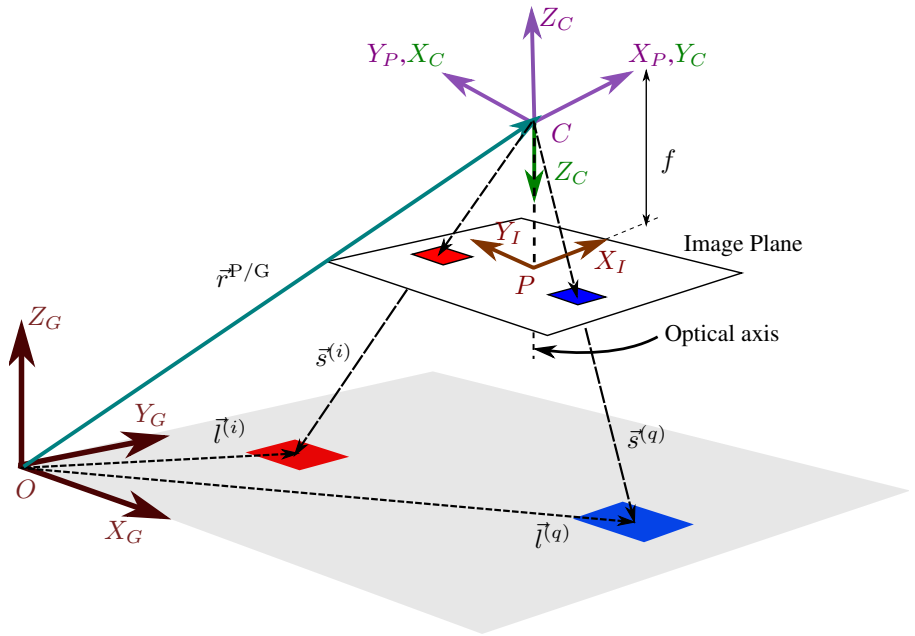


Figure 2: Navigation environment and a pinhole camera model.

is the attitude matrix of S_{G} with respect to S_{P}.

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

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

where

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

with

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

the known input vector defined as

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

and a state noise vector

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

The noise is assumed to be a zero-mean white Gaussian noise with covariance $\mathbf{Q} \in \mathbb{R}^{11}$. For short, $\mathrm{c}\psi$ and $\mathrm{s}\psi$ represents $\cos\psi$ and $\sin\psi$, respectively.

2.2 Measurement Equation

Figure 2 shows the navigation system with a pin-hole camera model (Ponce and Forsyth, 2012; Hartley and Zisserman, 2015) and a map of landmarks. We introduce a new CCS, the camera Cartesian Coordinate System denoted by $S_{\mathrm{C}} = \{X_{\mathrm{C}}, Y_{\mathrm{C}}, Z_{\mathrm{C}}\}$ 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 $i = 1, 2, ..., n$; where n is the total number of known landmarks in the environment. Similarly, the landmarks can be represented in S_{C} with $\vec{s}^{(i)}$, for $i = 1, 2, ..., n$. Hence, from Figure 2 one could geometrically obtain $\vec{r}^{\,\mathrm{C/G}}$ for the i-th landmark as

$$\vec{r}^{\,\mathrm{C/G}} = \vec{l}^{(i)} - \vec{s}^{(i)} \tag{12}$$

Representing Eq. (12) in S_{G}, we obtain

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

Hence

$$\mathbf{s}_C^{(i)} = \mathbf{D}^{\mathrm{C/P}}\mathbf{D}^{\mathrm{P/G}}(\psi)\left(\mathbf{l}_G^{(i)} - \mathbf{r}_G^{\mathrm{C/G}}\right) \tag{14}$$

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} as follows

$$\mathbf{D}^{\mathrm{C/P}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{15}$$

Using the definition of Eq. (6) and replacing by its measurements at time $t_k = kT$, with T representing the measurement sampling interval, it is possible to obtain

$$\check{\mathbf{s}}_{\mathrm{C}}^{(i)} = \begin{bmatrix} -\sin\psi & \cos\psi & 0 \\ \cos\psi & \sin\psi & 0 \\ 0 & 0 & -1 \end{bmatrix} \left(\mathbf{l}_{\mathrm{G}}^{(i)} - \mathbf{r}_{\mathrm{G}}^{\mathrm{C/G}}\right) + \mathbf{v}_k^{(i)} \tag{16}$$

where $\{\mathbf{v}^{(i)}\} \in \mathbb{R}^3$ is the measurement noise vector of the i-th landmark defined as a zero-mean Gaussian white noise, with covariance matrix $\mathbf{R}_k^{(i)}$. Furthermore, Eq. (16) is rewritten in a general form

$$\mathbf{y}_k^{(i)} = \mathbf{h}^{(i)}(\mathbf{x}_k) + \mathbf{v}_k^{(i)} \tag{17}$$

From Eq. (16) and Eq. (17), we obtain

$$\mathbf{y}_k^{(i)} \triangleq \mathbf{s}_{\mathrm{C}}^{(i)} \tag{18}$$

and

$$\mathbf{h}^{(i)}(\mathbf{x}_k) \triangleq \begin{bmatrix} -\sin\psi\left(l_x^{(i)} - r_{G,x}^{\mathrm{C/G}}\right) + \cos\psi\left(l_y^{(i)} - r_{G,y}^{\mathrm{C/G}}\right) \\ \cos\psi\left(l_x^{(i)} - r_{G,x}^{\mathrm{C/G}}\right) + \sin\psi\left(l_y^{(i)} - r_{G,y}^{\mathrm{C/G}}\right) \\ -l_z^{(i)} + r_{\mathrm{G},z}^{\mathrm{C/G}} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_k} \tag{19}$$

Equation (19) 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. Therefore, if we have q measured landmarks, the full measurement equation can be obtained as

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \tag{20}$$

where $\mathbf{y}_k \triangleq \begin{bmatrix} \mathbf{y}_k^{(1)} & \mathbf{y}_k^{(2)} & \dots & \mathbf{y}_k^{(q)} \end{bmatrix}^{\mathrm{T}}$, $\mathbf{v}_k \triangleq \begin{bmatrix} \mathbf{v}_k^{(1)} & \mathbf{v}_k^{(2)} & \dots & \mathbf{v}_k^{(q)} \end{bmatrix}^{\mathrm{T}}$ and

$$\mathbf{h}(\mathbf{x}_k) \triangleq \begin{bmatrix} \mathbf{h}^{(1)}(\mathbf{x}_k) \\ \mathbf{h}^{(2)}(\mathbf{x}_k) \\ \vdots \\ \mathbf{h}^{(q)}(\mathbf{x}_k) \end{bmatrix} \tag{21}$$

3 Problem Solution

This section aims to explain how the problem stated in Section 2 is solved. Therefore, Section 3.1 shows how to compute vector measurements from the image processing and the ultrasonic sensor data and Section **??** presents a review of Unscented Kalman Filter estimator and highlights some considerations for its implementation in the exposed problem.

3.1 Camera Model and Range Sensor

In order to compute the vector measurements $s_C^i = [x_c^i y_c^i z_c^i]^{\mathrm{T}}$, $i = 1, 2, ..., q$, a pinhole camera model is adopted (Hartley and Zisserman, 2015). See Figure 2. Define the image coordinate system $S_{\mathrm{I}} = \{X_{\mathrm{I}}, Y_{\mathrm{I}}\}$, with origin at point P and X_{I} and Y_{I} are parallel with X_{C} and Y_{C}, respectively. Denote the projection of s_{C}^i onto s_{I}^i by $s_I^i = \left[x_{\mathrm{I}}^i y_{\mathrm{I}}^i\right]^{\mathrm{T}}$. Thus, one can write

$$\check{x}_{\mathrm{C}}^{(i)} = \frac{\check{z}_{\mathrm{C}}^{(i)}}{f}\check{x}_{\mathrm{I}}^{(i)} \qquad\qquad \check{y}_{\mathrm{C}}^{(i)} = \frac{\check{z}_{\mathrm{C}}^{(i)}}{f}\check{y}_{\mathrm{I}}^{(i)}, \tag{22}$$

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

3.2 Extended Kalman Filter

Consider the dynamic's state-space given by Eq. (7) in discrete time

$$\mathbf{x}_{k+1} = \mathbf{f}\left(\mathbf{x}_k, \mathbf{u}_k\right) + \mathbf{w}_k \tag{23}$$

and its measurement equation given by Eq. (20)

$$\mathbf{y}_{k+1} = \mathbf{h}\left(\mathbf{x}_k\right) + \mathbf{v}_{k+1} \tag{24}$$

where $\{\mathbf{w}_k\} \in \mathbb{R}^{11}$ and $\{\mathbf{v}_{k+1}\} \in \mathbb{R}^{3q}$ are assumed additive uncorrelated zero-mean Gaussian white-noises with covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_{k+1}$, respectively.

The main objective of the Extended Kalman Filter is to obtain an estimate $\hat{\mathbf{x}}_{k|k}$ of the system state $\mathbf{x}_k$ that minimizes the following cost function

$$J_k \triangleq \mathbb{E}\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right)^{\mathrm{T}}\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right)\right], \forall k \tag{25}$$

where $\mathbb{E}$ denotes the expectation operator and $\hat{\mathbf{x}}_{k|k} \triangleq \mathbb{E}\left[\mathbf{x}_k|\mathbf{y_k}\right] \in \mathbb{R}^{11}$, for $\mathbb{E}\left[a|b\right]$ representing the expectation of a given the information of b. Furthermore, the state error covariance matrix is obtained by $\mathbf{P}_k \triangleq \mathbb{E}\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right)^{\mathrm{T}}\right] \in \mathbb{R}^{11\times 11}$. The initial condition $\mathbf{x}_0$ is also assumed a Gaussian random variable with mean and correlation matrix defined by $\overline{\mathbf{x}}$ and $\overline{\mathbf{P}}$, respectively; and uncorrelated to $\mathbf{w}_k$ and $\mathbf{v}_k$.

The Extended Kalman Filter consists of two main steps. The first step is defined as the propagation step and the goal is to find a predicted state estimate $\hat{\mathbf{x}}_{k+1|k} \triangleq \mathbb{E}\left[\mathbf{x}_{k+1}|\mathbf{y_k}\right]$, as well as its predicted state error covariance matrix $\mathbf{P}_{k+1} \triangleq \mathbb{E}\left[\left(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k}\right)\left(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k}\right)^{\mathrm{T}}\right]$, and a predicted measurement estimate $\hat{\mathbf{y}}_{k+1|k} \triangleq \mathbb{E}\left[\mathbf{y}_{k+1}|\mathbf{y_k}\right]$ using the dynamic and measurement equations. Assuming that the functions given by Eq. (23)-(24) are sufficiently smooth, the nonlinear functions $\mathbf{f}\left(\mathbf{x}_k, \mathbf{u}_k\right)$ and $\mathbf{h}\left(\mathbf{x}_k\right)$ can be linearized by expanding a first-order Taylor series about the estimate $\hat{\mathbf{x}}_{k|k}$ (Anderson *et al.*, 1982)

$$\mathbf{f}\left(\mathbf{x}_{k|k}, \mathbf{u}_k\right) \approx \mathbf{f}\left(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k\right) + \mathbf{F}_k\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right) \tag{26}$$

$$\mathbf{h}\left(\mathbf{x}_{k|k}\right) \approx \mathbf{h}\left(\hat{\mathbf{x}}_{k|k}\right) + \mathbf{H}_{k+1}\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right) \tag{27}$$

where

$$\mathbf{F}_k \triangleq \left[\frac{\partial}{\partial \mathbf{x}}\mathbf{f}\left(\mathbf{x}, \mathbf{u}_k\right)\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} \tag{28}$$

and

$$\mathbf{H}_{k+1} \triangleq \left[\frac{\partial}{\partial \mathbf{x}}\mathbf{h}\left(\mathbf{x}\right)\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k+1|k}} \tag{29}$$

are the Jacobian matrices of $\mathbf{f}\left(\mathbf{x}_{k|k}, \mathbf{u}_k\right)$ and $\mathbf{h}\left(\mathbf{x}_{k|k}\right)$, respectively. Therefore, it is possible to obtain the predicted estimate using a structure similar to the linear Kalman Filter as follows:

$$\hat{\mathbf{x}}_{k+1|k} \approx \mathbf{f}\left(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k\right) \tag{30}$$

$$\hat{\mathbf{y}}_{k+1|k} \approx \mathbf{h}\left(\hat{\mathbf{x}}_{k|k}\right) \tag{31}$$

Furthermore, the predicted state covariance matrix, also called the residual covariance, is obtained by the so-called discrete-time Riccati equation (Anderson *et al.*, 1982)

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k\mathbf{P}_{k|k}\mathbf{F}_k^{\mathrm{T}} + \mathbf{Q}_k \tag{32}$$

The second step of the EKF is defined as the update step. This step uses the measurement y_{k+1} at time t_{k+1} to update the predicted state estimate $\hat{\mathbf{x}}_{k+1|k}$ and state covariance $\mathbf{P}_{k+1|k}$, thereby obtaining the desired estimates $\hat{\mathbf{x}}_{k+1|k+1}$ and $\mathbf{P}_{k+1|k+1}$ with

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1|k}\right) \tag{33}$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\mathbf{P}_{k+1|k}, \tag{34}$$

where $\mathbf{K}_{k+1}$ is referred as the Kalman gain matrix, which is given by

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^{\mathrm{T}}\left(\mathbf{H}_{k+1}\mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^{\mathrm{T}} + \mathbf{R}_{k+1}\right)^{-1} \tag{35}$$

The previous EKF formulation is called the discrete EKF, however, in this work a continuous-discrete formulation of the EKF is desired. Consider the continuous state-space given by Eq. (7). The respective Taylor series approximation about $\hat{\mathbf{x}}_{k|k}$ and its state covariance matrix are given by (Gelb, 1974)

$$\dot{\mathbf{x}} = \mathbf{f}\left(\hat{\mathbf{x}}, \mathbf{u}\right) \tag{36}$$

$$\dot{\mathbf{P}} = \mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^{\mathrm{T}} + \mathbf{Q} \tag{37}$$

where from Eq. (8) we obtain

$$\mathbf{F} = \left[\frac{\partial}{\partial \mathbf{x}}\mathbf{f}\left(\mathbf{x}, \mathbf{u}\right)\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{I}_3 & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{B} & \mathbf{0}_{3\times1} & -\left(\mathbf{D}^{\mathrm{P/G}}\left(\psi\right)\right)^{\mathrm{T}} \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & 0 & -1 & \mathbf{0}_{1\times3} \\ \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times3} \end{bmatrix}_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} \tag{38}$$

where $\mathbf{I}_3$ is the 3×3 identity matrix and

$$\mathbf{B} = \begin{bmatrix} -c\psi\left(\breve{a}_{\mathrm{P,y}} - \beta_{a,y}\right) - s\psi\left(\breve{a}_{\mathrm{P,x}} - \beta_{a,x}\right) \\ c\psi\left(\breve{a}_{\mathrm{P,x}} - \beta_{a,x}\right) - s\psi\left(\breve{a}_{\mathrm{P,y}} - \beta_{a,y}\right) \\ 0 \end{bmatrix}$$

Similarly, for a i-th landmark it is possible to compute $\mathbf{H}_{k+1}^{(i)}$ from Eq. (28) using Eq. (17)-(19)

$$\mathbf{H}_{k+1}^{(i)} = \left[\frac{\partial}{\partial \mathbf{x}}\mathbf{h}^{(i)}\left(\mathbf{x}_k\right)\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k+1|k}} = \begin{bmatrix} -\left(\mathbf{D}^{\mathrm{C/G}}\left(\psi\right)\right)^{\mathrm{T}} & \mathbf{0}_{3\times3} & \mathbf{C} & \mathbf{0}_{3\times4} \end{bmatrix}_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}} \tag{39}$$

with

$$\mathbf{C} = \begin{bmatrix} -c\psi\left(l_x^{(i)} - r_{\mathrm{G,x}}^{\mathrm{C/G}}\right) - s\psi\left(l_y^{(i)} - r_{\mathrm{G,y}}^{\mathrm{C/G}}\right) \\ c\psi\left(l_y^{(i)} - r_{\mathrm{G,y}}^{\mathrm{C/G}}\right) - s\psi\left(l_x^{(i)} - r_{\mathrm{G,x}}^{\mathrm{C/G}}\right) \\ 0 \end{bmatrix}$$

The complete measurement Jacobian matrix is thus given by

$$\mathbf{H}_{k+1} = \begin{bmatrix} \mathbf{H}_{k+1}^{(1)} \\ \mathbf{H}_{k+1}^{(2)} \\ \vdots \\ \mathbf{H}_{k+1}^{(q)} \end{bmatrix}. \tag{40}$$

Note that the computational complexity of the Extended Kalman Filter algorithm is approximately $O\left(q^3\right)$ (Thrun, 2002), that is, the complexity grows in a cubic fashion proportionally to the number of processed landmarks. Thus, the number of used landmarks in the EKF algorithm is a trade-off between accuracy, as greater values of q leads to a preciser estimation, and computational efficiency, that could, depending on the available Hardware/Software, degrade the capability of estimating the desired states in real-time.

Finally, the complete Continuous-Discrete EKF algorithm is depicted next in Algorithm 1:

Algorithm 1 Continuous-Discrete Extended Kalman Filter

1: **Initial conditions**: $k = 0, \hat{\mathbf{x}}_{0|0} = \overline{\mathbf{x}}, \mathbf{P}_{0|0} = \overline{\mathbf{P}}$
2: **loop**
3: Prediction:
4: $\hat{\mathbf{x}}_{k+1|k} = \text{integrate}\left(\mathbf{f}\left(\hat{x}, \mathbf{u}\right)\right)$ from t_k to t_{k+1}, with initial conditions $\hat{\mathbf{x}}_{0|0}$
5: $\mathbf{P}_{k+1|k} = \text{integrate}\left(\mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^{\mathrm{T}} + \mathbf{Q}\right)$ from t_k to t_{k+1}, with initial conditions $\mathbf{P}_{0|0}$
6: Update:
7: $\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^{\mathrm{T}}\left(\mathbf{H}_{k+1}\mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^{\mathrm{T}} + \mathbf{R}_{k+1}\right)^{-1}$
8: $\hat{\mathbf{x}}_{k+1|k+1} = \hat{x}_{k+1|k} + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \mathbf{h}\left(\hat{\mathbf{x}}_{k+1|k}\right)\right)$
9: $\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\mathbf{P}_{k+1|k}$
10: $k = k + 1$
11: **end loop**

In this work, the integration is done numerically using a 4th order Runge-Kutta algorithm.

4 COMPUTATIONAL SIMULATION

To evaluate the proposed problem, a simulation platform is used. The simulation environment is composed with a nonlinear model of a quadrotor, modeled as in Mahony *et al.* (2012), using a saturated Proportional-Derivative (PD) controller, introduced in Santos *et al.* (2013), to follow a predefined position trajectory. 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 the visible landmarks. 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 quadrotor trajectory is defined by seven waypoints as displayed in Table 1. The route between the waypoints is a linear trajectory that connects these points in space. The translational velocity is constant, $v_w = 0.2$ m/s. The heading simulation is defined as a sinusoidal wave with frequency of $2\frac{\pi}{10}$ rad/sec and amplitude of $\frac{\pi}{6}$ meters.

Table 1: Position trajectory waypoints.

$\boldsymbol{r}(m)$	1	2	3	4	5	6	7
$\boldsymbol{r}_x$	0	0	1	1	0	0	0
$\boldsymbol{r}_y$	0	0	0	1	1	0	0
$\boldsymbol{r}_z$	1.5	2	2	2	2	2	1.5

The 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 covariance of the state noise is tuned in $\mathbf{Q} = \text{diag}\{\sigma_r^2, \sigma_v^2, \sigma_\psi^2, \sigma_g^2, \sigma_a^2\}$ with $\sigma_r^2 = 1.0 \times 10^{-8}\mathbf{I}_3$ m^2, $\sigma_v^2 = 1.0 \times 10^{-8}\mathbf{I}_3$ (m/s)^2, $\sigma_\psi^2 = 1.0 \times 10^{-6}$ rad^2, $\sigma_g^2 = 1.0 \times 10^{-8}$ (rad/s)^2 and $\sigma_a^2 = 1.0 \times 10^{-6}\mathbf{I}_3$ (m/s^2)^2. Likewise, the covariance of the measurement noise is tuned in $\mathbf{R}_k^{(i)} = 1.0 \times 10^{-4}\mathbf{I}_3$ m^2. In addition, to test the robustness of the filter, it is introduced in the ultrasonic range sensor measurement a variance of $\sigma_u^2 = 1.0 \times 10^{-6}$ m^2, as well as in the measurements of the camera $\check{x}_{\text{I}}^{(i)}$ and $\check{y}_{\text{I}}^{(i)}$, with variances of $\sigma_{\text{I,x}}^2 = \sigma_{\text{I,y}}^2 = 1.0 \times 10^{-4}$ m^2, simulating the image processing algorithm errors. The initial conditions is tuned in $\overline{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 1.5 & \mathbf{0}_{1\times 8} \end{bmatrix}^{\text{T}}$ and $\overline{\mathbf{P}} = \text{diag}\{\sigma_{0,r}^2, \sigma_{0,v}^2, \sigma_{0,\psi}^2, \sigma_{0,g}^2, \sigma_{0,a}^2\}$, where $\sigma_{0,r}^2 = \text{diag}\{1, 1, 1.0 \times 10^{-2}\}$ m^2, $\sigma_{0,v}^2 = 1.0 \times 10^{-4}\mathbf{I}_3$ (m/s)^2, $\sigma_{0,\psi}^2 = (0.25\pi)^2$ rad^2, $\sigma_{0,g}^2 = 1.0 \times 10^{-8}$ (rad/s)^2 and $\sigma_a^2 = 1.0 \times 10^{-8}\mathbf{I}_3$ (m/s^2)^2.

4.1 Results

The EKF filter is evaluated via extensive Monte Carlo (MC) simulations. In this work 100 simulation runs are performed for each realization, completing 4 Monte Carlo simulations. For each MC simulation the number of landmarks used in the update phase is changed in order to analyze the behavior of the proposed filter. At the first Monte Carlo simulation, the filter uses a unique landmark. Secondly, the filter uses two of the seen landmarks at a time t_k, taking care of choosing the two most distant landmarks in the image, thus providing better information about the desired states. The third MC simulation uses three landmarks and the fourth uses four. The result is depicted in Table 2.

Table 2: Monte Carlo simulations for different values of landmarks used.

Num. Landmarks (q)	ε_p	σ_p	ε_v	σ_v	ε_ψ	σ_ψ
1	9.4	26.2	8.4	9.8	15.7	132.3
2	7.7	4.6	8.1	11.5	8.3	20.4
3	5.2	4.2	6.1	11.4	6.2	10.1
4	4.7	3.1	6.1	11.3	5.7	9.7

The following figure of merit are used to evaluate the position estimation error:

$$\varepsilon_p \triangleq \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{k_f} e_p^{(i)}(k) \tag{41}$$

$$\sigma_p \triangleq \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{k_f}\left(e_p^{(i)}(k) - \varepsilon_p\right)^2} \tag{42}$$

with

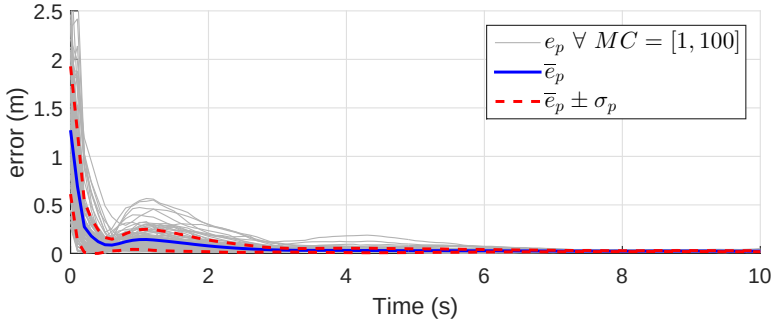
$$e_p^{(i)}(k) \triangleq \| \mathbf{r}(k) - \hat{\mathbf{r}}(k) \|_2^{(i)}, \;\; k = 1, 2, ..., k_f, \tag{43}$$

where the pair $\left(\varepsilon_p, \sigma_p^2\right)$ represents the sampling mean and variance of the MC simulation. The $\| \; . \; \|_2^{(i)}$ represents the Euclidean norm of the i-th realization, $\hat{\mathbf{r}}(k)$ represents the estimate of the position $\mathbf{r} \triangleq \mathbf{r}_G^{\text{G}/P}$ at the instant k, and N

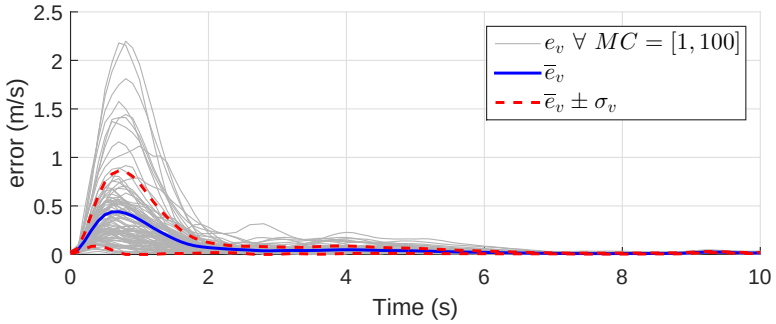
represents the number of simulation realizations. In this case, $N = 100$. Similarly, the velocity $(\varepsilon_v, \sigma_v)$ and heading $(\varepsilon_\psi, \sigma_\psi)$ estimation error are obtained using the same figure of merit, replacing $\mathbf{r}$ with $\mathbf{v}$ and ψ, respectively. The constant k_f represents the number of samples taken from each realization realization.

From Table 2 it is possible to notice the relationship between the number of landmarks used in the EKF and its accuracy. The more landmarks we use in the update phase of the filter, the preciser is the filter. Nonetheless, as stated in the end of Section 3.2, it comes with a higher computational effort. As can be noted, using only 1 landmark it is not possible to estimate the desired states. In fact, it was expected, since we need at least 2 vector measurements to estimate an unique position and heading estimate (Sun and Crassidis, 2002). Further, considering the other rows of Table 2, it can be seen that there is a significant increase in accuracy from 2 to 3 landmarks, but only a slight increase when 4 landmarks are used. For this reason, we found that using 3 landmarks in the EKF is sufficient to obtain a precise estimation solution to the proposed problem, since greater values of q shows lower tendency to increase the filter's accuracy.

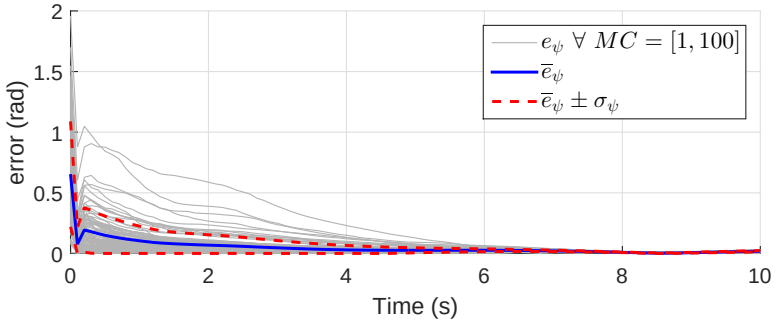
The Figure 3 shows the temporal response of the position, velocity and heading estimation error for the Monte Carlo simulations using 3 landmarks. As can be observed, the filter has higher errors between 0 and 2 seconds of simulation. This occurs mostly due to the large initial conditions errors in position and heading. These position initial conditions also 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. (19)) does not provide information about the vehicle's velocity. Moreover, it can be seen that the filter converges in approximately 2 seconds.



(a) Position Estimation error.



(b) Velocity Estimation error.



(c) Heading Estimation Error

Figure 3: Monte Carlo simulations using $q = 3$ landmarks.

5 CONCLUSION

The present paper has proposed a method of accurately and robustly estimating the position, velocity and heading of an Unmanned Aerial Vehicle in GNSS-denied environment using an embedded camera, inertial sensors and an aid of an ultrasonic range finder. The method is based on an Extended Kalman Filter that is responsible for the sensor fusion between the different type of sensors and the camera, that capture visible landmarks, whose positions are previously

known. The proposed solution was evaluated via computation simulations using a quadrotor vehicle model, but can be expanded to any UAV with a stabilizing gimbal platform. Moreover, a Monte Carlo simulation with 100 realizations was made to verify the relationship between the filter accuracy and the number of landmarks used. As expected, as higher number of landmarks is used, preciser the filter becomes, but with a higher computational cost. With these results, it could be shown that the filter could successfully estimate the desired states with the given set of sensors. In summary, the proposed estimation method represents a precise and low-cost alternative to the proposed problem. For future work, other filters are being developed and it will be possible to carry out a comparison of using different approaches to tackle the proposed problem.

6 REFERENCES

Anderson, B.D.O., Moore, J.B. and Eslami, M., 1982. "Optimal Filtering". *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 12, No. 2, pp. 235–236.

Arasaratnam, I. and Haykin, S., 2009. "Cubature kalman filters". *Automatic Control, IEEE Transactions on*, Vol. 54, No. 6, pp. 1254–1269.

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*, Vol. 94, No. 7, pp. 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*, Vol. 65, No. 1-4, pp. 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*, Vol. 28, pp. 32–37.

Evensen, G., 2003. "The Ensemble Kalman Filter: Theoretical formulation and practical implementation". *Ocean Dynamics*, Vol. 53, No. 4, pp. 343–367.

Gelb, A., 1974. *Applied Optimal Estimation*, Vol. 64.

Gini, G.C. and Marchi, A., 2002. "Indoor Robot Navigation With Single Camera Vision". *Pris*, pp. 67–76.

Hartley, R. and Zisserman, A., 2015. *Multiple view geometry in computer vision*, Vol. 1.

Hoffmann, G.M., Waslander, S.L. and Tomlin, C.J., 2008. "Quadrotor Helicopter Trajectory Tracking Control". *Electrical Engineering*, Vol. 44, No. August, pp. 1–14.

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

Mahony, R., Kumar, V. and Corke, P., 2012. "Multirotor Aerial Vehicles: Modeling, Estimation, and Control of Quadrotor". *IEEE Robotics & Automation Magazine*, Vol. 19, No. 3, pp. 20–32.

Ponce, J. and Forsyth, D., 2012. *Computer vision: a modern approach*.

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*, Vol. 13, No. 4, pp. 517–523.

Santos, D.A., Saotome, O. and Cela, A., 2013. "Trajectory control of multirotor helicopters with thrust vector constraints". *2013 21st Mediterranean Conference on Control and Automation, MED 2013 - Conference Proceedings*, pp. 375–379.

Sun, D. and Crassidis, J.L., 2002. "Observability Analysis of Six-Degree-of-Freedom Configuration Determination Using Vector Observations". *Jgcd*, Vol. 25, No. 6, pp. 1149–1157.

Thrun, S., 2002. "Probabilistic robotics". *Communications of the ACM*, Vol. 45, No. 3, pp. 1999–2000.

Trawny, N., Mourikis, A.I., Roumeliotis, S.I., Johnson, A.E. and Montgomery, J.F., 2007. "Vision-aided inertial navigation for pin-point landing using observations of mapped landmarks". *Journal of Field Robotics*, Vol. 24, No. 5, pp. 357–378.

Wendel, J., Meister, O., Schlaile, C. and Trommer, G.F., 2006. "An integrated GPS/MEMS-IMU navigation system for an autonomous helicopter". *Aerospace Science and Technology*, Vol. 10, No. 6, pp. 527–533.

Zhang, L., Li, T., Yang, H., Zhang, S., Cai, H. and Qian, S., 2015. "Unscented Kalman Filtering for Relative Spacecraft Attitude and Position Estimation". *Journal of Navigation*, Vol. 68, pp. 528–548.