Preliminary Results on Globally Asymptotically Stable Simultaneous Localization and Mapping in 3-D

Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre

Abstract— **This paper presents the design, analysis, performance evaluation, and preliminary experimental validation of a globally asymptotically stable (GAS) filter for simultaneous localization and mapping (SLAM) with application to unmanned aerial vehicles (UAVs). The SLAM problem is formulated in a sensor-based framework and modified in such a way that the system structure may be regarded as linear time-varying for observability purposes, from which a Kalman filter with GAS error dynamics follows naturally. The proposed solution includes the estimation of both body-fixed linear velocity and rate-gyro measurement biases. Both simulation results and preliminary experimental results, using an instrumented quadrotor equipped with a RGB-D camera, are included in the paper to illustrate the performance of the algorithm under realistic conditions.**

I. Introduction

Autonomous robot missions, including surveillance, critical infrastructure inspection, and search and rescue, raise the need for dependable navigation and relative positioning algorithms, particularly in environments where absolute positioning systems may not be used, either because of their absence or unreliability. These can take the form of aided localization algorithms, that make use of known characteristics of the environment such as maps or beacons typically employing ranging or, more recently, vision sensors (see [1] for an interesting view of the application of vision in robot navigation). There is, however, a more general solution to the problem of navigating an autonomous vehicle that does not require *a priori* knowledge of the environment: simultaneous localization and mapping (SLAM). Its importance to the research community is evident when looking at the wide set of solutions as diverse as EKF-SLAM, graph-based SLAM, or particle filters. A two-part interesting and thorough survey on SLAM techniques can be found in [2] and [3].

The algorithm proposed in this paper is a state-space SLAM filter and it is influenced by EKF-SLAM procedures such as [4], in which the filtering process is centered on the vehicle. The EKF approach uses a single filter to maintain estimates of the map and vehicle pose, as well as the cross-covariances. Even though it was one of the first solutions to the SLAM problem, this formulation has evidenced consistency and convergence problems studied in works like [5] and [6]. This paper proposes an alternate formulation that uses the linear time-varying Kalman filter and achieves global convergence results. Aside from the vehicle-centric approach of [4], the works of [7] and [8] are also related to the work here presented, as vision techniques are also employed: stereo images processed through SIFT for landmark detection in the former, and RGB-D images ran through SURF/SIFT algorithms in the latter.

The main contributions of this paper are the design, analysis, and validation of a novel sensor-based SLAM filter for tridimensional (3-D) environments. It is part of an integrated SLAM algorithm, extending the work in [9] and [10], by generalising many of its contributions. The proposed algorithm 1) has globally asymptotically stable (GAS) error dynamics; 2) resorts to the linear and angular motion kinematics, that are exact; 3) uses the low-cost *Microsoft Kinect*, in opposition to the 2-D landmark approach, which demands the use of considerably more expensive laser range finders; 4) builds on the well-established linear time-varying Kalman filter; and 5) explicitly estimates the rate-gyro bias.

The paper is organized as follows. Section II presents a short description of the problem, with the definition of the system dynamics. The observability analysis is performed in Section III. The filter design is described in Section IV including landmark detection, data association and loop closing procedures. Simulation results are presented in Section V and preliminary experimental results using an instrumented quadrotor are detailed in Section VI.

II. Description of the problem

This section presents the problem of designing a navigation system for a vehicle operating in an unknown environment. This problem is solved resorting to a novel SLAM algorithm, where no linearization or approximation is used. The only available sensors are a triaxial rate-gyro, and a RGB-D camera, such as the *Microsoft Kinect*, which provide angular rate measurements and RGB-D images, respectively, from where 3-D landmarks may be extracted.

A. Nonlinear System Dynamics

Let $\mathbf{R}(t) \in \mathrm{SO}(3)$ be the rotation matrix from the body-fixed frame $\{B\}$ to the inertial frame $\{I\}$, with $\dot{\mathbf{R}}(t) = \mathbf{R}(t)\ \mathbf{S}[\boldsymbol{\omega}(t)]$, where $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is the angular velocity expressed in body-fixed coordinates and $\mathbf{S}[\mathbf{a}]$ encodes the cross-product, i.e. $\mathbf{S}[\mathbf{a}]\mathbf{b} = \mathbf{a} \times \mathbf{b}$. Then, the position and velocity of a landmark expressed in the body-fixed frame, $\mathbf{p}_i(t) \in \mathbb{R}^3$ and $\dot{\mathbf{p}}_i(t) \in \mathbb{R}^3$, satisfy $\mathbf{p}_i(t) = \mathbf{R}^T(t)\ \left({}^I\mathbf{p}_i(t) - {}^I\mathbf{p}(t)\right)$ and

$$\dot{\mathbf{p}}_i(t) = -\mathbf{v}(t) - \mathbf{S}[\boldsymbol{\omega}(t)]\ \mathbf{p}_i(t), \tag{1}$$

respectively, where ${}^I\mathbf{p}(t) \in \mathbb{R}^3$ represents the vehicle position (as well as the origin of the body-fixed frame) in the inertial frame $\{I\}$ at time t, ${}^I\mathbf{p}_i(t) \in \mathbb{R}^3$ is the position of landmark $i \in \{1, \ldots, N_M\}$ expressed in the same frame,

This work was partially supported by the FCT [PEst-OE/EEI/LA0009/2011] and by the EU Project TRIDENT (Contract No. 248497).

The authors are with the Institute for Systems and Robotics, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal. Carlos Silvestre is also with the Department of Electrical and Computer Engineering, Faculty of Science and Technology of the University of Macau.

{plourenco,bguerreiro,pbatista,pjcro,cjs}@isr.ist.utl.pt

and $\mathbf{v}(t) \in \mathbb{R}^3$ denotes the velocity of the vehicle expressed in the body-fixed frame. Note that landmarks are assumed to be static in the inertial frame. It is important to notice that $\boldsymbol{\omega}(t)$ is available through noisy and biased rate-gyros measurements $\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_\omega(t) + \mathbf{n}_\omega(t)$, where the bias $\mathbf{b}_\omega(t) \in \mathbb{R}^3$ is assumed constant and $\mathbf{n}_\omega(t) \in \mathbb{R}^3$ corresponds to the rate-gyro noise, which is assumed to be zero-mean white Gaussian noise with standard deviation σ_{n_ω} in each component, i.e., $\mathbf{n}_\omega(t) \sim N(0, \sigma_{n_\omega}^2 \mathbf{I}_3)$. Taking this into account, and using the cross product property $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$, it is possible to rewrite (1) as

$$\dot{\mathbf{p}}_i(t) = -\mathbf{v}(t) - \mathbf{S}[\mathbf{p}_i(t)]\, \mathbf{b}_\omega(t) - \mathbf{S}[\boldsymbol{\omega}_m(t)]\, \mathbf{p}_i(t). \quad (2)$$

The vehicle-related variables, i.e., the linear velocity and the angular measurement bias will constitute the vehicle state, $\mathbf{x}_V(t) := \begin{bmatrix}\mathbf{v}^T(t) & \mathbf{b}_\omega^T(t)\end{bmatrix}^T \in \mathbb{R}^{n_V}$, with simple dynamics given by $\dot{\mathbf{x}}_V(t) = 0$, which means that both are assumed, in a deterministic setting, as constant. In the filtering framework, the inclusion of state disturbances allows to consider them as slowly time-varying.

It is now possible to derive the full state dynamics. For that purpose consider the position landmark dynamics (2), which may now be expressed as a function of the state vector, yielding

$$\dot{\mathbf{p}}_i(t) = \mathbf{A}_{MV_i}(\mathbf{p}_i(t))\, \mathbf{x}_V(t) - \mathbf{S}[\boldsymbol{\omega}_m(t)]\mathbf{p}_i(t),$$

where $\mathbf{A}_{MV_i}(\mathbf{p}_i(t)) = [-\mathbf{I}_3 \quad -\mathbf{S}[\mathbf{p}_i(t)]]$ and $\mathbf{I}_n$ is the identity matrix of dimension n. Finally, the N_O observed, also designated as visible, landmarks $\mathbf{x}_O(t) \in \mathbb{R}^{n_O}$ and the N_U unobserved or non-visible ones $\mathbf{x}_U(t) \in \mathbb{R}^{n_U}$ are concatenated in the landmark-based state vector, $\mathbf{x}_M(t) := \begin{bmatrix}\mathbf{x}_O^T(t) & \mathbf{x}_U^T(t)\end{bmatrix}^T \in \mathbb{R}^{n_M}$. The two state vectors here defined constitute the full state vector $\mathbf{x}_F(t) = \begin{bmatrix}\mathbf{x}_V^T(t) & \mathbf{x}_M^T(t)\end{bmatrix}^T$, with the full system dynamics reading

$$\begin{cases} \dot{\mathbf{x}}_F(t) = \mathbf{A}_F(t, \mathbf{x}_M(t))\mathbf{x}_F(t) \\ \mathbf{y}(t) = \mathbf{x}_O(t) \end{cases}, \quad (3)$$

with

$$\mathbf{A}_F(t, \mathbf{x}_M(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_M} \\ \mathbf{A}_{MV}(\mathbf{x}_M(t)) & \mathbf{A}_M(t) \end{bmatrix},$$

where

$$\mathbf{A}_{MV}(\mathbf{x}_M(t)) = \begin{bmatrix}\mathbf{A}_{MV_1}^T(\mathbf{p}_1(t)) & \cdots & \mathbf{A}_{MV_{N_M}}^T(\mathbf{p}_{N_M}(t))\end{bmatrix}^T,$$

$\mathbf{A}_M(t) = \operatorname{diag}(-\mathbf{S}[\boldsymbol{\omega}_m(t)], \cdots, -\mathbf{S}[\boldsymbol{\omega}_m(t)])$, and $\mathbf{0}_{n\times m}$ is a n by m matrix filled with zeros and, if m is omitted, the matrix is square. From (3) it follows that the system may be expressed in a way similar to the usual linear dynamical system form. However, it is obvious to conclude that the system above is nonlinear, as the dynamics matrix depends on the landmarks that may be visible or not.

B. Problem Statement

The problem addressed in this paper is the design of a SLAM filter in the space of the sensors, providing a sensor-based map and the velocity of the vehicle. The maps are represented by tridimensional position landmarks, which may also include up to 3 directions for each position. The pose of the vehicle is deterministic as it simply corresponds to the position and attitude of the body-fixed frame.

III. Observability Analysis

Observability is of the utmost importance in any filtering problem, and the work presented in this section aims at analysing the observability of the dynamical system previously exposed. It is important to notice that, although system (3) is inherently nonlinear, discarding the non-visible landmarks $\mathbf{x}_U(t)$ makes it possible to regard the resulting system as linear time-varying (LTV).

Consider the new state vector $\mathbf{x}(t) = \begin{bmatrix}\mathbf{x}_V^T(t) & \mathbf{x}_O^T(t)\end{bmatrix}^T$, which does not include the non-visible landmarks, for which the resulting system dynamics can be written as

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}(t, \mathbf{y}(t))\mathbf{x}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases}, \quad (4)$$

where $\mathbf{A}(t, \mathbf{y}(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \mathbf{A}_{MV_O}(\mathbf{y}(t)) & \mathbf{A}_{M_O}(t) \end{bmatrix}$ and $\mathbf{C} = [\mathbf{0}_{n_O \times n_V} \quad \mathbf{I}_{n_O}]$. Note that the matrix $\mathbf{A}(t, \mathbf{y}(t))$ depends not only on time but also on the system output. Nevertheless, the dependency on the system state is now absent and the system output is known, thus, the system can be seen as a linear time-varying system for observability analysis purposes. According to [11, Lemma 1, Section 3], if the observability Gramian associated with a system with a dynamics matrix depending on the system output is invertible, then the system is observable. This result will be used throughout this section. Before proceeding with this analysis the following assumption is needed.

Assumption 1: Any two detected position landmarks are assumed to be different and nonzero, i.e., $\mathbf{y}_i(t), \mathbf{y}_j(t) \neq \mathbf{0}$ and $\mathbf{y}_i(t) \neq \mathbf{y}_j(t)$ for all $t \geq t_0$ and $i, j \in \mathcal{I}_O$, where $\mathcal{I}_O$ denotes the set of visible landmarks.

It is important to notice that it is physically impossible to have two collinear visible landmarks at the same time, because of the intrinsic characteristics of the camera: the angle of view of the camera is always smaller than $180°$.

The following theorem states the analysis of the observability of system (4).

Theorem 1: Consider system (4) and let $\mathcal{T} := [t_0, t_f]$ and $\{t_1, t_2, t_3\} \in \mathcal{T}$. The system is observable on $\mathcal{T}$ in the sense that, given the system output, the initial condition is uniquely defined, if at least one of these conditions hold:

(i) There are, at least, three visible position landmarks at the same time t_1 that define a plane.
(ii) There exist two visible position landmarks in the interval $[t_1, t_2]$ such that at least one of the landmark sets $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_2(t_2)\}$ and $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_1(t_2)\}$ defines a plane.
(iii) There is a visible time-varying position landmark whose coordinates, $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$, define a plane.

Proof: The proof follows by transforming the system in question by means of a Lyapunov transformation [12, Chapter 1, Section 8], and then proving that the observability Gramian of the transformed system is non-singular.

Let $\mathbf{T}(t)$ be a Lyapunov transformation such that

$$\mathbf{z}(t) = \mathbf{T}(t)\, \mathbf{x}(t), \quad (5)$$

where $\mathbf{T}(t) = \operatorname{diag}(\mathbf{I}_{n_V}, \mathbf{R}_m(t), \cdots, \mathbf{R}_m(t))$ and $\mathbf{R}_m(t) \in \mathrm{SO}(3)$ is a rotation matrix respecting $\dot{\mathbf{R}}_m(t) = \mathbf{R}_m(t)\mathbf{S}[\boldsymbol{\omega}_m(t)]$. A Lyapunov transformation preserves the observability properties of a system, hence it suffices to prove that the new, transformed system is observable. This

approach has been used successfully in the past, see [11]. The computation of the new system dynamics and output is simple, yielding

$$\begin{cases} \dot{\mathbf{z}}(t) = \boldsymbol{\mathcal{A}}(t, \mathbf{y}(t))\mathbf{z}(t) \\ \mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\ \mathbf{z}(t), \end{cases} \tag{6}$$

with

$$\boldsymbol{\mathcal{A}}(t, \mathbf{y}(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \boldsymbol{\mathcal{A}}_{MV}(t, \mathbf{y}(t)) & \mathbf{0}_{n_O} \end{bmatrix},$$

and

$$\boldsymbol{\mathcal{C}}(t) = \begin{bmatrix} \mathbf{0}_{n_O \times n_V} & \text{diag}\left(\mathbf{R}_m^T(t), \cdots, \mathbf{R}_m^T(t)\right) \end{bmatrix},$$

where

$$\boldsymbol{\mathcal{A}}_{MV}(t, \mathbf{y}(t)) = \begin{bmatrix} \boldsymbol{\mathcal{A}}_{MV_1}^T(t, \mathbf{y}_1(t)) & \cdots & \boldsymbol{\mathcal{A}}_{MV_{N_O}}^T(t, \mathbf{y}_{N_O}(t)) \end{bmatrix}^T,$$

and

$$\boldsymbol{\mathcal{A}}_{MV_i}(t, \mathbf{y}_i(t)) = \begin{bmatrix} -\mathbf{R}_m(t) & -\mathbf{R}_m(t)\mathbf{S}[\mathbf{p}_i(t)] \end{bmatrix}.$$

Before proceeding to compute the observability Gramian associated with (6), it is necessary to know its transition matrix. A simple computation of $\mathbf{z}(t)$ as a function of $\mathbf{z}(t_0)$ by solving $\mathbf{z}(t) = \mathbf{z}(t_0) + \int_{t_0}^{t} \boldsymbol{\mathcal{A}}(\tau, y(\tau))\mathbf{z}(\tau)d\tau$ yields

$$\boldsymbol{\phi}(t, t_0) = \begin{bmatrix} \mathbf{I}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \boldsymbol{\phi}_{MV}(t, t_0) & \mathbf{I}_{n_O} \end{bmatrix},$$

where

$$\boldsymbol{\phi}_{MV}(t, t_0) = \begin{bmatrix} -\int_{t_0}^{t} \mathbf{R}_m(\sigma)d\sigma & -\int_{t_0}^{t} \mathbf{R}_m(\sigma)\mathbf{S}\left[\mathbf{p}_1(\sigma)\right] d\sigma \\ \vdots & \vdots \\ -\int_{t_0}^{t} \mathbf{R}_m(\sigma)d\sigma & -\int_{t_0}^{t} \mathbf{R}_m(\sigma)\mathbf{S}\left[\mathbf{p}_{N_O}(\sigma)\right] d\sigma \end{bmatrix}.$$

Finally, the observability Gramian is

$$\boldsymbol{\mathcal{W}}(t_0, t_f) = \int_{t_0}^{t_f} \boldsymbol{\phi}^T(\tau, t_0)\boldsymbol{\mathcal{C}}^T(\tau)\boldsymbol{\mathcal{C}}(\tau)\boldsymbol{\phi}(\tau, t_0)d\tau,$$

and, if $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is invertible, the system (6) is observable, in the sense that given the system input and output, the initial condition $\mathbf{z}(t_0)$ is uniquely defined. The next step is to prove, by contradiction, that this is the case, i.e, by assuming that $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is singular. In that case, there exists a unit vector $\mathbf{c} = \begin{bmatrix} \mathbf{c}_1^T & \mathbf{c}_2^T & \mathbf{c}_3^T & \cdots & \mathbf{c}_{2+N_O}^T \end{bmatrix}^T \in \mathbb{R}^{n_x}$, such that,

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0, t_f)\mathbf{c} = 0. \tag{7}$$

The objective now is to expand the above expression, in order to prove that there is no unit vector $\mathbf{c}$ that satisfies the equality. It is possible to see that

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0, t_f)\mathbf{c} = \int_{t_0}^{t_f} \left\|\mathbf{g}(\tau, t_0)\right\|^2 d\tau,$$

where $\mathbf{g}(\tau, t_0)$ and its derivative are given by

$$\mathbf{g}(\tau, t_0) := \begin{bmatrix} \boldsymbol{\phi}_{MV}(\tau, t_0) & \mathbf{I}_{n_O} \end{bmatrix} \mathbf{c},$$

and

$$\frac{d}{d\tau}\mathbf{g}(\tau, t_0) = \begin{bmatrix} -\mathbf{R}_m(\tau)\mathbf{c}_1 - \mathbf{R}_m(\tau)\mathbf{S}\left[\mathbf{p}_1(\tau)\right]\mathbf{c}_2 \\ \vdots \\ -\mathbf{R}_m(\tau)\mathbf{c}_1 - \mathbf{R}_m(\tau)\mathbf{S}\left[\mathbf{p}_{N_O}(\tau)\right]\mathbf{c}_2 \end{bmatrix}.$$

In order for (7) to be true, both $\mathbf{g}(\tau, t_0)$ and $\frac{d}{d\tau}\mathbf{g}(\tau, t_0)$ must be zero for all $\tau \in \mathcal{T}$, which implies that $\mathbf{c}_{2+i} = \mathbf{0}$ for all $i \in \mathcal{I}_O$ and

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}\left[\mathbf{p}_1(\tau)\right] \\ \vdots & \vdots \\ \mathbf{I}_3 & \mathbf{S}\left[\mathbf{p}_{N_O}(\tau)\right] \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \mathbf{0}, \quad \forall \tau \in \mathcal{T}. \tag{8}$$

Thus, it remains to show that, under the conditions of Theorem 1, the only possible solution is $\mathbf{c}_1 = \mathbf{c}_2 = \mathbf{0}$. Consider then the situation where there are three visible landmarks $\mathbf{p}_i(t_1)$, $i \in \{1, 2, 3\}$. In this case (8) can be rewritten as

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}\left[\mathbf{p}_1(t_1)\right] \\ \mathbf{0}_3 & \mathbf{S}\left[\mathbf{p}_2(t_1) - \mathbf{p}_1(t_1)\right] \\ \mathbf{0}_3 & \mathbf{S}\left[\mathbf{p}_3(t_1) - \mathbf{p}_1(t_1)\right] \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \mathbf{0}. \tag{9}$$

From this, it follows that either $\mathbf{c}_2 = \mathbf{c}_1 = \mathbf{0}$ or all three landmarks form a line, which contradicts the hypothesis of the theorem. Then, if condition (i) holds, $\mathbf{c}$ is not a unit vector. In the case where any of the remaining conditions applies, an equation similar to (9) may be constructed, this time with the sets $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_2(t_2)\}$ or $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_1(t_2)\}$, for condition (ii) and $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$ for condition (iii). Hence, if at least one of the conditions of Theorem 1 holds, then the observability Gramian is invertible on $\mathcal{T}$, and, using [11, Lemma 1, Section 3], it follows that (6) is observable. Moreover, as the Lyapunov transformation (5) preserves observability, the system (4) is also observable, thus concluding the proof of the theorem. ■

Given the sufficient conditions for observability, a Kalman Filter for the nonlinear system (3), with globally asymptotically stable error dynamics, can be designed following the classical approach. The following result addresses the equivalence between the state of the nonlinear system (4), regarded as LTV, and that of the nominal nonlinear system (3), when the non-visible landmarks are not considered.

Theorem 2: Consider that the conditions of Theorem 1 hold. Then,

(i) the initial state of the nonlinear system (3), discarding the non-visible landmarks, is uniquely determined, and is the same of the nonlinear system (4), regarded as LTV;
(ii) a state observer with uniformly globally asymptotically stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, with uniformly globally asymptotically stable error dynamics.

Proof: The proof of the first part of the theorem is made by comparison of the derivative of the output of the two systems in analysis, by computing $\mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\boldsymbol{\phi}(t, t_0)\mathbf{z}(t_0)$ and $\mathbf{y}(t) = \mathbf{C}_F \int_{t_0}^{t} \mathbf{A}_F(\mathbf{x}_F(\sigma), \sigma)\mathbf{x}_F(\sigma)d\sigma + \mathbf{C}_F\mathbf{x}_F(t_0)$, where $\mathbf{C}_F = [\mathbf{C} \quad \mathbf{0}_{n_O \times n_U}]$. It is omitted due to lack of space.

The first part of the theorem gives insight for the proof of the second part. An observer designed for a LTV system with GAS error dynamics has an estimation error convergent to zero, implying that the estimates asymptotically tend to the true state. Therefore, if the true state of the nonlinear system and the state of the LTV system are one and the same, the estimation error of the state observer for the nonlinear system converges to zero too. ■

Given this result, the design of a GAS observer for the LTV system follows. This step requires that the pair $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$ is uniformly completely observable as declared in [13]. The following theorem states the conditions under which this property is verified.

Theorem 3: Consider system (4). The pair $(\mathbf{A}(t,\mathbf{y}(t)),\mathbf{C})$ is uniformly completely observable, if there exists a $\delta > 0$ such that, for all $t \geq t_0$, it is possible to choose $\{t_1,t_2,t_3\} \in \mathcal{T}_\delta$, $\mathcal{T}_\delta = [t, t+\delta]$, such that at least one of the following conditions hold:

(i) There are at least three visible landmarks $\mathbf{p}_1(t_1)$, $\mathbf{p}_2(t_1)$ and $\mathbf{p}_3(t_1)$ such that they define a plane,
(ii) There exist two visible position landmarks at times t_1 and t_2 such that at least one of the landmark sets $\{\mathbf{p}_1(t_1),\mathbf{p}_2(t_1),\mathbf{p}_2(t_2)\}$ or the $\{\mathbf{p}_1(t_1),\mathbf{p}_2(t_1),\mathbf{p}_1(t_2)\}$ defines a plane,
(iii) There is a visible time-varying position landmark whose coordinates, at three different instants of time $\{t_1,t_2,t_3\}$, define a plane.

Proof: The proof follows similar steps to those of the proof of Theorem 1, considering uniform bounds for all $t \geq t_0$ and the intervals $\mathcal{T}_\delta := [t, t+\delta]$, and therefore it is omitted. The reader is referred to a similar proof with slightly different dynamics in [14]. ■

IV. SLAM Filter Design

This section addresses the design of the sensor-based 3D-SLAM filter. A discrete Kalman filter is designed, considering the sample-based/digital characteristics of both sensors needed for this work: an IMU (or more precisely a triad of rate-gyros) and a RGB-D camera (or other tridimensional relative position sensor). Hence, it is important to obtain the discrete-time version of the dynamic system under analysis.

1) Discrete dynamics: Denoting the synchronized sampling period of both sensors as T_s, the discrete time steps can be expressed as $t_k = kT_s + t_0$, where $k \in \mathbb{N}_0$ and t_0 is the initial time. Thus, the discretized system is characterized by the full state $\mathbf{x}_{F_k} := \mathbf{x}_F(t_k)$, the dynamics matrix $\mathbf{A}_{F_k} := \mathbf{A}_F(t_k,\mathbf{y}(t_k),\mathbf{x}_U(t_k))$ and the output matrix $\mathbf{C}_{F_k} := \mathbf{C}_F(t_k)$. Finally, the Euler discretization of the system dynamics (3), including system disturbance and measurement noise, yields

$$\begin{cases} \mathbf{x}_{F_{k+1}} = \mathbf{F}_k \mathbf{x}_{F_k} + \boldsymbol{\xi}_k \\ \mathbf{y}_{k+1} = \mathbf{H}_{k+1}\mathbf{x}_{F_{k+1}} + \boldsymbol{\theta}_{k+1} \end{cases} \tag{10}$$

where $\mathbf{F}_k = \mathbf{I}_{n_x} + T_s\mathbf{A}_{F_k}$ and $\mathbf{H}_{k+1} := \mathbf{C}_{F_{k+1}}$. The disturbance vector $\boldsymbol{\xi}_k$ and the measurement noise vector $\boldsymbol{\theta}_k$ are both zero-mean discrete white Gaussian noise, with $\langle \boldsymbol{\xi}_k \boldsymbol{\xi}_k^T \rangle = \boldsymbol{\Xi}_k$ and $\langle \boldsymbol{\theta}_k \boldsymbol{\theta}_k^T \rangle = \boldsymbol{\Theta}_k$, where $\langle . \rangle$ denotes the expected value of its arguments. Note that this system includes the non-visible landmarks, which are propagated in open-loop.

2) Prediction Step: The prediction step uses the full system (10) and the measurements of the rate-gyros, propagating the state every time a reading is available. Its equations are standard, and are therefore omitted.

3) Update Step: This step occurs every time 2-D colour and depth images are available from the Kinect. Then, an implementation of SURF [15] detects features in the 2-D picture of the environment that are matched to a pointcloud built with the depth image. This matching returns a set of observed tridimensional landmarks in cartesian coordinates. The SLAM filter does not know a priori if a landmark from this set is in the current map or if it is the first time it is seen. This is when data association takes place, associating the measured data with the existing landmarks. The algorithm used, the Joint Compatibility Branch and Bound [16], performs a depth-first search only expanding nodes when the joint associations are jointly compatible in a probabilistic sense. Both the landmark detection and association algorithms may be substituted by others as they are independent from the filtering technique. The association algorithm provides the innovation vector and its covariance matrix, and also redefines the new sets of visible and non-visible landmarks. The update equations are standard.

4) Loop closing: Loop closing enables the recognition of previously visited places, allowing the reduction of the uncertainty associated with the landmarks. The rather naive loop closing algorithm used consists of using only a subset of the state landmarks in the association algorithm, namely the more recent ones $\mathcal{I}_{rec}$, and separating the full state into three subsets, the recent, the old, $\mathcal{I}_{old}$, and the ones in between, $\mathcal{I}_{gap}$. This allows the duplication of landmarks when an area is revisited. Then, periodically, the algorithm tries to associate landmarks in $\mathcal{I}_{rec}$ and $\mathcal{I}_{old}$ using an adapted version of the association algorithm. If the number of jointly compatible associations passes a certain predefined threshold, a loop closure takes place and is incorporated in the filter by means of a noise free measurement.

V. Simulation results

The simulated environment consists of 70 landmarks spread throughout a $16\mathrm{m}\times16\mathrm{m}\times3\mathrm{m}$ map, including a closed 2 m wide corridor in outer borders of the map. The trajectory is simply a loop through the corridors at half-height, with the vehicle starting on the floor.

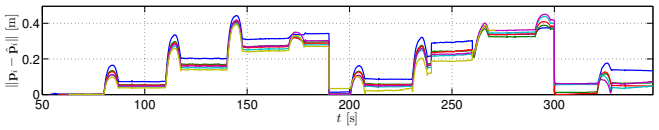
1) SLAM filter parameters: The SLAM filter parameters include the usual Kalman filter parameters, as well as time thresholds for the sets, and tuning knobs for the loop closure and state maintenance procedures. The output noise covariance is $\boldsymbol{\Theta}_k = 10^{-3}\mathbf{I}_3$ and the state disturbance covariance is given by $\boldsymbol{\Xi}_k = T_s \operatorname{diag}\left(25\mathbf{I}_3, 10^{-6}\mathbf{I}_3, 10^{-4}\mathbf{I}_3, \ldots, 10^{-4}\mathbf{I}_3\right) 10^{-4}$. The initial estimates of velocity and angular bias are set to zero. As to the SLAM-specific parameters, the recent landmark set is composed by landmarks seen in the last 15 s, and the old by landmarks not seen in more than 100 s. The first loop closure is tried at 100 s and it is triggered if at least 6 landmarks are associated. Finally, any landmark not visible for more than 200 s is discarded.

2) Results: The simulation starts with the vehicle stopped for 50 seconds, which then takes off and circles through the map for around 350 seconds at an average speed of 0.45 $\mathrm{m/s}$. The zero-mean noise added to the angular velocity measurements is normal-distributed with a standard deviation of $\sigma_{\omega_m} = 5 \times 10^{-4}$ $\mathrm{rad/s}$ at each coordinate, and the noise included in the landmark observations is also zero-mean Gaussian white noise with a standard deviation of $\sigma_y = 10^{-3}$ m.

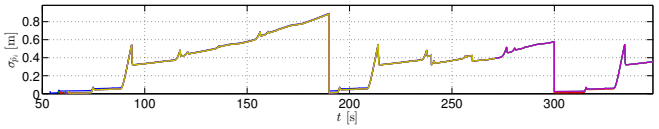
Firstly, the Kalman filter performance can be evaluated through Fig. 1 that depicts the evolution of 5 landmarks, with Fig. 1(b) showing the standard deviation of each sensor-based landmark growing to around 1m, when a loop closure is triggered at $t = 190$ s and the uncertainty diminishes considerably. In Fig. 1(a) the norm of estimation error of those 5 landmarks is shown. The velocity estimation error has mean under 10^{-3} $\mathrm{m/s}$ with standard deviation below 5×10^{-3} $\mathrm{m/s}$, and both the mean and standard deviation

of the estimation error of the angular measurement bias are lower than 10^{-4} rad/s.

The statistics regarding the number of landmarks can be found in Fig. 2. In blue, the number of visible landmarks. The stems represent loop closure trials and events, and the dashed line the minimum number of landmark associations necessary for loop closing. In the whole simulation, the state can contain over 90 landmarks.



(a) Estimation error norm.



(b) Standard deviation.

Fig. 1. Estimation error and standard deviation of the first 5 landmarks.

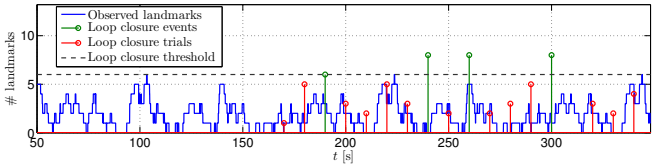
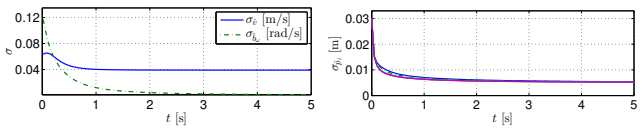


Fig. 2. Evolution of the number of landmarks used in the Kalman filter (in blue), and the number of visible landmarks (in green). Loop closure trials, threshold and events also present.



(a) Vehicle variables. (b) State landmarks.

Fig. 3. Evolution of the standard deviations of all the variables in the first 20 seconds of simulation.

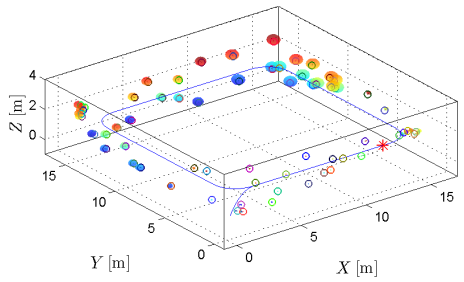


Fig. 4. Picture of estimated map rotated and translated using the true parameters.

The convergence of the standard deviation of the state variables when the observability conditions are satisfied is presented in Fig. 3. In these first 5 seconds of simulation, the vehicle is immobilized on the ground, and 5 different landmarks are visible.

Finally, the estimated map at $t = 200s$, $10s$ after a loop closing event, rotated and translated using the true quantities along the true trajectory is shown in Fig. 4. The coloured ellipsoids represent the 2σ bounds of each landmark, the red star denotes the position of the vehicle at the time, and the small circles mark the real positions. Note that the older landmarks have greater uncertainty, and that the landmarks closer to the position of the vehicle (the ones more affected by the loop closure) have low uncertainty.

This simulation was designed to allow the practical validation of the consistency of the algorithm, by exposing the vehicle to previously visited terrain after exploring new areas, in order to trigger a loop closing. The results show that the sensor-based map is consistent, allowing the loop to be closed repeatedly (see Fig. 2). Moreover, the simulation results here presented demonstrate that the uncertainty is coherent with the estimation errors, as shown in Fig. 1.

VI. Preliminary Experimental Results

The simulation results were consolidated by a preliminary experiment in the Sensor-Based Cooperative Robotics Research Laboratory - SCORE Lab of the Faculty of Science and Technology of the University of Macau.

The experimental setup consists of an *AscTec Pelican* quadrotor, which is equipped with an *Intel Atom* processor board, and into which was added a *Microstrain 3DM-GX3-25* inertial measurement unit working at 200Hz and a *Microsoft Kinect* camera, at 30Hz. The experiment consisted in moving the quadrotor inside a $6\text{m}\times6\text{m}$ room (usable area of 16m^2). The room was equipped with a *VICON* motion capture system, which provides accurate estimates of the position, attitude, linear and angular velocities of any vehicle placed inside the working area with the correct markers. Figure 5 represents the flow of information in the algorithm used in the experimental setup. The last block named *Inertial*

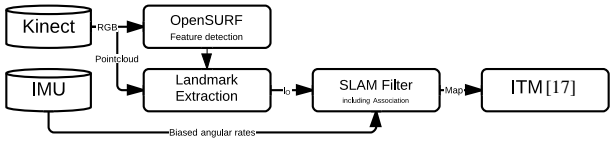


Fig. 5. The full algorithm used in the experiments.

Trajectory and Map estimation represents the algorithm presented in [17] that estimates the vehicle pose and the map in the inertial frame using only the sensor-based map.

In the first 15 seconds the vehicle was stopped and in the following it was hand-driven in a small lap around the room. Figure 6 shows the comparison of the estimation of the inertial trajectory, provided by the ITM algorithm with the ground truth provided by *VICON*, in blue. It can be seen that the estimated trajectory follows very closely the true trajectory of the vehicle.

Figure 7 compares the ground truth with the estimation of the body-fixed velocity of the vehicle. The former is obtained by rotating the inertial estimate by the *VICON* system with the rotation matrix it provides. Again, the estimates follow the ground truth very closely, aside from the fact that the *VICON* estimate is somewhat noisier.

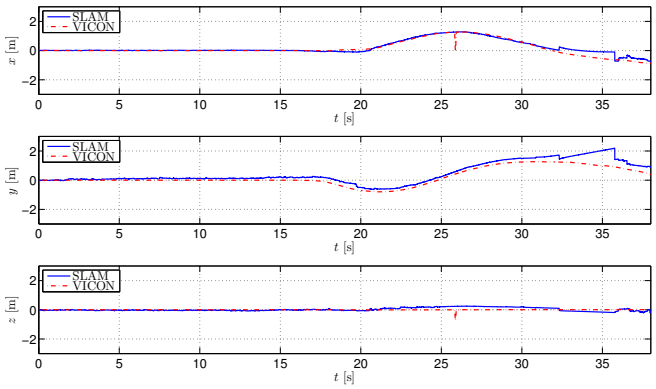


Fig. 6. Time evolution of the real and estimated trajectory

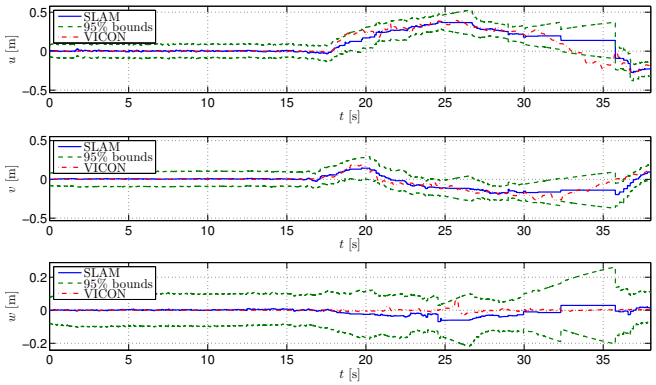


Fig. 7. Time evolution of the real and estimated velocity

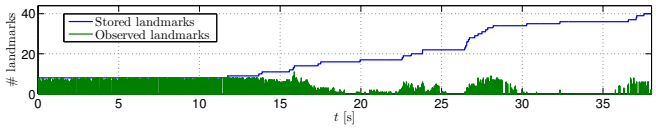


Fig. 8. Evolution of the number of landmarks used in the Kalman filter (in blue) and the visible (in green).

Finally, the evolution of the number of landmarks involved in the algorithm are shown in Fig. 8. In blue, the number of landmarks in the SLAM filter state are presented. Note the constant number while the vehicle is stopped. The number of visible landmarks in each instant is shown in green. Note that just after the first 30 s the number of observation instants is small (due to technical problems with the equipment) and the number of visible landmarks is also very small, which explains the degradation in the estimation that occurs around that time.

VII. Conclusions and Future Work

This paper presented the design, analysis, simulation, and preliminary experimental validation of a novel globally asymptotically stable sensor-based SLAM filter. The framework in which all work is based is intended to filter within the space of the sensors, thus avoiding the attitude representation in the filter state. The main focus of this work was the observability analysis, which provided theoretical results in observability and, subsequently, on the convergence of the error dynamics of the proposed nonlinear system. Furthermore, the performance and consistency of the algorithm were validated in simulation showing the convergence of the uncertainty in every variable except the non-visible landmarks, as well as the production of a consistent map which allows the closure of a loop with nearly 50 meters. Preliminary experimental results, with ground truth data, showed also the good performance of the SLAM filter and hinted at the good estimation of the transformation to the inertial frame.

The complete algorithm involves two different main blocks, the SLAM filter on one side, and an Inertial Trajectory and Map Estimation algorithm on the other, using as inputs the sensor-based estimated map [17]. Future work may also address the inclusion of landmark directions and the establishment of necessary conditions in the observability analysis.

References

[1] G. N. DeSouza and A. C. Kak, "Vision for Mobile Robot Navigation: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, 2002.

[2] H. Durrant-Whyte and T. Bailey, "Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[3] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robotics & Automation Magazine*, vol. 13, no. 3, pp. 108–117, 2006.

[4] J. Castellanos, R. Martinez-Cantin, J. Tardós, and J. Neira, "Robocentric map joining: Improving the consistency of EKF-SLAM," *Robotics and Autonomous Systems*, vol. 55, no. 1, pp. 21 – 29, 2007.

[5] S. Huang and G. Dissanayake, "Convergence and Consistency Analysis for Extended Kalman Filter Based SLAM," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 1036 –1049, October 2007.

[6] S. Julier and J. Uhlmann, "A counter example to the theory of simultaneous localization and map building," in *Proc. of the 2001 IEEE Int. Conf. on Robotics and Automation (ICRA)*, vol. 4, Seul, South Korea, May 2001, pp. 4238–4243.

[7] S. Se, D. Lowe, and J. Little, "Mobile Robot Localization and Mapping with Uncertainty using Scale-invariant Visual Landmarks," *The International Journal of Robotics Research*, vol. 21, no. 8, pp. 735–758, 2002.

[8] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An Evaluation of the RGB-D SLAM System," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, St. Paul, MA, USA, May 2012.

[9] B. Guerreiro, P. Batista, C. Silvestre, and P. Oliveira, "Sensor-based Simultaneous Localization and Mapping - Part I: GAS Robocentric Filter," in *Proc. of the 2012 American Control Conference*, Montréal, Canada, Jun. 2012, pp. 6352–6357.

[10] ——, "Sensor-based Simultaneous Localization and Mapping - Part II: Online Inertial Map and Trajectory Estimation," in *Proc. of the 2012 American Control Conference*, Montréal, Canada, Jun. 2012, pp. 6334–6339.

[11] P. Batista, C. Silvestre, and P. Oliveira, "Single range aided navigation and source localization: Observability and filter design," *Systems & Control Letters*, vol. 60, no. 8, pp. 665 – 673, 2011.

[12] R. Brockett, *Finite Dimensional Linear Systems*, ser. Series in decision and control. John Wiley & Sons, 1970.

[13] B. Anderson, "Stability properties of Kalman-Bucy filters," *Journal of the Franklin Institute*, vol. 291, no. 2, pp. 137 – 144, 1971.

[14] P. Batista, C. Silvestre, and P. Oliveira, "Sensor-based Complementary Globally Asymptotically Stable Filters for Attitude Estimation," in *Proc. of the 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 7563–7568.

[15] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346 – 359, 2008, Similarity Matching in Computer Vision and Multimedia.

[16] J. Neira and J. Tardós, "Data Association in Stochastic Mapping Using the Joint Compatibility Test," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 890 –897, dec 2001.

[17] P. Lourenço, B. J. Guerreiro, P. Batista, P. Oliveira, and C. Silvestre, "3-D Inertial Trajectory and Map Online Estimation: Building on a GAS Sensor-based SLAM filter," in *European Control Conference 2013*, July 2013, accepted.