Proceedings of the 2011 IEEE
International Conference on Mechatronics and Automation
August 7 - 10, Beijing, China

UAV-based Marine Mammals Positioning and Tracking System

Tiago Gaspar, Paulo Oliveira, and Carlos Silvestre
Institute for Systems and Robotics, Instituto Superior Técnico
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
Email: {tgaspar, pjcro, cjs}@isr.ist.utl.pt

***Abstract*— In this paper, a new strategy to localize and track marine mammals moving at the ocean surface is proposed, resorting to measurements provided by a GPS, an Attitude and Heading Reference System (AHRS), and images provided by a camera installed onboard an Unmanned Aerial Vehicle (UAV). The segmentation of the marine mammals in the images is tackled resorting to methodologies based on active contours. The measurements of the position of the target are combined with data from the position and attitude of the UAV, provided by the GPS and AHRS, respectively, leading to target position estimates, in an inertial reference frame. To obtain these estimates, two Kalman Filters are proposed: i) a time-invariant filter, that estimates only the target position, and ii) a time-varying filter, that combines estimates of the position of the target with those of the position of the UAV, which is shown to improve the overall performance. Simulation results illustrating the behavior of the system in realistic conditions are presented and discussed. Results are also reported in the case where occlusions in the images occur, namely during the periods where the marine mammals dive, to allow the filter robustness assessment.**

I. Introduction

In the last years, there has been a massification of the use of systems such as acoustic telemetry [1] or pop-up satellite archival tags [2] to study the daily movement patterns and behaviour of marine animals [1], [3], and migration patterns in marine protected areas [1], for instance. More recently, the development and use of Unmanned Aerial Vehicles (UAVs) [4] as tools for ocean surface data acquisition has also been exploited. However, most vehicles have been designed to conduct simple survey missions that in general do not require close interaction between the operator and the environment. It is by now felt that the effective use of UAVs in demanding marine science applications must be clearly demonstrated, namely by evaluating UAV-based systems in terms of their performance and adaptability to different missions scenarios.

One of the marine science areas that can profit from the use of UAVs is location and tracking of marine mammals [5]. Other applications including, for instance, the use of UAVs for sea surface temperature measurement and for directing research vessels to new areas of interest, enabling a more efficient use of ship time, are also foreseen as useful tools that can help marine scientists in facing nowadays challenges.

In this work, the problem of tracking and estimating the position of a marine mammal moving at the ocean surface is addressed [5], [6], [7]. With this purpose, an UAV instrumented with a GPS, an Attitude and Heading Reference System (AHRS), and an image acquisition module, which consists of a digital video camera mounted on a pan-tilt unit, is used. The intrinsically nonlinear problem that results is casted into a linear one by rewriting the measurements provided by the camera in an alternative form. Based on these measurements, and on the measurements provided by the GPS and AHRS, two linear Kalman Filters [8] that estimate the position of the target are proposed. One, time-invariant, that estimates only the target position, and other, time-varying, that combines estimates of the target position with estimates of the position of the UAV to improve its performance. These filters are used in a Multiple-Model Adaptive Estimation (MMAE) strategy [9], [10], which copes with the uncertainty associated with one of the parameters of the model considered for the target, in this case its angular turn rate. A set of simulations carried out under realistic noise conditions is provided, and results illustrating the performance of the system are presented. The case where the marine animal dives, leading to occlusions in images, is addressed and validated in one of the simulations.

This document is organized as follows. In section II, the problem addressed in this work is described, and the models considered for the target, UAV, and sensors are presented. The proposed MMAE Kalman Filters are derived in section III, and in section IV simulation results are presented and discussed. In section V, some concluding remarks are outlined.

II. Problem Statement

Consider an UAV instrumented with an AHRS which provides estimates of the orientation ${}^I\mathbf{R}_p(t) \in SO(3)$ of a body-fixed frame $\{P\}$, attached to the UAV, with respect to an inertial reference frame $\{I\}$, along the time t (${}^I\mathbf{R}_p(t)$ is the rotation matrix that rotates the coordinates of points from frame $\{P\}$ to frame $\{I\}$). The Special Orthogonal group is here denoted by $SO(3)$. A GPS installed onboard provides estimates of the position of the origin ${}^I\mathbf{X}_p(t) \in \mathbb{R}^3$ of $\{P\}$ expressed in the inertial frame $\{I\}$. Without loss of generality, the inertial reference frame is considered to have its origin in the vicinity of the mission scenario and at the sea surface, with the z-axis orthogonal to it and pointing upwards, see Fig. 1. The UAV is also instrumented with an image acquisition module, which consists of a digital video camera mounted on a pan-tilt unit. The position ${}^p\mathbf{X}_c \in \mathbb{R}^3$ of the origin of a body-fixed frame $\{C\}$, attached to the camera, expressed in the aircraft frame $\{P\}$ and the orientation ${}^p\mathbf{R}_c \in SO(3)$ of $\{C\}$ with respect to $\{P\}$ are known (${}^p\mathbf{R}_c$ is the rotation

This work was partially funded by FCT (ISR/IST plurianual funding) through the PIDDAC Program and by the FCT project PTDC/MAR/64546/2006 OBSERVFLY. The work of Tiago Gaspar was supported by the PhD Student Scholarship SFRH/BD/46860/2008, from FCT.

matrix that rotates the coordinates of points from frame $\{C\}$ to frame $\{P\}$). The origin of $\{C\}$ is at the camera optical centre and its z-axis is aligned with the camera optical axis.

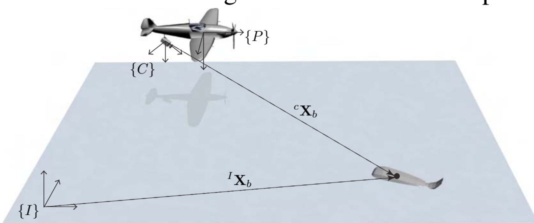


Fig. 1. Mission scenario.

Moreover, consider a marine mammal moving at the sea surface and denote the inertial coordinates of its position by ${}^{I}\mathbf{X}_b(t) \in \mathbb{R}^3$. According to the setup described, the position of the target in the inertial reference frame is given by the vector ${}^{I}\mathbf{X}_b(t) = [{}^{I}x_b(t)\ {}^{I}y_b(t)\ 0]^T$.

The problem considered in this paper is that of tracking the marine mammal and obtaining estimates $[{}^{I}\hat{x}_b(t)\ {}^{I}\hat{y}_b(t)]^T$ of its position, using measurements provided by the GPS, AHRS, and image acquisition module. The measurements provided by the AHRS are considered to be ideal, i.e., not corrupted by noise, whereas those provided by the GPS and image acquisition module are considered to be corrupted by noise.

A. Marine mammals model

State-space models have been used in the characterization of the movement of several animals, as reported in [11]. In the sequel, the dynamical model chosen for the marine mammal is described. Given the trajectories expected for this type of targets, which, as stated before, are assumed to lie in a plane coincident with the sea surface, the *2D Horizontal Constant-Turn Model with Known Turn Rate*, as presented in [7], is selected. This model assumes that the target moves with constant speed and constant angular (turn) rate ω. Assuming that ω is known, a four-dimensional state vector $\mathbf{x}_b(t) = [{}^{I}x_b(t)\ {}^{I}\dot{x}_b(t)\ {}^{I}y_b(t)\ {}^{I}\dot{y}_b(t)]^T$ results, where $[{}^{I}x_b(t)\ {}^{I}y_b(t)]^T$ and $[{}^{I}\dot{x}_b(t)\ {}^{I}\dot{y}_b(t)]^T$ are, respectively, the position and velocity of the target expressed in inertial coordinates. Let $\mathbf{x}_b(t_k)$ denote the values taken by the state $\mathbf{x}_b(t)$ of the target at time instants $t_k = kT$, $k \in \mathbb{N}$, where $T > 0$ is the sampling interval. In this case, the target state dynamics assumes the following linear parametrically varying form

$$\begin{aligned}\mathbf{x}_b(t_k) &= \mathbf{F}_b(\omega)\mathbf{x}_b(t_{k-1}) + \mathbf{w}_b(t_{k-1})\\ &= \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & -\frac{1-\cos(\omega T)}{\omega} \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & \frac{1-\cos(\omega T)}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix}\mathbf{x}_b(t_{k-1}) + \mathbf{w}_b(t_{k-1}).\end{aligned}$$

The process noise $\mathbf{w}_b(t_k) \in \mathbb{R}^4$ is assumed to be Gaussian and zero-mean, with covariance matrix

$$\mathbf{Q}_b(\omega) = S_{w_b}\begin{bmatrix} \frac{2(\omega T-\sin(\omega T))}{\omega^3} & \frac{1-\cos(\omega T)}{w^2} & 0 & \frac{\omega T-\sin(\omega T)}{\omega^2} \\ \frac{1-\cos(\omega T)}{\omega^2} & T & -\frac{\omega T-\sin(\omega T)}{\omega^2} & 0 \\ 0 & -\frac{\omega T-\sin(\omega T)}{\omega^2} & \frac{2(\omega T-\sin(\omega T))}{\omega^3} & \frac{1-\cos(\omega T)}{\omega^2} \\ \frac{\omega T-\sin(\omega T)}{\omega^2} & 0 & \frac{1-\cos(\omega T)}{w^2} & T \end{bmatrix}, \tag{1}$$

where S_{w_b} corresponds to the power spectral density of the continuous-time process noises that affect the two components of the velocity of the target. More details about this model and the relations presented above can be found in [7].

Note that the state dynamics and the process noise covariance matrix depend explicitly on the target turn rate, whose actual value is unknown. A solution that copes with this problem will be proposed in section III.

B. UAV model

In what follows, we avoid writing explicitly the dynamical equations of the UAV and rely only on its kinematic equations of motion. Thus, a general solution that suits different kinds of aircrafts is obtained.

Let $\mathbf{x}_p(t) = [{}^{I}x_p\ {}^{I}\dot{x}_p\ {}^{I}\ddot{x}_p\ {}^{I}y_p\ {}^{I}\dot{y}_p\ {}^{I}\ddot{y}_p\ {}^{I}z_p\ {}^{I}\dot{z}_p\ {}^{I}\ddot{z}_p]^T$, where $[{}^{I}x_p\ {}^{I}y_p\ {}^{I}z_p]^T$, $[{}^{I}\dot{x}_p\ {}^{I}\dot{y}_p\ {}^{I}\dot{z}_p]^T$, and $[{}^{I}\ddot{x}_p\ {}^{I}\ddot{y}_p\ {}^{I}\ddot{z}_p]^T$ denote, respectively, the UAV position, velocity, and acceleration expressed in inertial coordinates, and consider a Wiener process model for the acceleration of the UAV, see [6]. The dependence of the scalar quantities on the time t was omitted for simplicity of presentation. The process noise that affects the three components of the acceleration of the UAV is assumed to be Gaussian and zero-mean, with power spectral density matrix diag$[S_{w_p}, S_{w_p}, S_{w_p}]$. Under these assumptions, the motion of the UAV can be described by the discrete-time state equation

$$\mathbf{x}_p(t_k) = \mathbf{F}_p\mathbf{x}_p(t_{k-1}) + \mathbf{w}_p(t_{k-1}), \tag{2}$$

where $t_k = kT$, $t_{k-1} = (k-1)T$, $k \in \mathbb{N}$, and $T > 0$ is the sampling interval. In this expression, $\mathbf{F}_p =$ diag$[\mathbf{F}, \mathbf{F}, \mathbf{F}]$ and the value of the discrete-time process noise $\mathbf{w}_p(t_k)$ is associated with the covariance matrix $\mathbf{Q}_p =$ diag$[S_{w_p}\mathbf{Q}, S_{w_p}\mathbf{Q}, S_{w_p}\mathbf{Q}]$, with

$$\mathbf{F} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{Q} = \begin{bmatrix} T^5/20 & T^4/8 & T^3/6 \\ T^4/8 & T^3/3 & T^2/2 \\ T^3/6 & T^2/2 & T \end{bmatrix}.$$

C. Measurements model

As stated before, the UAV is instrumented with a GPS and an AHRS, which provide measurements of its position and attitude with respect to $\{I\}$, and with an image acquisition module that acquires images of a marine mammal moving at the sea surface. In the sequel, the measurements provided by these sensors will be described.

Data obtained from the GPS and AHRS correspond to direct measurements of the position ${}^{I}\mathbf{X}_p$ and orientation ${}^{I}\mathbf{R}_p$ of the UAV with respect to the inertial reference frame. Measurements provided by the image acquisition module are more intricate. Given the high complexity of camera optical systems, and the consequent high number of parameters required to model the whole image acquisition process, it is common to use a linear model for cameras. In this work, the pinhole model is considered, see [12] for details about this model.

Let $\mathbf{M} = [x, y, z, 1]^T$ be the homogeneous coordinates of a point with Euclidean coordinates $\mathbf{X} = [x, y, z]^T$, expressed in the inertial reference frame, and $\mathbf{m} = [u, v, 1]^T$ the homogeneous coordinates of the projection $\mathbf{x} = [u, v]^T$ of that point into the image frame. According to the pinhole model, the relation between the coordinates expressed in both frames is given by

$$\lambda\mathbf{m} = \mathbf{PM}, \tag{3}$$

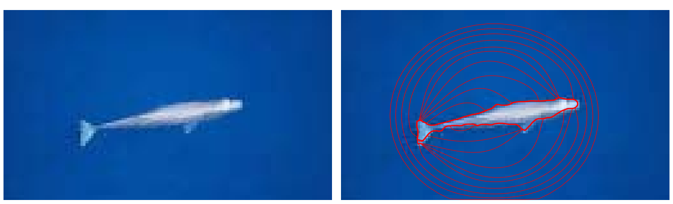
where λ is a multiplicative constant related with the distance from the point in the world to the camera, and $\mathbf{P}$ is the 3×4 projection matrix that relates 3D inertial coordinates with 2D image coordinates.

Consider that the centre of the marine mammal in the image is the projection of the point used to define the position ${}^I\mathbf{X}_b$ of the target in the inertial reference frame. Let the coordinates of this point in the image frame be denoted by (u, v). Then, according to the model selected, measurements provided by the camera can be written in the form

$$u = \frac{p_{11}{}^I x_b + p_{12}{}^I y_b + p_{14}}{p_{31}{}^I x_b + p_{32}{}^I y_b + p_{34}} \text{ and } v = \frac{p_{21}{}^I x_b + p_{22}{}^I y_b + p_{24}}{p_{31}{}^I x_b + p_{32}{}^I y_b + p_{34}}, \tag{4}$$

where p_{ij} is the projection matrix element in the i-th line and j-th column. As can be seen, these measurements are a nonlinear function of the state ${}^I\mathbf{X}_b$ of the target. The absence of a z-component in this expression is due to the fact that, according to the topology of the setup described in the beginning of this section, the 3D position of the target is given by ${}^I\mathbf{X}_b = [{}^I x_b \ {}^I y_b \ 0]^T$.

The marine mammal is segmented and isolated resorting to active contours, see [13] and [14]. Once segmented the mammal (see example in Fig. 2), the coordinates (u, v) of the point used to define the position of the target in the world, in this case the centre of the target boundary, are easily computed as the mean of the coordinates of the points that belong to the estimated contour.



(a) Marine mammal. (b) Target segmentation.

Fig. 2. Example of segmentation of a marine mammal using active contours.

III. Filters Design

In this section, the estimation methodologies adopted to locate the marine mammal are detailed. Two Kalman Filters with different structures are derived. These filters are here called isolated and joint, depending on the use, or not, of estimates of the position of the UAV to help in the estimation of the position of the target. A Multiple-Model Adaptive Estimation approach is used to deal with the unknown nature of the target turn rate. At the end of the section, the Baram Proximity Measure (BPM) is used to provide some insight into how to choose the nominal turn rate values for each underlying Kalman Filter model.

As stated before, the measurements used to estimate the position of the target with respect to the inertial reference frame are the UAV position (provided by the GPS) and orientation (provided by the AHRS), and the coordinates, in images acquired by the image acquisition module, of the point that defines the position of the target. Moreover, the proposed estimation strategies do not consider noise in the measurements provided by the AHRS.

A. *Isolated Kalman Filter*

Kalman Filters (KFs) [8] provide an optimal solution (in the minimum mean square error sense) to the problem of estimating the state of a discrete-time gaussian process that is described by a linear stochastic difference equation. However, this approach is not valid when the process and/or the measurements are nonlinear. One of the most successful approaches in these situations consists in applying a linear time-varying Kalman Filter to a system that results from the linearization of the original nonlinear one, along the estimates. These filters, usually referred to as Extended Kalman Filters (EKFs) [8], do not provide, however, the performance and stability guarantees that linear Kalman Filters ensure under the proper observability and controlability conditions, see [15].

As can be seen from (4), the relation between the coordinates ${}^I\mathbf{X}_b$ of the point that represents the target, and the coordinates (u, v) of the projection of that point into the image frame, is nonlinear. In the sequel, a strategy that circumvents the undesired characteristics of such relation is proposed.

The transformation given by the projection matrix in (3) can be decomposed into the form $\mathbf{P} = \mathbf{K}[{}^c\mathbf{R}_I \mid {}^c\mathbf{X}_I]$, where $\mathbf{K}$ is the 3×3 intrinsic parameters matrix, i.e., a matrix that depends on the camera internal parameters, such as its focal length, and $[{}^c\mathbf{R}_I \mid {}^c\mathbf{X}_I]$ is the 3×4 external parameters matrix, i.e., corresponds to the Euclidean transformation from the inertial reference frame $\{I\}$ to the camera reference frame $\{C\}$: ${}^c\mathbf{R}_I$ denotes a 3×3 rotation matrix from $\{I\}$ to $\{C\}$, and ${}^c\mathbf{X}_I$ corresponds to the Euclidean coordinates of the origin of $\{I\}$ expressed in $\{C\}$. If $\mathbf{m} = [u, v, 1]^T$, then, according to the pinhole model described in section II-C, and using the properties of rigid body transformations, see [16], it is possible to rewrite (3) in the form

$$\begin{aligned} \lambda \mathbf{m} &= \mathbf{K}({}^c\mathbf{R}_I {}^I\mathbf{X}_b + {}^c\mathbf{X}_I) \\ &= \mathbf{K}{}^p\mathbf{R}_c^T[{}^I\mathbf{R}_p^T({}^I\mathbf{X}_b - {}^I\mathbf{X}_p) - {}^p\mathbf{X}_c]. \end{aligned} \tag{5}$$

Apart from λ and ${}^I\mathbf{X}_b$, all other quantities on this equation are either known ($\mathbf{K}$, ${}^p\mathbf{R}_c$, and ${}^I\mathbf{X}_b$) or measured ($\mathbf{m}$, ${}^I\mathbf{R}_p$, and ${}^I\mathbf{X}_p$). Given the constraint ${}^I z_b = 0$ on the third component of ${}^I\mathbf{X}_b$, this system is a linear system with three equations and three unknowns (λ, ${}^I x_b$, and ${}^I y_b$). By replacing the values of $\mathbf{m}$, ${}^I\mathbf{R}_p$, and ${}^I\mathbf{X}_p$, in (5), by the values of their measurements, and by solving the system that results, we transform the measurements of (u, v), which depend nonlinearly on the state of the target, into measurements $({}^I x_{b_m}, {}^I y_{b_m})$ of $({}^I x_b, {}^I y_b)$, that are a linear function of the state of the marine mammal.

In order to obtain estimates $[{}^I\hat{x}_b \ {}^I\hat{y}_b]^T$ of the marine mammal position $[{}^I x_b \ {}^I y_b]^T$, consider a system with state $\mathbf{x}_I = \mathbf{x}_b = [{}^I x_b \ {}^I\dot{x}_b \ {}^I y_b \ {}^I\dot{y}_b]^T \in \mathbb{R}^4$, whose evolution is governed by the linear stochastic difference equation (2), and consider that at a given instant t_k of time, measurements $\mathbf{z}_I(t_k) = [{}^I x_{b_m}(t_k) \ {}^I y_{b_m}(t_k)]^T \in \mathbb{R}^2$, given by

$$\mathbf{z}_I(t_k) = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{C}_I} \mathbf{x}_I(t_k) + \mathbf{v}_I(t_k), \tag{6}$$

are available. The process $\mathbf{w}_I = \mathbf{w}_b$ and measurement $\mathbf{v}_I$ noises are assumed to be white, zero-mean, Gaussian, independent ($E[\mathbf{w}_I(t_k)\mathbf{v}_I^T(t_j)] = 0$, for all t_k, t_j, where $E[.]$ denotes the expected value operator), and with covariance matrices $\mathbf{Q}_I(\omega) = \mathbf{Q}_b(\omega)$ and $\mathbf{V}_I$ given, respectively, by (1) and $\mathbf{V}_I(t_k) = \mathrm{diag}[\sigma_{x_b}^2,\ \sigma_{y_b}^2]$, where σ_{x_b} and σ_{y_b} denote the standard deviations of the measurements ${}^I x_{b_m}$ and ${}^I y_{b_m}$.

The following notation will be used in the sequel: $\mathbf{x}_I^-(t_k)$ and $\mathbf{x}_I^+(t_k)$ denote the *a priori* and *a posteriori* state estimates, and $\mathbf{P}_I^-(t_k)$ and $\mathbf{P}_I^+(t_k)$ denote the *a priori* and *a posteriori* state covariance matrices, respectively. Moreover, let $\mathbf{x}_I^+(t_0) = E[\mathbf{x}_I(t_0)]$ and $\mathbf{P}_I^+(t_0) = E[(\mathbf{x}_I(t_0) - \mathbf{x}_I^+(t_0))(\mathbf{x}_I(t_0) - \mathbf{x}_I^+(t_0))^T]$. Then, the proposed Kalman Filter is given by the following equations

Predict step:

$$\begin{aligned}
\mathbf{x}_I^-(t_k) &= \mathbf{F}_I(\omega)\mathbf{x}_I^+(t_{k-1}) \\
\mathbf{P}_I^-(t_k) &= \mathbf{F}_I(\omega)\mathbf{P}_I^+(t_{k-1})\mathbf{F}_I^T(\omega) + \mathbf{Q}_I(\omega)
\end{aligned}$$

Update step:

$$\begin{aligned}
\mathbf{K}_I(t_k) &= \mathbf{P}_I^-(t_k)\mathbf{C}_I^T[\mathbf{C}_I\mathbf{P}_I^-(t_k)\mathbf{C}_I^T + \mathbf{V}_I]^{-1} \\
\mathbf{x}_I^+(t_k) &= \mathbf{x}_I^-(t_k) + \mathbf{K}_I(t_k)(\mathbf{z}_I(t_k) - \mathbf{C}_I\mathbf{x}_I^-(t_k)) \\
\mathbf{P}_I^+(t_k) &= [\mathbf{I} - \mathbf{K}_I(t_k)\mathbf{C}_I]\mathbf{P}_I^-(t_k),
\end{aligned} \tag{7}$$

which are computed for each $k \in \mathbb{N}$, see [8] for details. In these expressions, $\mathbf{I}$ denotes the identity matrix and $\mathbf{F}_I(\omega)$ the matrix that defines the dynamics of $\mathbf{x}_I$, i.e., $\mathbf{F}_I(\omega) = \mathbf{F}_b(\omega)$.

According to the strategy described, estimates $[{}^I\hat{x}_b(t_k)\ \ {}^I\hat{y}_b(t_k)]^T$ of $[{}^I x_b\ {}^I y_b]^T$ at time instants t_k can be obtained from the first and third components of $\mathbf{x}_I^+(t_k)$.

B. Joint Kalman Filter

The methodology detailed in the previous section makes direct use of the measurements provided by the GPS in the transformation of the measurements of (u, v) into measurements of $({}^I x_b,\ {}^I y_b)$, which depend linearly on the state of the target. In this section, a different approach is pursued, as the measurements provided by the GPS are also incorporated into the filtering process. A filter with a different structure results, and new estimates of the state of the target, as well as estimates of the state of the UAV, are provided.

In order to derive the new relation between the measurements provided to the filter and the state to be estimated, rewrite (5) in the form

$$\lambda\mathbf{K}^{-1}\mathbf{m} + {}^p\mathbf{R}_c^T\, {}^p\mathbf{X}_c = {}^p\mathbf{R}_c^T\, {}^I\mathbf{R}_p^T({}^I\mathbf{X}_b - {}^I\mathbf{X}_p). \tag{8}$$

Moreover, let matrices $\mathbf{T}$ and $\mathbf{C}_p$ be defined in such a way that ${}^I\mathbf{X}_b = \mathbf{T}\mathbf{x}_b$ and ${}^I\mathbf{X}_p = \mathbf{C}_p\mathbf{x}_p$, i.e., $\mathbf{T} = [\ \mathbf{e}_1\ \ \mathbf{0}\ \ \mathbf{e}_2\ \ \mathbf{0}\ ]$ and $\mathbf{C}_p = [\ \mathbf{e}_1\ \ \mathbf{0}\ \ \mathbf{0}\ \ \mathbf{e}_2\ \ \mathbf{0}\ \ \mathbf{0}\ \ \mathbf{e}_3\ \ \mathbf{0}\ \ \mathbf{0}\ ]$, where $\mathbf{e}_i$, $i = 1, 2, 3$, and $\mathbf{0}$ are the i-th vector of the canonical basis of $\mathbb{R}^3$ and a column vector with 3 zeros, respectively. Combining these relations with (8), yields

$$\underbrace{\lambda\mathbf{K}^{-1}\mathbf{m} + {}^p\mathbf{R}_c^T\, {}^p\mathbf{X}_c}_{\mathbf{y}} = {}^p\mathbf{R}_c^T\, {}^I\mathbf{R}_p^T(\mathbf{T}\mathbf{x}_b - \mathbf{C}_p\mathbf{x}_p). \tag{9}$$

By replacing $\mathbf{m}$ by the value of its measurements and λ by the value obtained according to the procedure described in section III-A, we transform the measurements of (u, v), which are a nonlinear function of the state of the target, into measurements $\mathbf{y}_m$, of $\mathbf{y}$, that depend linearly on the state $\mathbf{x}_b$ of the target and on the state $\mathbf{x}_p$ of the UAV.

If the state of the new system is considered to be $\mathbf{x}_J = [\mathbf{x}_b^T\ \mathbf{x}_p^T]^T \in \mathbb{R}^{13}$, then, at a given instant t_k of time, measurements $\mathbf{z}_J(t_k) = [\mathbf{y}_m^T(t_k)\ \ {}^I\mathbf{X}_{p_m}^T(t_k)]^T \in \mathbb{R}^6$, given by

$$\mathbf{z}_J(t_k) = \underbrace{\begin{bmatrix} {}^p\mathbf{R}_c^T\, {}^I\mathbf{R}_p^T(t_k)\mathbf{T} & -{}^p\mathbf{R}_c^T\, {}^I\mathbf{R}_p^T(t_k)\mathbf{C}_p \\ \mathbf{0}_{3\times 4} & \mathbf{C}_p \end{bmatrix}}_{\mathbf{C}_J(t_k)} \mathbf{x}_J(t_k) + \mathbf{v}_J(t_k), \tag{10}$$

are available. The values of $\mathbf{y}_m(t_k)$, ${}^I\mathbf{X}_{p_m}(t_k)$, and $\mathbf{v}_J(t_k)$ denote the measurements of $\mathbf{y}$, the measurements of ${}^I\mathbf{X}_p$, and the measurement noise, respectively, at instant t_k. In this equation, and throughout the remaining of this document, the term $\mathbf{0}_{i\times j}$ represents a matrix of zeros with dimension $i \times j$. Since the value of ${}^I\mathbf{R}_p$ depends on time, in this case the measurement equation is time varying.

The state $\mathbf{x}_J$ of the system corresponds to the concatenation of the states of the target and UAV, therefore its evolution over time is modelled by the linear stochastic difference equation

$$\mathbf{x}_J(t_k) = \underbrace{\begin{bmatrix} \mathbf{F}_b(\omega) & \mathbf{0}_{4\times 9} \\ \mathbf{0}_{9\times 4} & \mathbf{F}_p \end{bmatrix}}_{\mathbf{F}_J(\omega)} \mathbf{x}_J(t_{k-1}) + \underbrace{\begin{bmatrix} \mathbf{w}_b(t_{k-1}) \\ \mathbf{w}_p(t_{k-1}) \end{bmatrix}}_{\mathbf{w}_J(t_{k-1})}.$$

The process $\mathbf{w}_J$ and measurement $\mathbf{v}_J$ noises are assumed to be white, zero-mean, Gaussian, independent ($E[\mathbf{w}_J(t_k)\mathbf{v}_J^T(t_j)] = 0$, for all t_k, t_j), and with covariance matrices $\mathbf{Q}_J(\omega)$ and $\mathbf{V}_J$, respectively, of the form

$$\mathbf{Q}_J(\omega) = \begin{bmatrix} \mathbf{Q}_b(\omega) & \mathbf{0}_{4\times 9} \\ \mathbf{0}_{9\times 4} & \mathbf{Q}_p \end{bmatrix} \text{ and } \mathbf{V}_J = \begin{bmatrix} \mathbf{V}_{\mathbf{y}_m} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{V}_p \end{bmatrix},$$

where $\mathbf{V}_{\mathbf{y}_m}$ and $\mathbf{V}_p$ are 3×3 diagonal matrices with the variances of the components of $\mathbf{y}_m$ and ${}^I\mathbf{X}_{p_m}$ in its diagonal.

By adopting a reasoning and notation similar to the ones in the end of section (III-A), a joint Kalman Filter that estimates the state of the target and the state of the UAV results. Since the measurements equation (10) is time varying, in opposition to the time invariant equation (6), when replacing the value of $\mathbf{C}_I$, in (7), by the value of $\mathbf{C}_J$, attention must be payed to the fact that $\mathbf{C}_J$ depends on time, i.e., $\mathbf{C}_I$ must be replaced by $\mathbf{C}_J(t_k)$. All other substitutions are straightforward.

According to the strategy described, estimates $[{}^I\hat{x}_b(t_k)\ \ {}^I\hat{y}_b(t_k)]^T$ of $[{}^I x_b\ {}^I y_b]^T$ at time instants t_k can be obtained from the first and third components of $\mathbf{x}_J^+(t_k)$. Moreover, estimates of the position of the UAV with respect to the inertial reference frame, at the same time instants, can be obtained from the 5-th, 8-th, and 11-th components of $\mathbf{x}_J^+(t_k)$.

C. Multiple-Model Adaptive Estimation

The model considered for the target, see expression (2), requires the knowledge of its turn rate. However, this value is not known in real applications, which led us to the adoption of a multiple-model based approach that simultaneously identifies some parameters of the system and estimates its state. The strategy implemented in this work, known as Multiple-Model Adaptive Estimation [10], considers several models for a system that differ in a set of parameters (in this case the target turn rate). For each one of these models, an isolated (see section III-A) or joint (see section III-B) Kalman Filter, depending on the strategy in use, is designed. The individual estimates provided by each filter are then combined using a weighted sum. By designing two Multiple-Model Adaptive Estimators, one based on a bank of isolated KFs and other based on a bank of joint KFs, two different strategies that estimate the position of a marine mammal moving with an unknown turn rate result.

The use of multiple-model approaches requires the definition of a criterion to determine the number of models that must be considered, the subsets (of the set in which the unknown parameter is assumed to be) associated with each model, and the nominal parameter value of each KF. With this purpose, the BPM [17] is used. A detailed description of the use of this technique in multiple-model architectures can be found in [9].

In order to find the number of models to use and the corresponding nominal parameter values, let the null turn rate, that corresponds to a straight trajectory, be the nominal parameter of one of the models. Then search the remaining parameter set for the turn rate nominal values that lead to a situation in which there is always a filter whose BPM, in relation to the filter based upon the true model, does no exceed a certain value. The boundaries of each subset are defined by the points of intersection of the BPM curves. For the system proposed, a total of $N = 4$ models results, with the nominal turn rate values presented in the sequel.

In order to gain some insight into how to choose the turn rate nominal values to use in each KF, a set of simulations were carried out. Figure 3 depicts the BPM for each one of the 4 models obtained for the isolated Kalman filtering approach. For the joint strategy, the same models are considered. The

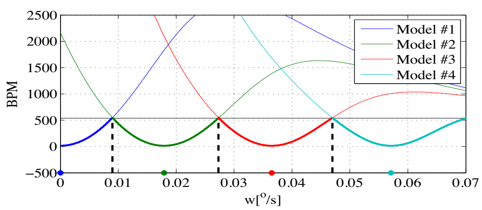


Fig. 3. BPM for the four models. Dots in different colours correspond to the turn rate values that minimize the BPM in each subset.

four turn rate nominal values used in each Kalman Filter are the ones that minimize the BPMs presented in Fig.3, $2\pi[0, 0.018, 0.037, 0.057]$ rad/s, which lead to the division of the original set ($\Omega = [0, 2\pi 0.07]$ rad/s) into the following subsets: $\Omega_1 = 2\pi[0, 0.009]$ rad/s, $\Omega_2 = 2\pi[0.009, 0.027]$ rad/s, $\Omega_3 = 2\pi[0.027, 0.047]$ rad/s, and $\Omega_4 = 2\pi[0.047, 0.070]$ rad/s. As can be seen, the regions of validity of the four subsets have different dimensions, which is a consequence of the use of the BPM instead of a methodology based on the Euclidean distance between the four nominal turn rate values.

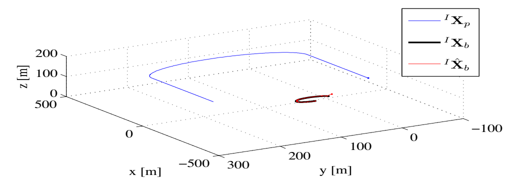
IV. Simulation Results

In this section, simulation results that illustrate the performance of the proposed isolated and joint MMAE Kalman Filters are presented. Experiments depicting the behavior of the system when the marine mammal submerges are provided.

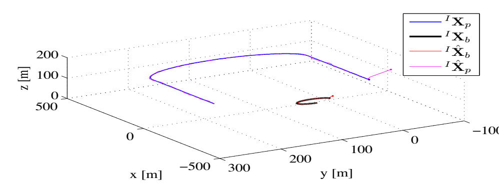
In order to keep simulations as close as possible to reality, the trajectories described by the UAV are generated according to the aircraft dynamic model *SymAirDyn*, proposed in [4]. The image acquisition module is modelled by a pinhole camera with position ${}^{p}\mathbf{X}_c = (200, 0, 100)$ and orientation ${}^{p}\mathbf{R}_c = {}^{p}\mathbf{R}_{c_0}\mathbf{R}(\alpha_p)\mathbf{R}(\theta_p)$ with respect to $\{P\}$, where $\mathbf{R}(\alpha_p)$, $\mathbf{R}(\theta_p)$, and ${}^{p}\mathbf{R}_{c_0}$ are rotation matrices that express rotations of α_p, θ_p, and $-\pi/2$ rad about the x, y, and z axes, respectively. In these expressions, $\mathbf{R}(\alpha_p)$ and $\mathbf{R}(\theta_p)$ denote the rotation matrices that express the camera pan and tilt movements (α_p and θ_p denote the camera pan and tilt angles, respectively), and ${}^{p}\mathbf{R}_{c_0}$ corresponds to the orientation of the camera with respect to $\{P\}$ when $\alpha_p = \theta_p = 0$ rad. The camera control strategy proposed by the authors in [18] was used to keep the target visible in acquired images.

In the simulations reported in this section, the sampling interval is $T = 0.2$s. Moreover, the power spectral densities $S_{w_b} = 10^{-3}$ m^2/s and $S_{w_p} = 10^{-3}$ m^2/s^3 are considered. The covariance matrices of the measurements provided to the isolated and joint filters are $\mathbf{V}_I = \text{diag}[10^2,\ 10^2]\,\text{m}^2$ and $\mathbf{V}_J = \text{diag}[10^2,\ 10^2,\ 10^2,\ 2^2,\ 2^2,\ 5^2]\,\text{m}^2$, respectively. The measurements of the yaw, pitch, and roll angles of the UAV, provided by the AHRS, are considered to be corrupted by zero-mean, white Gaussian noises with standard deviation $1°$, $0.5°$, and $0.5°$, respectively. An error superior to 50% in the state initial conditions of both filters is considered.

In Fig. 4, the real position of the marine mammal, ${}^{I}\mathbf{X}_b$, and UAV, ${}^{I}\mathbf{X}_p$, are depicted. The evolution over time of



(a) Isolated MMAE Kalman Filter.



(b) Joint MMAE Kalman Filter.

Fig. 4. Evolution over time of the real and estimated positions of the marine mammal and UAV.

the estimates ${}^{I}\hat{\mathbf{X}}_b$ of the target position, provided by both the isolated and joint filters, is also depicted, as well as the evolution of the estimates of the UAV position ${}^{I}\hat{\mathbf{X}}_p$ provided by the joint filtering strategy. In this simulation, the marine mammal describes a circular trajectory with the turn rate of the third model, i.e., $\omega = 0.037$ rad/s, and the UAV moves along the trajectory depicted in blue.

The performance of the proposed filters is addressed in Fig. 5. The values in brackets, σ_x and σ_y, correspond to the

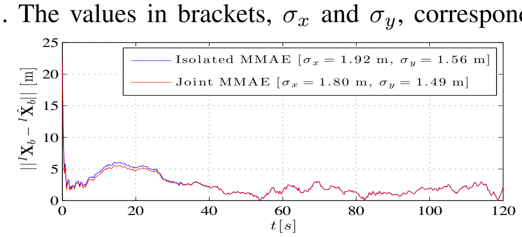


Fig. 5. Euclidean norm of the marine mammal position estimation error.

steady-state standard deviations of the errors of the estimates of ${}^{I}x_b$ and ${}^{I}y_b$. Both estimates converge to the vicinity of the real values. As can be seen, despite being very similar (the steady-state errors of the marine mammal position estimates have both standard deviations on the order of 2 m), the performance of the joint MMAE Kalman Filter is slightly better than the one of the isolated filter. This result is a consequence of the structure of the joint strategy, that resorts to estimates of the position of the UAV to help in estimating the position of the marine mammal. The main source of errors in these kind of experiments comes from the high altitude of the UAV with respect to the mammal position, moving at the sea surface. In this experiment, the UAV moved 150 m above the sea surface. Thus, small uncertainties in the measurements of its position and orientation, and as a consequence on the measurements of the position and orientation of the camera, have a strong impact on the estimates of the target position.

In Fig. 6, the performance of both filters when estimating the position of a marine mammal that moves along the trajectory in Fig. 4, but which submerges during the time interval $[40, 70]$ s, is depicted. As before, the values in brackets, σ_x

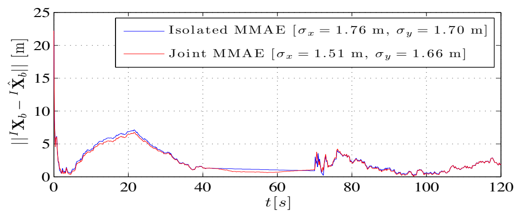
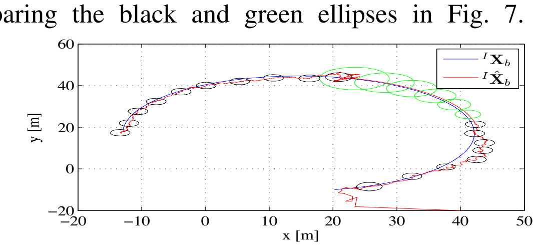
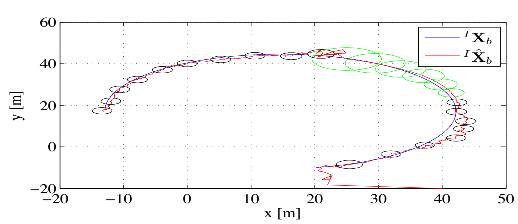


Fig. 6. Euclidean norm of the estimation error of the marine mammal position when there are occlusions.

and σ_y, correspond to the steady-state standard deviations of the errors of the estimates of ${}^I x_b$ and ${}^I y_b$, respectively. The performance of both filters is similar to their performance when the target is always visible in images acquired by the camera, since when the marine mammal submerges, estimates of its position are obtained resorting to the predict step of the Kalman Filters. When measurements are not used to improve the estimates provided by the predict step, the covariance of the estimation errors increases, as we can confirm by comparing the black and green ellipses in Fig. 7. In this



(a) Isolated MMAE Kalman Filter.



(b) Joint MMAE Kalman Filter.

Fig. 7. Evolution of the covariance of the target position estimation error when there are occlusions. Black and green ellipses correspond to moments when the target is visible by the camera and submerged, respectively.

figure, ellipses represent the covariance of the target position estimation error. The two semi-axes of these ellipses have direction and length given, respectively, by the eigenvectors and square root of the eigenvalues of the covariance matrices obtained from the entries of the covariance matrices estimated by the filters that correspond to the errors in the estimation of the target position.

V. CONCLUSIONS

This work presented a new strategy to track and locate targets moving at the ocean surface, resorting to measurements provided by a GPS, an Attitude and Heading Reference System (AHRS), and an image acquisition module, all installed on an Unmanned Aerial Vehicle (UAV). Measurements of the position of the target in acquired images were computed resorting to methodologies based on active contours, and were combined with measurements of the position and attitude of the UAV, provided by the GPS and AHRS, respectively, to obtain estimates of the position of the target with respect to an inertial reference frame. To obtain these estimates, two Kalman Filters were proposed. One, time-invariant, that estimates only the position of the target, and other, time-varying, that combines estimates of the position of the target with estimates of the position of the UAV to improve its performance. To assess the performance of the proposed strategies, a set of simulations, carried out under realistic conditions, were presented and discussed. In the near future, this system will be implemented on a real platform, and tracking and localization of real marine mammals will be pursued.

REFERENCES

[1] D. Zeller, "Ultrasonic telemetry: its application to coral reef fisheries," *Fisheries Bulletin*, vol. 97, pp. 1058–1065, 1999.
[2] B. Block et al., "Migratory movements, depth preferences, and thermal biology of atlantic bluefin tuna," *Science*, vol. 293, no. 5533, pp. 1310–1314, 2001.
[3] J. Sibert and J. Nielsen, *Electronic Tagging and Tracking in Marine Fisheries*. Springer, 2001.
[4] P. Rosa, C. Silvestre, D. Cabecinhas, and R. Cunha, "Autolanding controller for a fixed wing unmanned air vehicle," in *Proc. of AIAA Guidance and Control Conference*, 2007.
[5] P. Giraudeta and H. Glotin, "Real-time 3d tracking of whales by echo-robust precise tdoa estimates with a widely-spaced hydrophone array," *Applied Acoustics*, vol. 67, pp. 1106–1117, 2006.
[6] Y. Bar-Shalom, X. Rong-Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, Inc., 2001.
[7] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. part I: Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1333–1364, 2003.
[8] A. Gelb, *Applied Optimal Estimation*. Cambridge, Massachusetts: MIT Press, 2001.
[9] S. Fekri, M. Athans, and A. Pascoal, "Issues, progress and new results in robust adaptive control," *International Journal of Adaptive Control and Signal Processing*, vol. 20, pp. 519–579, Mar 2006.
[10] D. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Transactions on Automatic Control*, vol. 10, no. 4, pp. 434 – 439, Oct 1965.
[11] T. A. Patterson, L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos, "State-space models of individual animal movement," *Trends in Ecology and Evolution*, vol. 23, no. 2, p. 8, Feb 2008.
[12] O. Faugeras and Q. Luong, *The geometry of multiple images*. MIT Press, 2001.
[13] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
[14] F. Leymarie and M. D. Levine, "Tracking deformable objects in the plane using an active contour model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 617–634, 1993.
[15] B. Anderson and J. Moore, *Optimal Filtering*. USA: Englewood Cliffs: Prentice-Hall, 1979.
[16] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1994.
[17] Y. Baram, "Information, consistent estimation and dynamic system identification," Ph.D. dissertation, MIT, Cambridge, MA, USA, 1976.
[18] S. Brás, R. Cunha, J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear attitude observer based on active vision and inertial measurements," *IEEE Transactions on Robotics*, 2011.