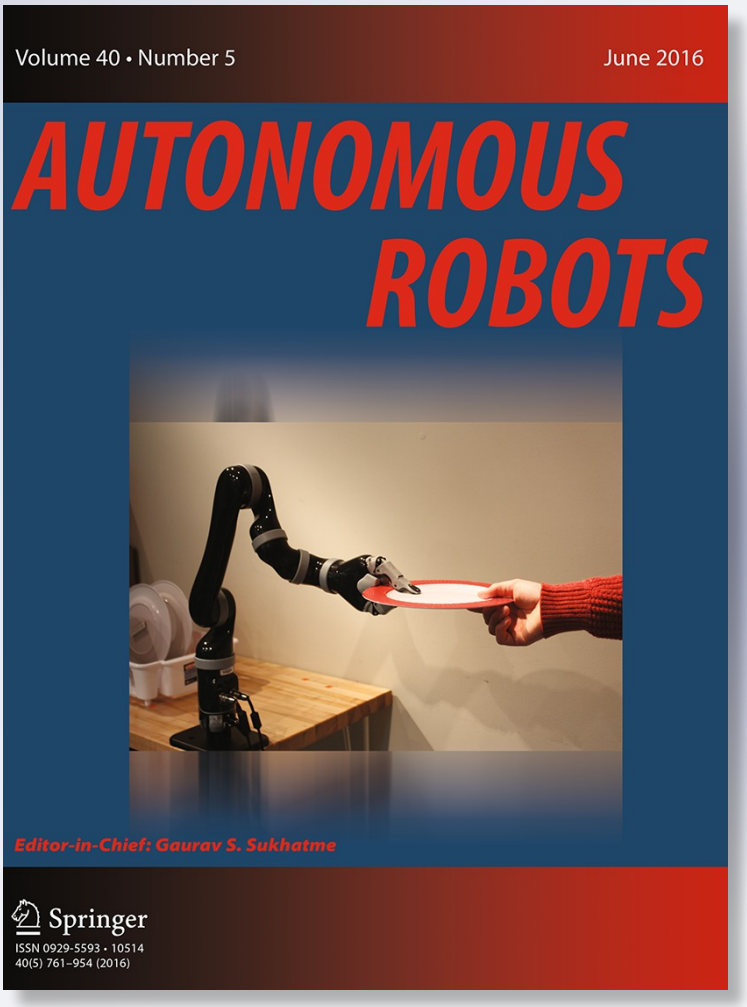
Simultaneous localization and mapping for aerial vehicles: a 3-D sensor-based GAS filter

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

Autonomous Robots

ISSN 0929-5593
Volume 40
Number 5

Auton Robot (2016) 40:881-902
DOI 10.1007/s10514-015-9499-z



Springer

Auton Robot (2016) 40:881–902
DOI 10.1007/s10514-015-9499-z

Simultaneous localization and mapping for aerial vehicles: a 3-D sensor-based GAS filter

Pedro Lourenço[1] · **Bruno J. Guerreiro**[2,4] · **Pedro Batista**[1,2] · **Paulo Oliveira**[1,3,5] · **Carlos Silvestre**[1,4]

Received: 2 March 2014 / Accepted: 9 September 2015 / Published online: 24 September 2015

Abstract This paper presents the design, analysis, and experimental validation of a globally asymptotically stable (GAS) filter for simultaneous localization and mapping (SLAM) with application to unmanned aerial vehicles. The main contributions of this paper are the results of global convergence and stability for SLAM in tridimensional (3-D) environments. The SLAM problem is formulated in a sensor-based framework and modified in such a way that the 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. Experimental results from several runs, 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.

Keywords Simultaneous localization and mapping · Aerial robotics · 3-D mapping · Sensor fusion · RGB-D camera

✉ Pedro Lourenço
plourenco@isr.tecnico.ulisboa.pt

Bruno J. Guerreiro
bguerreiro@isr.tecnico.ulisboa.pt

Pedro Batista
pbatista@isr.tecnico.ulisboa.pt

Paulo Oliveira
pjcro@isr.tecnico.ulisboa.pt

Carlos Silvestre
csilvestre@umac.mo

1 Institute for Systems and Robotics, Laboratory of Robotics and Engineering Systems, Lisbon, Portugal

2 Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

3 LAETA - Associated Laboratory for Energy, Transports and Aeronautics, Institute of Mechanical Engineering, Lisbon, Portugal

4 Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, China

5 Institute of Mechanical Engineering, Associated Laboratory for Energy, Transports and Aeronautics, Lisbon, Portugal

1 Introduction

Autonomous robot missions, such as surveillance, search and rescue, or critical infrastructure inspection, raise the need for dependable navigation and relative positioning algorithms. These are particularly important in environments where absolute positioning systems such as GPS may not be used either due to their absence or unreliability. Aided localization algorithms are a possible solution for this problem, and make use of known characteristics of the environment such as maps or beacons typically employing acoustic or radio ranging, or, more recently, light detection and ranging (LIDAR) and vision sensors. However, when the system does not assume any a priori knowledge of the environment, a far more complex problem must be considered, usually known as the simultaneous localization and mapping (SLAM) problem: to navigate a vehicle in an unknown environment while building a map and localizing itself within that map.

1.1 Related work

Since its introduction in the scientific community during the 1980s, SLAM has been the subject of many research efforts, interestingly and thoroughly reported in the two-part survey

found in Durrant-Whyte and Bailey (2006) and Bailey and Durrant-Whyte (2006). There is a vast diversity of research topics in SLAM, all stemming from the myriad of filtering techniques, mapping sensors, and mission profiles associated with it. In the light of this work, the most relevant ones are the application of vision to SLAM, given that vision techniques such as landmark detection in RGB-D images are utilized here, the consistency and convergence of SLAM algorithms, and the sensor-based (or robocentric) approach to SLAM.

The applications of vision to SLAM have flourished in recent years, after much attention and effort was put into making the most of the potentialities of laser ranging sensors. Although both cameras and LIDARs have become more and more affordable, the latter is still considerably more expensive than the former. That fact, combined with very good results obtained in feature detection and image processing [for which the SURF, presented in Bay et al. (2008), is an example], led to the increasing interest in vision-based SLAM. There are four main approaches in the literature differing in the number or type of cameras: mono (Weiss et al. 2013), stereo (Se et al. 2002), time-of-flight (May et al. 2009), and RGB-D—a camera with a depth sensor (Endres et al. 2012). All these algorithms utilize feature extraction and association to detect interest points in the collected images and associate them in two different frames (either from different cameras or at different time instants). Some use the great number of extractable features in two different time instants to compute the camera poses in a visual-inertial nonlinear optimization process, as Leutenegger et al. (2013), while others, such as Se et al. (2002) and Davison et al. (2007), employ a filtering framework to track landmarks. The combined use of cameras and sensors such as inertial measurement units (IMU) raises the need for proper calibration of camera parameters as well as rate-gyros and accelerometer bias. This topic is addressed in Jones and Soatto (2011) and Kelly and Sukhatme (2011) as visual-inertial sensor fusion using extended or unscented Kalman filters.

Even though there have been several topics of intense research within the scope of SLAM, there is still important work to do regarding topics such as consistency or convergence. For example, due to the nonlinear nature of the problem, the widely used extended Kalman filter has well-known problems of consistency and optimistic estimates (see Bailey et al. (2006) and Julier and Uhlmann (2001) for an analysis of this issue) due to the use of linearized equations evaluated at the estimated states that can be erroneous and mislead the filter. In order to cope with this inherent characteristic of EKF-SLAM, the authors of Huang et al. (2010) investigate the unobservable space of the linearized error-state system which they show is of smaller dimensionality than the actual unobservable space of the underlying nonlinear system, thus obtaining spurious information from a direction where none is available. They then argue that selecting the linearization points of the EKF to ensure that both unobservable spaces are of the same dimension improves the consistency of the filter, and propose a constrained optimization problem with a closed-form solution that minimizes the linearization error subject to conditions that ensure an unobservable subspace of appropriate dimensions.

Apart from dealing with consistency problems, there are also some notable results regarding the convergence of EKF-SLAM [such as Huang and Dissanayake (2007) and Bishop and Jensfelt (2009)]. These study the evolution of the covariance of the filter and derive conditions in which the covariances of the filter converge, as long as it is consistent. From a different point of view, Jones and Soatto (2011) and Kelly and Sukhatme (2011) address the observability properties of nonlinear SLAM, as it is a necessary (but not sufficient) condition for the convergence of any filtering algorithm. In Jones and Soatto (2011) an algorithm with monocular vision and inertial measurements is proposed, aiming to perform SLAM and calibration of parameters such as the local gravity vector and the six degrees-of-freedom IMU-camera rigid transformation, whereas in Kelly and Sukhatme (2011) the authors also estimate the biases of the accelerometer and rate-gyro measurements in an algorithm more focused on calibration. They perform a thorough observability analysis based on differential geometry considerations resulting in conditions on the vehicle trajectory for local weak observability of the underlying nonlinear system [see Hermann and Krener (1977) for further details]. All these represent important advances in the study of convergence and observability of SLAM algorithms.

The simultaneous localization and mapping problem can be formulated in several different ways, and many approaches follow the idea of performing the estimation in an Earth-fixed frame. Nevertheless, the sensor-based approach has precedents in SLAM history, most importantly the robocentric idea in Castellanos et al. (2007), designed to tackle the consistency problems while still considering the estimation of the (unobservable) incremental pose in the filter state. In the light of this, the work in Guerreiro et al. (2013) proposed a different solution where the filter is purely sensor-based, suppressing pose representation in the state and therefore avoiding its pitfalls—an approach deeply rooted in the well-established theory of sensor-based control which explores the fact that vehicle-fixed sensors provide measurements that otherwise would need to be transformed to an earth-fixed frame [see Weiss et al. (1987), Kermorgant and Chaumette (2011), and references therein for more on this subject]. This means that error variables are derived directly in the sensor space and therefore use only body-fixed quantities. These can be provided directly by sensors, as in Victorino et al. (2000), or by state observers such as sensor-based SLAM. This strategy countermands consistency problems and finally

provides theoretical guarantees of global convergence for two-dimensional SLAM, paving the way for the contributions in this paper.

Despite the great research efforts devoted to the SLAM problem as documented previously, a formal theoretical result on global convergence for a three-dimensional SLAM algorithm is absent from the literature to the best of the authors' knowledge. Addressing that problem is the focus of this paper.

1.2 Contributions

The algorithm proposed in this paper is rooted in EKF-SLAM procedures such as Castellanos et al. (2007), 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. This paper proposes an alternate formulation that uses a Kalman filter and achieves global convergence results by exploring the linear time-varying nature of the sensor-based SLAM system and analysing its observability. Differing from the observability-heavy works in Jones and Soatto (2011) and Kelly and Sukhatme (2011) that tackle monocular vision SLAM with angular rate and accelerometer measurements mentioned previously, this sensor-based system assumes that 3-D landmark positions and the angular rates are measured, and only includes the local relative map, the linear velocity and the rate-gyro bias. It does not include accelerometer readings, thus avoiding the need to estimate its bias and the gravity vector and rendering the problem less complex. However, and most importantly, the observability results here presented are a sufficient condition for the global asymptotical stability of the sensor-based SLAM filter, the main contribution of this paper. This relieves the filter design process of the careful consideration usually devoted to filter initialization. The sensor-based approach on SLAM has been addressed by the authors in previous work. The conference version of this paper, Lourenço et al. (2013) built on the work in Guerreiro et al. (2013) by generalising its contributions, and presented a sensor-based SLAM filter with sufficient theoretical guarantees for the observability of a nonlinear system with preliminary experimental results. This filter was complemented by a second paper, Lourenço et al. (2013), in which a methodology was proposed to obtain Earth-fixed estimates of the pose of the vehicle and of the landmark map using only the sensor-based map. This paper extends this work by presenting several new results on the observability of the underlying nonlinear system. Sufficient and necessary conditions are introduced for observability, global stability and convergence. Furthermore, new and comprehensive experimental results are detailed to illustrate the performance of the algorithm in real-world conditions. The proposed globally asymptotically stable (GAS) sensor-based SLAM filter: (i) resorts to the linear and angular motion kinematics, which are exact; (ii) uses the low-cost Kinect, in opposition to the bidimensional (2-D) landmark approach, which demands the use of considerably more expensive laser range finders; (iii) builds on the well-established Kalman filter for linear time-varying systems; and (iv) explicitly estimates the rate gyro measurement bias.

1.3 Paper organization and notation

The paper is organized as follows. Section 2 presents a short description of the problem, with the definition of the system dynamics. The observability analysis is performed in Sect. 3. The filter design is described in Sect. 4 including landmark detection, data association, and loop closing procedures. Experimental results using an instrumented quadrotor are detailed in Sect. 6, whereas in Sect. 7 the concluding remarks and directions for further research are provided.

The following symbol convention is used in this paper: vectors are represented in bold small letters, matrices in bold capital letters and scalar symbols are expressed in italic, constants by capital letters, and variables in small letters. The superscript ^{E} indicates a vector or matrix expressed in the Earth-fixed frame $\{E\}$. For the sake of clarity, when no superscript is present, the vector is expressed in the body-fixed frame $\{B\}$. $\mathbf{I}_n$ is the identity matrix of dimension n, and $\mathbf{0}_{n\times m}$ is a n by m matrix filled with zeros. If m is omitted, the matrix is square. $\mathbf{S}[\mathbf{a}]$ is a special skew-symmetric matrix, henceforth called the cross-product matrix, as $\mathbf{S}[\mathbf{a}]\mathbf{b} = \mathbf{a}\times\mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

2 Problem formulation

Building on the sensor-based idea explained in the previous section, this paper addresses the problem of designing a navigation system in a sensor-based framework for an aerial vehicle capable of sensing landmarks in a previously unknown environment. This is done resorting to a novel SLAM algorithm, where no linearization or approximation is used whatsoever. The only available sensors are a triaxial rate gyro and an RGB-D camera, such as the *Microsoft Kinect*, which provide angular rate measurements and RGB-D images, from where 3-D landmarks may be extracted.

2.1 Nonlinear system dynamics

Let the pair $\left(\mathbf{R}(t), {}^E\mathbf{p}(t)\right) \in \mathrm{SO}(3)\times\mathbb{R}^3$ encode the transformation from the body-fixed frame $\{B\}$ to an Earth-fixed frame $\{E\}$. $\mathbf{R}(t)$ is a rotation matrix satisfying $\dot{\mathbf{R}}(t) = \mathbf{R}(t)\mathbf{S}[\boldsymbol{\omega}(t)]$, where $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is the angular velocity, expressed in body-fixed coordinates, and ${}^E\mathbf{p}(t)$ represents the vehicle position (as well as the origin of the body-fixed frame) in the Earth-fixed frame. Consider also the existence of natural static

landmarks in the environment whose coordinates can be perceived by the vehicle. 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

$$\mathbf{p}_i(t) = \mathbf{R}^T(t) \left({}^E\mathbf{p}_i(t) - {}^E\mathbf{p}(t) \right)$$

and

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

respectively, where ${}^E\mathbf{p}_i(t) \in \mathbb{R}^3$ is the position of landmark i expressed in the Earth-fixed frame, assuming ${}^E\dot{\mathbf{p}}_i(t) = 0$, and $\mathbf{v}(t) \in \mathbb{R}^3$ denotes the velocity of the vehicle, relative to the Earth-fixed frame, expressed in the body-fixed frame.

Consider that the vehicle is equipped with a triad of orthogonally mounted rate gyros, rendering the angular velocity of the vehicle available through the biased rate gyros measurements $\boldsymbol{\omega}_m(t)$ given by

$$\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_\omega(t),$$

where the bias $\mathbf{b}_\omega(t) \in \mathbb{R}^3$ is assumed constant. Taking this into account, and using the cross product property $\mathbf{S}[\mathbf{a}]\mathbf{b} = -\mathbf{S}[\mathbf{b}]\mathbf{a}$, it is possible to rewrite (1) as

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

The variables presented so far can be divided in the vehicle-related or landmark-related categories. The former, i.e., the linear velocity and the angular measurement bias will constitute the vehicle state, denoted by $\mathbf{x}_V(t) := \left[\mathbf{v}^T(t)\ \mathbf{b}_\omega^T(t)\right]^T \in \mathbb{R}^{n_V}$, 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, however, the inclusion of state disturbances allows to consider them as slowly time-varying.

The landmarks define the map and can be separated in two complementary sets, $\mathcal{I}_O$ and $\mathcal{I}_U$. The former contains the N_O observed, also designated as visible, landmarks, while the latter contains the N_U unobserved or non-visible. The landmarks in each set can be agglomerated in state vectors denoted as $\mathbf{x}_O(t) = \{\mathbf{p}_i(t) \in \mathcal{I}_O\} \in \mathbb{R}^{n_O}$ and $\mathbf{x}_U(t) = \{\mathbf{p}_i(t) \in \mathcal{I}_U\} \in \mathbb{R}^{n_U}$, which together form the landmark state vector $\mathbf{x}_M(t) = \left[\mathbf{x}_O^T(t)\ \mathbf{x}_U^T(t)\right]^T \in \mathbb{R}^{n_M}$.

With the previous definitions, 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 $\mathbf{A}_{MV_i}(\mathbf{p}_i(t)) = \left[-\mathbf{I}_3\ -\mathbf{S}[\mathbf{p}_i(t)]\right]$.

The two state vectors here defined constitute the full state vector $\mathbf{x}_F(t) = \left[\mathbf{x}_V^T(t)\ \mathbf{x}_M^T(t)\right]^T$, with the full system dynamics reading as

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

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

$$\mathbf{A}_{MV}(\mathbf{x}_M(t)) = \left[\mathbf{A}_{MV_1}^T(\mathbf{p}_1(t)) \cdots \mathbf{A}_{MV_{N_M}}^T(\mathbf{p}_{N_M}(t))\right]^T$$

and $\mathbf{A}_M(t) = \mathrm{diag}\,(-\mathbf{S}[\boldsymbol{\omega}_m(t)], \ldots, -\mathbf{S}[\boldsymbol{\omega}_m(t)])$. From (3) it follows that the system may be expressed in a way similar to the usual linear system form. However, it can be seen that the system above is nonlinear, as the dynamics matrix depends on the landmarks that constitute the state. On the other hand, it must be noted that, as $\mathbf{y}(t) = \mathbf{x}_O(t)$, the dynamics matrix can be written as a function of the system output and the non-visible landmarks.

2.2 Problem statement

The problem addressed in this paper is the design of a SLAM filter in the space of the sensors by means of a filter for system (3), providing a sensor-based map and the velocity of the vehicle. Besides the map, which is defined by tridimensional landmarks, the filter shall have as by-product the estimation of the angular measurement bias. The pose of the vehicle is deterministic as its position corresponds to the origin of the body-fixed frame, which also defines its attitude, i.e, the vehicle is always at the origin of the sensor-based map.

3 Observability analysis

Observability is of the utmost importance in any filtering problem and it is the main focus of the theoretical work presented in this paper. This section aims at analysing the dynamical system previously defined, by establishing several observability results, with the ultimate purpose of designing a SLAM filter with stability and convergence guarantees. 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). This is done because the non-visible

landmarks are clearly unobservable, and it is shown here that they do not influence the observability of the visible landmarks and of the vehicle state.

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

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

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

$$\mathbf{C} = \left[\mathbf{0}_{n_O \times n_V}\ \mathbf{I}_{n_O}\right].$$

The submatrices with subscript $(.)_O$ are equal to their subscript-less counterparts but use only the visible landmarks and have the appropriate dimensions. Furthermore, as indicated in Sect. 2, the visible part of the landmark state coincides with the output, and can therefore be substituted, eliminating the dependency of the dynamics matrix on the system state. However, the matrix $\mathbf{A}(t, \mathbf{y}(t))$ depends not only on time but also on the system output, even though the system can be seen as a linear time-varying system for observability analysis purposes as the dependency on the system state is now absent and the system output is known. This property is established in (Batista et al. 2011b, Lemma 1), according to which if the observability Gramian associated with a system with a dynamics matrix depending on the system input and output is invertible, then the system is observable. This result will be used throughout this section. However, before proceeding with this analysis the following assumption is introduced.

Assumption 1 Any two detected position landmarks are assumed to be different and nonzero, i.e., $\mathbf{y}_i(t) \neq \mathbf{y}_j(t)$ and $\mathbf{y}_i(t) \neq \mathbf{0}$ for all $t \geq t_0$ and $i, j \in \mathcal{I}_O$, where $\mathbf{y}_i(t) = \mathbf{p}_i(t)$ is i-th vector component of the output, $i \in \{1, \ldots, N_O\}$.

It is important to notice that it is physically impossible to have two collinear landmarks, let alone equal, visible at the same time, as the angle of view of the camera is always smaller than 180°, as shown in Fig. 2b. Furthermore, a landmark cannot be coincident with the vehicle. All this makes it a very mild assumption.

The analysis starts by investigating the observability of the LTV system (4) and then by showing that the original nonlinear system (3) is equivalent to the LTV. In short, the results of this section explore the observability and uniform complete observability of the LTV system and extend the observability result to the original nonlinear system. It is found that the geometric properties of the set of landmark observations in time drive the observability of both systems, in the sense that the complete set of observations of all time must contain sensor-based landmarks that define a plane, either naturally or through the motion of the vehicle. The following theorem states the sufficient and necessary conditions for the observability of system (4).

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

(i) *there are, at least, three visible landmarks at the same time t_1 that define a plane;*
(ii) *there exist two visible 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; or*
(iii) *there is a visible time-varying 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 analysis by means of a Lyapunov transformation [see Brockett (1970)], and then proving that the observability Gramian of the transformed system is non-singular in the conditions of the theorem, which, as (Batista et al. 2011b, Lemma 1) states, implies the observability of the transformed system. 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 for example Batista et al. (2011b) and Batista et al. (2010).

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

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

where $\mathbf{T}(t) = \text{diag}\left(\mathbf{I}_{n_V}, \mathbf{R}_m(t), \ldots, \mathbf{R}_m(t)\right)$ and the rotation matrix $\mathbf{R}_m(t) \in \text{SO}(3)$ satisfies the differential equation $\dot{\mathbf{R}}_m(t) = \mathbf{R}_m(t)\mathbf{S}[\boldsymbol{\omega}_m(t)]$. The computation of the new system dynamics and output is simple, yielding

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

The dynamics matrix is given by

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

where

$$\mathcal{A}_{MV}(t, \mathbf{y}(t)) = \begin{bmatrix} -\mathbf{R}_m(t) & -\mathbf{R}_m(t)\mathbf{S}[\mathbf{y}_1(t)] \\ \vdots & \vdots \\ -\mathbf{R}_m(t) & -\mathbf{R}_m(t)\mathbf{S}[\mathbf{y}_{N_O}(t)] \end{bmatrix}.$$

The output matrix is simply

$$\mathcal{C}(t) = \left[\mathbf{0}_{n_O \times n_V} \ \text{diag}\left(\mathbf{R}_m^T(t), \cdots, \mathbf{R}_m^T(t)\right)\right].$$

Before proceeding to computing the observability Gramian associated with the transformed system (6), it is necessary to know its transition matrix. Computing $\mathbf{z}(t)$ as a function of $\mathbf{z}(t_0)$ by solving $\boldsymbol{\phi}(t, t_0)\mathbf{z}(t_0) = \mathbf{z}(t_0) + \int_{t_0}^{t} \mathcal{A}(\tau, y(\tau))\mathbf{z}(\tau)d\tau$, or using the Peano-Baker series, 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

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

It is noted that $\mathbf{y}_i(t)$ was replaced by $\mathbf{p}_i(t)$ as it is the same for all $i \in \mathcal{I}_O$. The observability Gramian is defined by

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

If $\mathcal{W}(t_0, t_f)$ is invertible, then 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 contraposition, that this is the case, i.e, by assuming that the system is unobservable, which implies that the observability Gramian on $\mathcal{T}$ is singular, and showing that the conditions of the theorem do not hold under that hypothesis. Suppose then that $\mathcal{W}(t_0, t_f)$ is not invertible, which means that there exists a unit vector

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_v^T & \mathbf{c}_b^T & \mathbf{c}_1^T & \cdots & \mathbf{c}_{N_O}^T \end{bmatrix}^T \in \mathbb{R}^{(n_V + n_O)},$$

such that,

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

Expanding (7) gives

$$\mathbf{c}^T \mathcal{W}(t_0, t_f)\mathbf{c} = \int_{t_0}^{t_f} \|\mathbf{f}(\tau, t_0)\|^2 \, d\tau, \tag{8}$$

where $\mathbf{f}(\tau, t_0) := \text{diag}\left(\mathbf{R}_m(\tau), \ldots, \mathbf{R}_m(\tau)\right)\mathcal{C}(\tau)\boldsymbol{\phi}(\tau, t_0)\mathbf{c}$. The evaluation of $\mathbf{f}(\tau, t_0) = \left[\boldsymbol{\phi}_{MV}(\tau, t_0)\ \mathbf{I}_{n_O}\right]\mathbf{c}$ and its derivative yields

$$\mathbf{f}(\tau, t_0) = \begin{bmatrix} \mathbf{c}_1 - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{c}_v d\sigma - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{S}[\mathbf{p}_1(\sigma)]\mathbf{c}_b d\sigma \\ \vdots \\ \mathbf{c}_{N_O} - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{c}_v d\sigma - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{S}[\mathbf{p}_{N_O}(\sigma)]\mathbf{c}_b d\sigma \end{bmatrix} \tag{9}$$

and

$$\frac{\partial \mathbf{f}(\tau, t_0)}{\partial \tau} = \begin{bmatrix} -\mathbf{R}_m(\tau)\mathbf{c}_v - \mathbf{R}_m(\tau)\mathbf{S}[\mathbf{p}_1(\tau)]\mathbf{c}_b \\ \vdots \\ -\mathbf{R}_m(\tau)\mathbf{c}_v - \mathbf{R}_m(\tau)\mathbf{S}[\mathbf{p}_{N_O}(\tau)]\mathbf{c}_b \end{bmatrix}, \tag{10}$$

respectively. In order for (7) to be true, both quantities must be zero for all $\tau \in \mathcal{T}$. For the first case, this restriction immediately leads to $\mathbf{c}_i = 0$ for all $i \in \mathcal{I}_O$, when $\mathbf{f}(\tau, t_0) = 0$ is evaluated at $\tau = t_0$. Equating the first time derivative in (10) to zero implies that

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}[\mathbf{p}_1(\tau)] \\ \vdots & \vdots \\ \mathbf{I}_3 & \mathbf{S}[\mathbf{p}_{N_O}(\tau)] \end{bmatrix} \begin{bmatrix} \mathbf{c}_v \\ \mathbf{c}_b \end{bmatrix} = \mathbf{0}, \quad \forall \tau \in \mathcal{T}. \tag{11}$$

Thus, to prove the sufficiency of the conditions of the theorem it remains to show that those conditions do not hold. For that purpose, consider the situation where there are three visible landmarks $\mathbf{p}_i(t_1)$, $i \in \{1, 2, 3\}$. In this case (11) can be rewritten as

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}[\mathbf{p}_1(t_1)] \\ \mathbf{0}_3 & \mathbf{S}[\mathbf{p}_2(t_1) - \mathbf{p}_1(t_1)] \\ \mathbf{0}_3 & \mathbf{S}[\mathbf{p}_3(t_1) - \mathbf{p}_1(t_1)] \end{bmatrix} \begin{bmatrix} \mathbf{c}_v \\ \mathbf{c}_b \end{bmatrix} = \mathbf{0}. \tag{12}$$

From this, it is simple to find either that $\mathbf{c}_b = 0$ or that all three landmarks form a line. The first case implies that $\mathbf{c}_v$ is also zero, contradicting the hypothesis that $\mathbf{c}$ is a unit vector. Thus, all the landmarks must form a line and (i) cannot hold.

In the case where any of the remaining conditions applies, an equation similar to (12) 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 the observability Gramian is singular, none of the

hypothesis of the theorem can hold, which means that, if at least one of the conditions of Theorem 1 holds, then $\mathcal{W}(t_0, t_f)$ is invertible on $\mathcal{T}$, and, using Batista et al. (2011b, Lemma 1, Sect. 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 sufficiency.

Proving the necessity of the conditions of the theorem for the observability follows naturally by contraposition, i.e., by assuming that neither (i), (ii), or (iii) hold and showing that system (6) is not observable. In this situation, there is no possible combination of landmarks in the time interval $\mathcal{T}$ that defines a plane, which is equivalent to say that either there are no visible landmarks or all the visible ones are static or move in the same line. Hence, all the N_O observed landmarks must respect

$$\mathbf{p}_i(t) = p_{d_i}(t)\mathbf{d} + \mathbf{p}_i(t_0), \tag{13}$$

where $\mathbf{d} \in \mathbb{R}^3$ defines the line, $p_{d_i}(t) \in \mathbb{R}$, and $i \in \mathcal{I}_O$. The static case is trivially provided by a constant $p_{d_i}(t)$ for all t. The proof starts by computing the explicit time evolution of the output of the transformed system (6), which is given by $\mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\boldsymbol{\phi}(t, t_0)\mathbf{z}(t_0)$, whose i-th element is

$$\begin{aligned}\mathbf{y}_i(t) = {} & \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_i(t_0) \\ & + \mathbf{R}_m^T(t)\int_{t_0}^{t} \mathbf{R}_m(\sigma)\left(\mathbf{v}(t_0) - \mathbf{S}[\mathbf{b}(t_0)]\mathbf{p}_i(\sigma)\right)d\sigma.\end{aligned} \tag{14}$$

Substituting the negation of the theorem conditions expressed by (13) in this expression yields

$$\begin{aligned}\mathbf{R}_m(t)\mathbf{y}_i(t) = {} & \mathbf{R}_m(t_0)\mathbf{p}_1(t_0) \\ & - \int_{t_0}^{t} \mathbf{R}_m(\sigma)\big[\mathbf{v}(t_0) - \mathbf{S}[\mathbf{b}(t_0)](p_{d_i}(\sigma)\mathbf{d} \\ & + \mathbf{p}_1(t_0))\big]d\sigma,\end{aligned}$$

for all $i \in \mathcal{I}_O$. If there are two different initial states that lead to the same output, then the system is not observable. With that purpose in mind, let

$$\bar{\mathbf{z}}_0 := \left[\bar{\mathbf{v}}_0^T \ \bar{\mathbf{b}}_0^T \ \mathbf{p}_1^T(t_0) \cdots \mathbf{p}_{N_O}^T(t_0)\right]^T, \tag{15}$$

and

$$\tilde{\mathbf{z}}_0 := \left[\tilde{\mathbf{v}}_0^T \ \tilde{\mathbf{b}}_0^T \ \mathbf{p}_1^T(t_0) \cdots \mathbf{p}_{N_O}^T(t_0)\right]^T \tag{16}$$

be two different initial conditions for system (6), such that $\bar{\mathbf{b}}_0 = a\mathbf{d}$, $\bar{\mathbf{v}}_0 = \mathbf{S}[\bar{\mathbf{b}}_0]\mathbf{p}_1(t_0)$, $\tilde{\mathbf{b}}_0 = b\mathbf{d}$, and $\tilde{\mathbf{v}}_0 = \mathbf{S}[\tilde{\mathbf{b}}_0]\mathbf{p}_1(t_0)$, with $a \neq b$ and $a, b \in \mathbb{R}\backslash\{0\}$. It is possible to see that given either initial states (15) and (16) the output of the system is the same and given by

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_1(t_0) \\ \vdots \\ \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_{N_O}(t_0) \end{bmatrix}.$$

Given that the two initial states are different and lead to the same output for all $t \in \mathcal{T}$, the system is rendered unobservable if the conditions of the theorem are not satisfied. Hence, if the system (6) is observable, at least one of the conditions must hold, which, combined with the fact that the Lyapunov transformation preserves the observability properties, thus concludes the proof of necessity. □

Given the sufficient and necessary conditions for observability, this section aims at designing a filter for the nonlinear system (3), with globally asymptotically stable error dynamics. However, further results are necessary to achieve this goal. The following theorem 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 Consider the transformed system (6), whose state, and therefore initial condition, is related with the state of the nonlinear system (4), regarded as LTV. The proof follows with the transformed system for simplicity of analysis.

Let the initial condition for this system be given by

$$\bar{\mathbf{z}}(t_0) = \left[\bar{\mathbf{v}}^T(t_0) \ \bar{\mathbf{b}}_\omega^T(t_0) \ \bar{\mathbf{z}}_1^T(t_0) \cdots \bar{\mathbf{z}}_{N_O}^T(t_0)\right]^T,$$

where $\bar{\mathbf{z}}_i(t) = \mathbf{R}_m(t)\mathbf{y}_i(t)$ for all $i \in \mathcal{I}_O$, resulting from $\mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\mathbf{z}(t)$. It is noted that $\mathbf{z}(t) = \boldsymbol{\phi}(t, t_0)\bar{\mathbf{z}}(t_0)$, which then yields $\mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\boldsymbol{\phi}(t, t_0)\bar{\mathbf{z}}(t_0)$. Recall (14) and that the i-th component of the output of the system is given by

$$\begin{aligned}\mathbf{y}_i(t) = {} & \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{y}_i(t_0) \\ & - \mathbf{R}_m^T(t)\int_{t_0}^{t} \mathbf{R}_m(\sigma)\left(\bar{\mathbf{v}}(t_0) + \mathbf{S}[\mathbf{y}_i(\sigma)]\bar{\mathbf{b}}_\omega(t_0)\right)d\sigma,\end{aligned}$$

where the relation between $\mathbf{y}_i(t)$ and $\bar{\mathbf{z}}_i(t_0)$ was used. The next steps include left multiplying this expression by $\mathbf{R}_m(t)$, differentiating, and further simplifying by left multiplying both sides by $\mathbf{R}_m(t)^T$, which yields the first derivative of $\mathbf{y}_i(t)$,

$$\dot{\mathbf{y}}_i(t) = -\bar{\mathbf{v}}(t_0) - \mathbf{S}[\mathbf{y}_i(t)]\left(\bar{\mathbf{b}}_\omega(t_0) - \boldsymbol{\omega}_m(t)\right). \tag{17}$$

Now consider the nonlinear system (3). Let the initial condition of this system be given by

$$\mathbf{x}(t_0) = \left[\mathbf{v}^T(t_0)\ \mathbf{b}_\omega^T(t_0)\ \mathbf{p}_1^T(t_0)\ \cdots\ \mathbf{p}_N^T(t_0)\right]^T,$$

where $\mathbf{p}_i(t_0) = \mathbf{y}_i(t_0)$ for all $\mathbf{p}_i(t_0) \in \mathcal{I}_O$. The output of the system is related to the state by

$$\mathbf{y}(t) = \mathbf{C}\int_{t_0}^{t} \mathbf{A}(\mathbf{x}(\sigma), \sigma)\mathbf{x}(\sigma)d\sigma + \mathbf{C}\mathbf{x}(t_0),$$

which yields

$$\mathbf{y}_i(t) = \mathbf{y}_i(t_0) - \int_{t_0}^{t} \left(\mathbf{v}(t_0) + \mathbf{S}[\mathbf{p}_i(\sigma)]\left(\mathbf{b}_\omega(t_0) - \boldsymbol{\omega}_m(\sigma)\right)\right)d\sigma, \tag{18}$$

with first time derivative given by

$$\dot{\mathbf{y}}_i(t) = -\mathbf{v}(t_0) - \mathbf{S}[\mathbf{p}_i(t)]\left(\mathbf{b}_\omega(t_0) - \boldsymbol{\omega}_m(t)\right). \tag{19}$$

Comparison of (17) with (19) after substituting $\mathbf{p}_i(t)$ by $\mathbf{y}_i(t)$ yields

$$\mathbf{0} = \left(\bar{\mathbf{v}}(t_0) - \mathbf{v}(t_0)\right) - \mathbf{S}[\mathbf{y}_i(t)]\left(\bar{\mathbf{b}}_\omega(t_0) - \mathbf{b}_\omega(t_0)\right)$$

for all t in $\mathcal{T}$ and $i \in \mathcal{I}_O$. When the conditions of Theorem 1 hold, this system yields $\bar{\mathbf{v}}(t_0) = \mathbf{v}(t_0)$ and $\bar{\mathbf{b}}_\omega(t_0) = \mathbf{b}_\omega(t_0)$ by a similar reasoning to that which was used to prove the sufficiency of those conditions to ensure the observability of the system. As mentioned before, the initial condition $\bar{\mathbf{z}}(t_0)$ is related to that of the nonlinear system (4) by the Lyapunov transformation $\mathbf{T}(t)$. Hence, under the conditions of Theorem 1, the initial state of the nonlinear system (4), regarded as LTV, and the initial state of the nonlinear system (3), discarding the non-visible landmarks, are the same and uniquely defined.

The first part of the theorem, now proven, gives insight for the proof of the second part. An observer designed for a LTV system with globally asymptotically stable error dynamics has an estimation error convergent to zero, implying that the estimates asymptotically tend to the true state. Therefore, as the true state of the nonlinear system and the state of the LTV system are shown to be the same under the conditions of the theorem, the estimation error of the same observer applied to the nonlinear system converges to zero asymptotically fast. □

Theorem 2 established sufficient conditions for the observability of the nominal nonlinear system by relating it to the LTV system (4). This can be complemented with the following theorem to ensure a stronger result, by showing that those conditions are not only sufficient but in fact necessary.

Theorem 3 *The nonlinear system (4), discarding the non-visible landmarks, is observable if and only if the conditions of Theorem 1 hold.*

Proof Theorems 1 and 2 establish the ground for this proof, as the sufficiency part comes directly from those results. For the necessity part, consider the output of the nonlinear system expressed in terms of time, as presented in (18). The proof is made by substituting the negation of the conditions of the theorem in the output equation and showing that two different initial conditions lead to the same output, following very similar terms to that of the necessity part of Theorem 1 and is therefore omitted. □

Given that a GAS observer for system (4) is an observer for the nominal nonlinear system, the design of such an 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 shown in Anderson (1971).

The concept of uniform complete observability implies uniform bounds on the observability Gramian in time intervals of length δ. It is a stronger form of observability for linear time-varying systems needed in the steps to design a GAS observer. The following theorem states the conditions for this property to be verified, and requires the introduction of a new assumption.

Assumption 2 Any detected landmark is upper and lower bounded in norm, i.e.

$$\underset{t\geq t_0}{\forall}\ \underset{i\in\mathcal{I}_O}{\forall}\ \underset{P_m, P_M>0}{\exists}: \quad P_m \leq \|\mathbf{y}_i(t)\| \leq P_M.$$

This imposes upper and lower bounds on the norm of each landmark, which may be seen as a demanding assumption. However, given that the vehicle cannot be arbitrarily distant to any landmark and that, furthermore, the range of the RGB-D camera has also a lower and upper limit, these bounds are perfectly justified. With the introduction of this assumption, the analysis can now proceed with the following result.

Theorem 4 *Consider system (4) regarded as LTV. The pair $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$ is uniformly completely observable if and only if there exist $\delta > 0$ and $\alpha_* > 0$ 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* $\mathbf{p}_1(t)$, $\mathbf{p}_2(t)$ *and* $\mathbf{p}_3(t)$ *such that* $\| (\mathbf{p}_1(t_1) - \mathbf{p}_2(t_1)) \times (\mathbf{p}_1(t_1) - \mathbf{p}_3(t_1)) \| \geq \alpha_*$;

(ii) *there exist two visible landmarks at times* t_1, t_2 *such that* $\| (\mathbf{p}_1(t_1) - \mathbf{p}_2(t_1)) \times \left(\mathbf{p}_j(t_1) - \mathbf{p}_j(t_2)\right) \| \geq \alpha_*$ *for some* $j \in \{1, 2\}$; *or*

(iii) *there exists one visible time-varying landmark such that its coordinates, at three different instants of time* $\{t_1, t_2, t_3\}$, *satisfy* $\| (\mathbf{p}_1(t_1) - \mathbf{p}_1(t_2)) \times (\mathbf{p}_1(t_1) - \mathbf{p}_1(t_3)) \| \geq \alpha_*$.

Proof The proof, provided in Appendix 1, follows similar steps to the proofs of Theorem 1, but considering uniform bounds for all $t \geq t_0$ and intervals $[t, t + \delta]$. □

Remark 1 The conditions of Theorem 4 together with Assumption 2 imply that the vectors defined by the set of observed landmarks must be sufficiently away from collinearity so that the plane defined by them does not degenerate in time. Moreover, conditions (ii) and (iii) can be seen as persistent excitation conditions.

The results shown in this section lead towards the design of a state observer, such as the Kalman filter for linear time-varying systems, with globally asymptotically stable error dynamics for the nominal nonlinear system (3). This property can be shown using the Lyapunov function $V(t, \tilde{\mathbf{x}}(t)) = \tilde{\mathbf{x}}^T(t)\mathbf{P}^{-1}(t)\tilde{\mathbf{x}}(t)$, where $\tilde{\mathbf{x}}(t)$ denotes the estimation error and $\mathbf{P}(t)$ is the error covariance matrix. Under the conditions of Theorem 4, it is possible to use Anderson (1971, Lemma 3.2) to show that $\mathbf{P}^{-1}(t)$ is positive definite . Using this fact and following the steps in Khalil (2002, Example 8.11) for this Lyapunov function, it can then be shown that the conditions of Khalil (2002, Theorem 8.5) are satisfied globally, which means that the error dynamics of the Kalman filter for the LTV system (4) are indeed globally asymptotically stable. The discrete-time implementation of such a filter is discussed in the following section.

4 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 depth camera (or other tridimensional relative position sensor). Hence, it is important to obtain the discrete-time version of the dynamic system under analysis.

4.1 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 state $\mathbf{x}_k := \mathbf{x}(t_k)$, the dynamics matrix $\mathbf{A}_k(\mathbf{y}_k) := \mathbf{A}(t_k, y(t_k))$ and the output matrix $\mathbf{C}_k := \mathbf{C}(t_k)$. Finally, the Euler discretization of the system dynamics (4), including system disturbance and measurement noise, yields

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \boldsymbol{\xi}_k \\ \mathbf{y}_{k+1} = \mathbf{H}_{k+1}\mathbf{x}_{k+1} + \boldsymbol{\theta}_{k+1} \end{cases} \tag{20}$$

where $\mathbf{F}_k := \mathbf{I}_{n_x} + T_s\mathbf{A}_k(\mathbf{y}_k)$ and $\mathbf{H}_{k+1} := \mathbf{C}_{k+1}$. The disturbance vector $\boldsymbol{\xi}_k$ and the measurement noise vector $\boldsymbol{\theta}_k$ are both zero-mean discrete white Gaussian noise, with $\langle \boldsymbol{\xi}_k \boldsymbol{\xi}_l^T \rangle = \boldsymbol{\Xi}_k \delta_{k-l}$ and $\langle \boldsymbol{\theta}_k \boldsymbol{\theta}_l^T \rangle = \boldsymbol{\Theta}_k \delta_{k-l}$, where $\langle . \rangle$ denotes the expected value of its argument and δ_k is the Dirac delta function at instant k.

4.2 Prediction step

The discrete system (20) does not include the non-visible landmarks, which must be propagated in open-loop using the nonlinear equations in (3). Thus, with the full state vector, $\mathbf{x}_{F_k} := \begin{bmatrix} \mathbf{x}_k^T & \mathbf{x}_{U_k}^T \end{bmatrix}^T$, the prediction step of the filter for the visible landmarks is that of the Kalman filter, see Gelb (1974), and for non-visible ones the non-visible state must be used in the equations, given by

$$\begin{cases} \hat{\mathbf{x}}_{F_{k+1|k}} = \mathbf{F}_{F_{k|k}} \hat{\mathbf{x}}_{F_{k|k}} \\ \boldsymbol{\Sigma}_{F_{k+1|k}} = \hat{\mathbf{F}}_{F_{k|k}} \boldsymbol{\Sigma}_{F_{k+1|k}} \hat{\mathbf{F}}_{F_{k|k}}^T + \boldsymbol{\Xi}_{F_k} \end{cases}$$

where the full discrete dynamics matrix is given by $\mathbf{F}_{F_{k|k}} = \mathbf{I}_{n_x} + T_s\mathbf{A}_{F_{k|k}}(\mathbf{y}_k, \hat{\mathbf{x}}_{U_{k|k}})$ and is in fact an estimated version, as the block that relates to the non-visible landmarks, namely $\mathbf{A}_{MV}(\mathbf{x}_M(t))$, is created with the estimated non-visible landmarks, $\hat{\mathbf{x}}_{U_{k|k}}$. The covariance prediction is made with the Jacobian of the system dynamics,

$$\begin{aligned} \hat{\mathbf{F}}_{F_{k|k}} &= \frac{\partial \mathbf{F}_{F_{k|k}} \hat{\mathbf{x}}_{F_{k|k}}}{\partial \hat{\mathbf{x}}_{F_{k|k}}} \\ &= \begin{bmatrix} \mathbf{F}_k & \mathbf{0}_{n_x \times n_U} \\ \begin{bmatrix} \mathbf{A}_{MV_U}(\hat{\mathbf{x}}_{U_{k|k}}) & \mathbf{0}_{n_U \times n_O} \end{bmatrix} & -\operatorname{diag}(\mathbf{S}[\boldsymbol{\omega}_{m_k} - \hat{\mathbf{b}}_{\omega_k}]) \end{bmatrix}, \end{aligned}$$

and with $\boldsymbol{\Xi}_{F_k} = \operatorname{diag}(\boldsymbol{\Xi}_k, \boldsymbol{\Xi}_{U_k})$. This prediction step uses the measurements of the rate gyros, propagating the state every time a reading is available. Note that, even though the visible landmarks and the vehicle state have LTV-like dynamics, that is only achieved by using directly the landmark and angular velocity measurements in the dynamics. For this reason, this approach is not guaranteed to be optimal, and its noise characterization is not exact. However, the noise parameters $\boldsymbol{\Xi}$ can be calibrated a priori with Monte Carlo analysis and with actual measurements, to better cope with this matter.

4.3 Update step

The update step is divided in two different stages, landmark association and the update equations. This step occurs every time a RGB-D (colour and depth) image is available from the *Kinect*. The RBG two-dimensional frame is processed by an implementation of SURF (Bay et al. 2008), which detects and provides a 64-dimensional descriptor to points-of-interest on the 2-D pictures of the environment. The resulting features are then matched to a pointcloud built with the depth image. This matching returns a set of observed tridimensional landmarks in cartesian coordinates. One of the main issues in performing SLAM with natural untagged landmarks, or in mapping in general, is the problem of data association: as the landmarks are collected from the environment, there is no way to know a priori with certainty if a landmark from the set of measurements is in the current map or if it is being detected for the first time. At this point, some form of landmark association must take place to match the measured data with the existing landmarks in the filter state.

Landmark association Wrong associations may have a very negative effect on the estimation, so this is a field where a lot of research effort has been put, yielding a multitude of algorithms from the community. There are two ways to ascertain the correspondence between measured and observed landmarks, either by comparison of their positions, usually through computation of the distances between landmarks in different sets, or by comparison of distinctive characteristics. The use of image processing procedures permits the description of the landmarks by labels with a great deal of information, and which are invariant and robust to several transformations. This fact enhances greatly the data association, as in landmark-filled environments spurious associations when using only the coordinates are very common. For this reason, the filter proposed in this paper employs a greedy search algorithm with mutual exclusion which looks for the landmarks with closest descriptor. Landmark detection provides tens of landmarks that cannot possibly be all incorporated in the filter, and as such there must be a trimming down process. First the scene is divided in quadrants, and the landmarks in each quadrant are organized by the value of the determinant of their Hessian [see Bay et al. (2008) for more information on these quantities]. For the desired number N_O of landmarks being chosen, it is only a matter of picking from each quadrant in equal parts so that an even distribution of features is achieved. These landmarks constitute the measured set and are then associated with the state landmarks whose predicted coordinates are within the field of view of the camera (see Fig. 2b). Besides the 64-D descriptor that differentiates landmarks and that provides the main information upon which the association decision is based, each SURF feature has a boolean Laplacian, which is used to prune the search tree, by dividing it in two. The association process can then proceed regularly for each search tree (depending on the Laplacian). The candidate pairs of landmarks to associate are chosen according to the lowest absolute difference of descriptors, and the decision is supported upon the Mahalanobis distance of the landmarks themselves, which follows a χ^2 distribution with three degrees of freedom. A statistical test based on the 95 % percentile is performed to decide whether an association is valid of not. Finally, when a landmark is associated, the measured descriptor is combined with the descriptor of the state landmark with a mean filter. It must be remarked that both the landmark detection and association algorithms may be substituted by others, as they are independent from the filtering technique described in this paper.

The association algorithm redefines the new sets of visible and non-visible landmarks, and also provides the innovation vector $\boldsymbol{\nu}_{k+1}$ and its covariance matrix $\boldsymbol{\Sigma}_{\nu_{k+1}}$ given by

$$\begin{cases} \boldsymbol{\nu}_{k+1} = \mathbf{y}_{k+1} - \mathbf{H}_{k+1}\hat{\mathbf{x}}_{k+1|k} \\ \boldsymbol{\Sigma}_{\nu_{k+1}} = \mathbf{H}_{k+1}\boldsymbol{\Sigma}_{k+1|k}\mathbf{H}_{k+1}^T + \boldsymbol{\Theta}_{k+1} \end{cases}.$$

The update equations are standard, and given by

$$\begin{cases} \hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\boldsymbol{\nu}_{k+1} \\ \boldsymbol{\Sigma}_{k+1|k+1} = \boldsymbol{\Sigma}_{k+1|k} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\boldsymbol{\Sigma}_{k+1|k}, \\ \mathbf{K}_{k+1} = \boldsymbol{\Sigma}_{k+1|k}\mathbf{H}_{k+1}^T\boldsymbol{\Sigma}_{\nu_{k+1}}^{-1} \end{cases}$$

where $\mathbf{K}_{k+1}$ is the Kalman gain.

4.4 State maintenance

The number of landmarks and therefore the size of the state, the size of the state covariance, and the size of the search trees for data association can grow unbounded with the mission time. Due to the computational complexity of SLAM algorithms, this limits the operation time and the size of the environment. For this reason, it is usual to introduce state maintenance procedures to clean the state. In this algorithm, each time a measured landmark is not associated with any already existing landmark, it is automatically added to the state which is the mechanism of incorporation of new landmarks to the state. However, this can also lead to the inclusion of spurious measurements that do not repeat themselves in time an are therefore superfluous to the filter. To tackle this issue, and prevent the state from growing indefinitely in size, there is a counter for each time a landmark is observed, as well as a timer. This way it is possible to keep track of how long has passed since the last observation of a specific landmark was observed and how many times it appeared. A landmark is removed from the state if two conditions are met simultaneously: it must not have been observed for more than some

time $T_{\max}$ and it must have been observed less than $N_{\min}$ times.

5 Earth-fixed trajectory and map

Simultaneous localization and mapping is intrinsically a nonlinear problem and, as such, assumes many interpretations, approaches, and solutions. However, a great number of algorithms choose to describe the problem in an Earth-fixed frame. To be statistically consistent, this formulation must include the pose of the vehicle either in its state or in the probability distribution that describes the belief of the algorithm for particle filter implementations. Nevertheless, in these formulations the onboard sensors provide vehicle-centered measurements that need to be transformed to the Earth-fixed frame. Working directly on a sensor-based framework enables to leave the Earth-fixed pose representation out of the filtering process and avoids the transformation of the measurements. However, there may still be need for an estimate of the Earth-fixed pose of the vehicle, as well as the map of the environment. The algorithm here proposed encompasses the SLAM filter working on the sensor-based frame, on one hand, and an Earth-fixed Trajectory and Map (ETM) estimation methodology for the Earth-fixed map and vehicle pose on the other. The full algorithm is outlined in Fig. 1, and this section summarizes the last part, a methodology proposed in Guerreiro et al. (2012) and extended in Lourenço et al. (2013), where an algorithm is presented to solve the problem of obtaining an estimate of the pose of the vehicle and of the Earth-fixed map using only the sensor-based map provided by the SLAM filter. For that purpose, an optimization problem is proposed, with a closed-form solution that provides an estimate of the transformation between the body-fixed frame $\{B\}$ and the Earth-fixed frame $\{E\}$, defined by the pair $\left(\hat{\mathbf{R}}_k, {}^{E}\hat{\mathbf{p}}_k\right)$. Moreover an approximate characterization of the uncertainty of the outputs of the algorithm is also performed. The main idea behind this methodology is to solve a Procrustes problem in each iteration (Schönemann 1966), yielding the vehicle pose and allowing the computation of the Earth-fixed map obtained using

$$ {}^{E}\hat{\mathbf{p}}_{i_k} = {}^{E}\hat{\mathbf{p}}_k + \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} \tag{21} $$

for all $i \in \{1, \ldots, N\}$. Then, it is possible to define the error function

$$ {}^{E}\mathbf{e}_{i_k} = {}^{E}\hat{\mathbf{p}}_{i_{k-1}} - \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} - {}^{E}\hat{\mathbf{p}}_k, $$

that represents the error between the Earth-fixed landmark estimate i and its sensor-based homologous rotated and translated with the estimated transformation. It is noticeable that there is an algebraic loop in (21) that can only be averted if it is remarked that, as the landmarks in $\{E\}$ are static, this is also valid as a correspondence between ${}^{E}\hat{\mathbf{p}}_{i_{k-1}}$ and $\hat{\mathbf{p}}_{i_k}$. Minimizing this error is the purpose of the optimization problem

$$ \left(\hat{\mathbf{R}}_k^*, {}^{E}\hat{\mathbf{p}}_k{}^*\right) = \arg \min_{\substack{\hat{\mathbf{R}}_k \in SO(3) \\ {}^{E}\hat{\mathbf{p}}_k \in \mathbb{R}^3}} \sum_{i=1}^{N} \sigma_i^{-2} \left\| {}^{E}\mathbf{e}_{i_k} \right\|^2, \tag{22} $$

where σ_i serves as weight for each landmark pair, and encodes its uncertainty.

The closed-form solution for this family of problems was derived in Umeyama (1991) and subsequent works. It is done by computing the rotation between the two pointclouds after making their centroids coincide, and then computing the optimal translation between the two centroids. The optimal rotation matrix is given by

$$ \hat{\mathbf{R}}_k^* = \mathbf{U}\,\mathrm{diag}(1, 1, |\mathbf{U}|\,|\mathbf{V}|)\mathbf{V}^T, \tag{23} $$

where $\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathrm{svd}(\mathbf{B}_k^T)$, $\mathbf{B}_k^T := \mathbf{Y}_k \mathbf{W}_k \mathbf{X}_k^T$, $\mathbf{W}_k$ is a weight matrix, and $\mathbf{Y}_k = \left[{}^{E}\hat{\mathbf{p}}_{1_k} \cdots {}^{E}\hat{\mathbf{p}}_{N_k}\right]$ and $\mathbf{X}_k = \left[\hat{\mathbf{p}}_{1_k} \cdots \hat{\mathbf{p}}_{N_k}\right]$, $\mathbf{Y}_k, \mathbf{X}_k \in \mathbb{R}^{3\times N}$, are the concatenation of the landmark vectors expressed in the Earth-fixed and body-fixed frames, respectively. The optimal translation vector follows directly:

$$ {}^{E}\hat{\mathbf{p}}_{B_k}^* = \frac{1}{N_{W_k}} \left(\mathbf{Y}_k - \hat{\mathbf{R}}_k^* \mathbf{X}_k\right) \boldsymbol{\Sigma}_{{}^{E}\mathbf{e}_k}^{-1} \mathbf{1}, \tag{24} $$

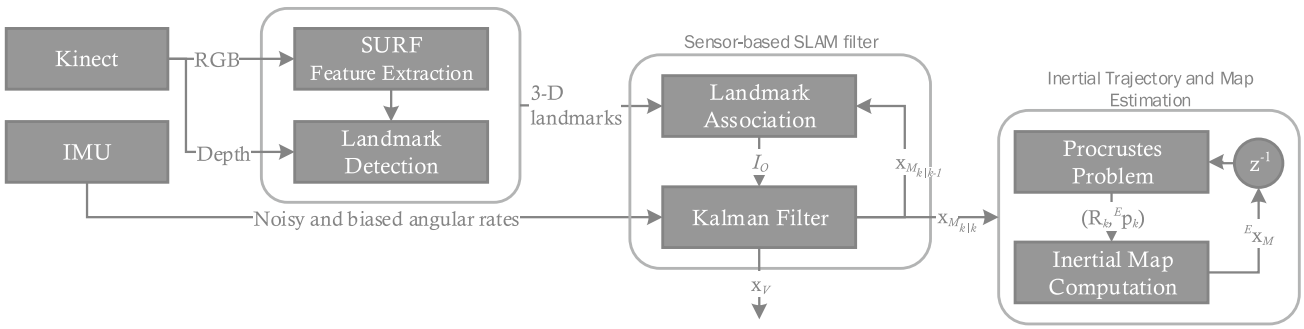


Fig. 1 The SLAM algorithm used in the experiments of Sect. 6

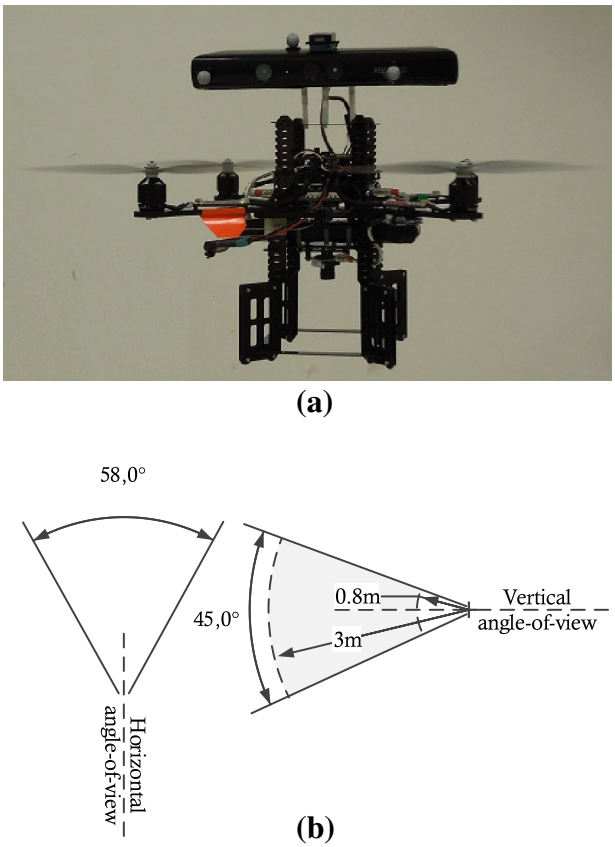


Fig. 2 The experimental setup. **a** The AscTec Pelican equipped with a *Microsoft Kinect*, a *Microstrain 3DM-GX3-25* and the *VICON* markers, **b** The field of view of the *Kinect*

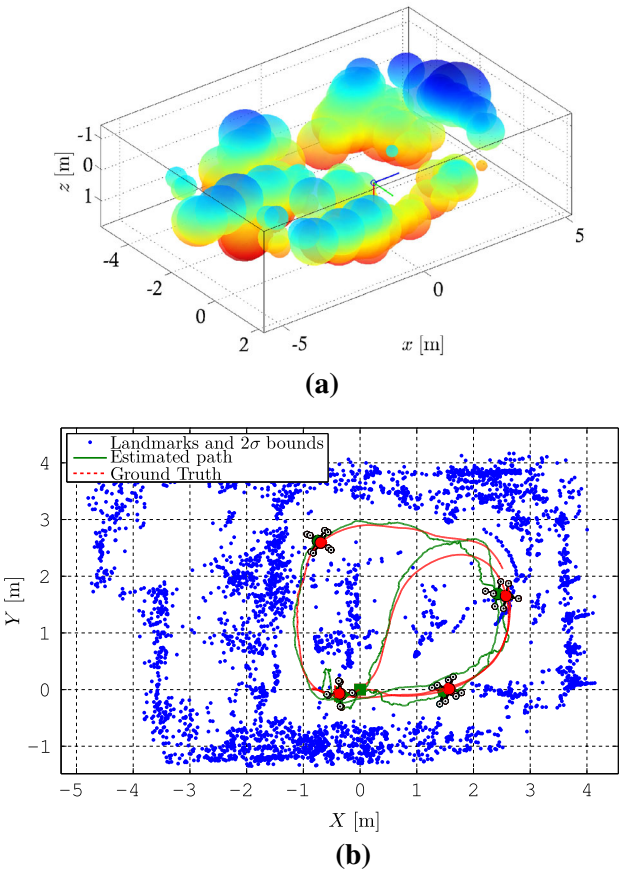


Fig. 3 Run #1—Landmark map of the environment with the estimated and real trajectory. **a** Sensor-based map with uncertainty ellipsoids, **b** top view of the Earth-fixed map with vehicle trajectory and ground truth

where $\mathbf{\Sigma}_{E_{\mathbf{e}_k}} = \text{diag}\left(\sigma_1^2, \ldots, \sigma_N^2\right)$, $\mathbf{1}$ is a column vector of ones, and $N_{W_k} = \mathbf{1}^T \mathbf{\Sigma}_{E_{\mathbf{e}_k}}^{-1} \mathbf{1}$. This solution to (22) is valid as long as $\mathbf{B}_k$ has rank 2, as reported in Umeyama (1991). This means that the set of landmarks must span $\mathbb{R}^2$, i.e., it must contain at least three landmarks whose connecting vectors define a plane. Therefore, it follows that as long as the initial pose is known and these conditions are kept throughout time, it is possible to recover the pose of the vehicle at all times.

In Lourenço et al. (2013) the authors proposed an approximate uncertainty description, founded on perturbation theory, of the estimates provided by the ETM algorithm, namely (21), (23), and (24). This uncertainty is characterized in the form of covariance matrices that depend solely on the sensor-based map and its covariance provided by the GAS SLAM filter proposed in this paper.

6 Experimental results

The theoretical results of the previous sections were consolidated by a series of experiments in the Sensor-Based Cooperative Robotics Research Laboratory - SCORE Lab of the University of Macau. The experimental setup consists of an *Asctec Pelican* quadrotor (see Fig. 2a), into which was added a *Microstrain 3DM-GX3-25* inertial measurement unit working at 50 Hz and a *Microsoft Kinect* camera, at 10 Hz. The experiments consisted in moving the quadrotor inside a 8 m × 6 m room (usable area of 16 m^2) 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. These estimates are used in this paper as ground truth for validation of the proposed SLAM algorithm.

This section provides experimental results for the performance and consistency evaluation of the sensor-based SLAM filter in connection with the ETM algorithm. The results in the Earth-fixed frame are obviously dependent on the individual performance of each algorithm. Conversely, the sensor-based filter does not depend on the ETM algorithm. Furthermore, the results of the ETM algorithm are naturally affected by the nonlinearity intrinsic to the problem of translating and rotating a map arbitrarily between coordinate frames, also found in EKF-based SLAM algorithms. By proposing this algorithm, it is argued that it may be possible to obtain a less

Table 1 SLAM filter parameters

Θ	Ξ_p	Ξ_v	Ξ_{b_ω}	Desired N_O
$(0.05\ m)^2\mathbf{I}_{n_O}$	$(0.05\ m)^2\mathbf{I}_{n_M}$	$(0.05\ m/s)^2\mathbf{I}_3$	$(5 \times 10^{-6}\ ^\circ/s)^2\mathbf{I}_3$	15

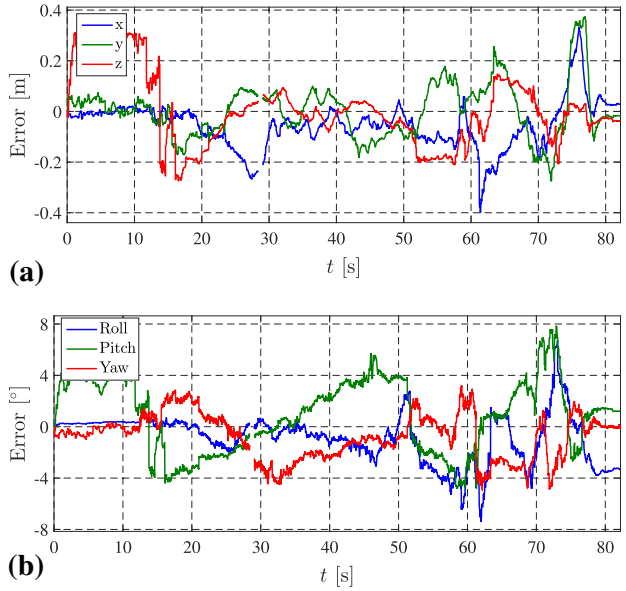


Fig. 4 Run #1—Time evolution of the estimation error of the Earth-fixed estimates. **a** Position, **b** attitude

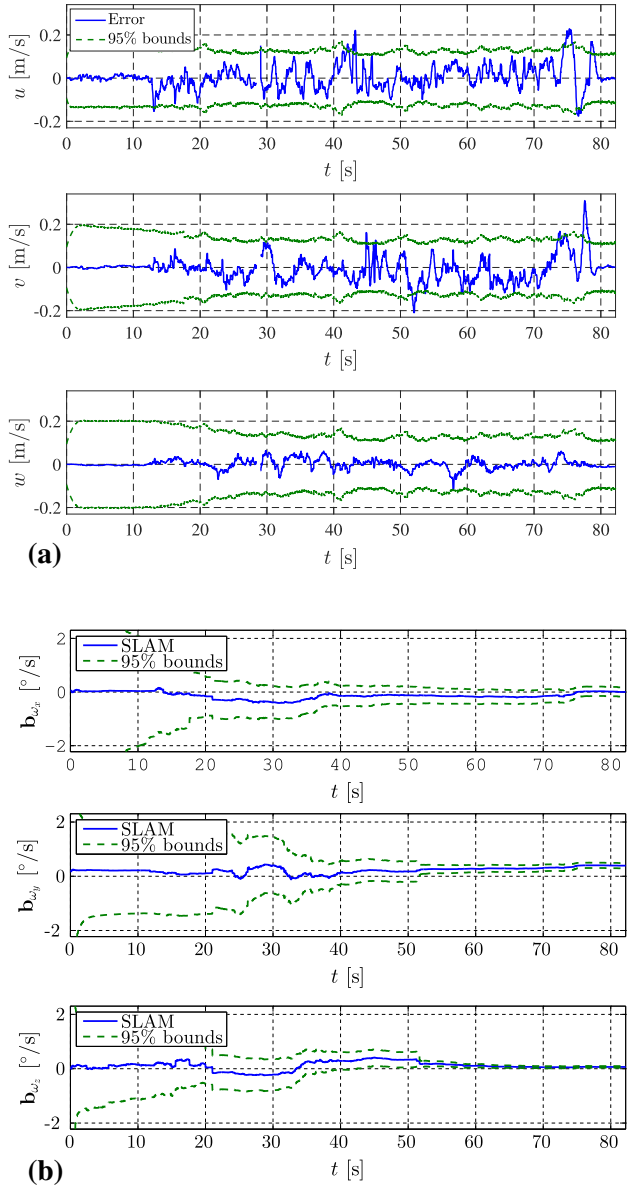


Fig. 5 Run #1—Time evolution of the sensor-based estimates with 2σ bounds. **a** Velocity estimation error, **b** angular measurement bias

uncertain Earth-fixed trajectory and landmark map. With this approach, all the landmark association, loop-closing, control, and decision procedures can be made in the sensor-based frame, minimizing the effects of nonlinearities in the consistency of the filter, whereas the Earth-fixed trajectory and map estimation is used only for completeness of results. As there is no ground truth for the sensor-based map (which is where the SLAM problem is posed and solved), the results of the ETM are used for comparison with the ground truth.

In this section, the results of two different experiments are presented: (i) a short path inside the lab, as shown in Fig. 3, with ground truth available at all times, and (ii) a larger path starting inside the lab but moving outside where no ground truth is available. The parameters of the algorithm used in these experiments are presented in Table 1, that includes the observation noise covariance, the process noise covariance, and the desired N_O, which is an indicative parameter that is used when choosing which and how many features in the RGB image to input as measured landmarks in the filter.

6.1 Run #1

The first run is simply a circular-like lap inside the working area of the lab. The vehicle is hand-driven at constant height and average speed of 0.3 m/s in a 25 m path, which can be seen in Fig. 3. In Fig. 3a the map in the sensor frame is presented with the respective 95 % (or 2σ) uncertainty ellipsoids. One important aspect to retain from this figure is that the uncertainty of the non-visible landmarks increases with time, noting that the most recent landmarks, i.e., those close to the vehicle, have much smaller uncertainty ellipsoids.

In Fig. 3b, a top view of the Earth-fixed map is shown along with the estimated trajectory (solid line) and the ground truth trajectory (dashed line) obtained from the *VICON*. The

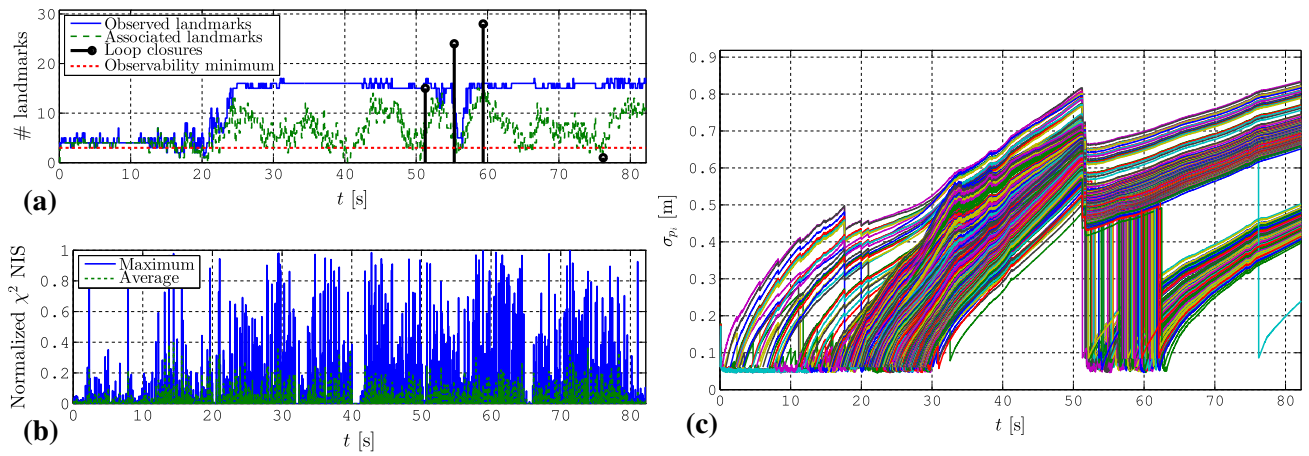


Fig. 6 Run #1—Time evolution of landmark-related variables. **a** Number of landmarks used in the Kalman filter, **b** NIS association, **c** Landmark standard deviation

coloured squares, that coincide by construction, and triangles indicate the start and end of the run, respectively. The ellipses are the 2-D projection of the 2σ uncertainty ellipsoids. The small quadrotors represent the pose of the vehicle in several instants, both with ground truth (dashed red) and SLAM estimates (solid green).

These results can be better explained by looking at Fig. 4, where the estimation error of the Earth-fixed quantities are plotted against time. First, Fig. 4a depicts the evolution of the position estimation error. Note that the vertical estimates are worse than the horizontal ones, which are quite accurate (standard deviation of 0.08 m against 0.14 m). Given that there was no motion in the vertical direction, and that the horizontal angle-of-view of the *Kinect* is greater than the vertical one, these results are not unexpected (see Fig. 2b). This is confirmed in Fig. 4b, where the attitude of the vehicle is presented in the form of the Euler angles, estimated with small error, except for the pitch angle (θ) which is off up to 8° (standard deviations of 1.7° for roll and yaw, and 2.8° for pitch). It must be noted that the calibration of the rigid transformation between the camera and the rate-gyros frames was not performed automatically, as in Kelly and Sukhatme (2011), but manually. This may explain the errors in the attitude estimation, together with the lack of vertical motion by the vehicle, and the reduced vertical angle-of-view of the *Kinect* that limits the vertical separation of landmarks, and consequently the information extractable therein.

The body-fixed velocity estimation errors and the angular rate measurement bias estimates are shown in Fig. 5. The velocity estimation error is depicted alongside the 95 % uncertainty bounds, and, even though the velocity is modelled to be constant, it follows the velocity accurately (standard deviation of 0.02 m/s in the vertical axis and 0.05 m/s in the horizontal ones). Furthermore, its uncertainty converges while generally maintaining the consistency throughout the run. The measurement bias on the right is obviously presented without ground truth, but its uncertainty can be seen to converge, confirming the results of Sect. 3.

Finally, Fig. 6 presents information about the sensor-based landmark estimates: the number of landmarks in Fig. 6a, the association quality index in Fig. 6b, and the standard deviation representing the uncertainty in Fig. 6c. The landmark uncertainty plot is of interest because it allows to identify the moments where loop closures occur, at $t \approx 50$ s, as well as the growth rate of the uncertainty of the non-visible landmarks that is smaller than 1 m after a displacement of 12 m. Loop closures occur when the vehicle is exposed to previously visited terrain, and the old landmarks reappear in its estimated field-of-view being associated with freshly obtained measurements. The solid blue line in Fig. 6a represents the number of measured landmarks in each instant. From the association process results a smaller number of actual associations, presented in dashed green. It can be noticed that the number of landmarks is almost always greater than three, which means that at least one of the observability conditions derived in Sect. 3 is fulfilled. At last, Fig. 6b presents the average and maximum normalized innovation squared (NIS) value for each set of associations, given by

$$\mathrm{NIS}_{i_k} = \boldsymbol{\nu}_{i_k}^T \boldsymbol{\Sigma}_{\nu_{i_k}} \boldsymbol{\nu}_{i_k}.$$

This value is employed as an association index, and is known to follow a chi-squared distribution with three degrees of freedom for 3-D measurements. It is compared to the 95 % percentile threshold to determine if an association is valid or not. In the figure, the average and maximum NIS values are also normalized with the value corresponding to the 95 % percentile. It is observed that although the maximum values approach the threshold, the average association is well below

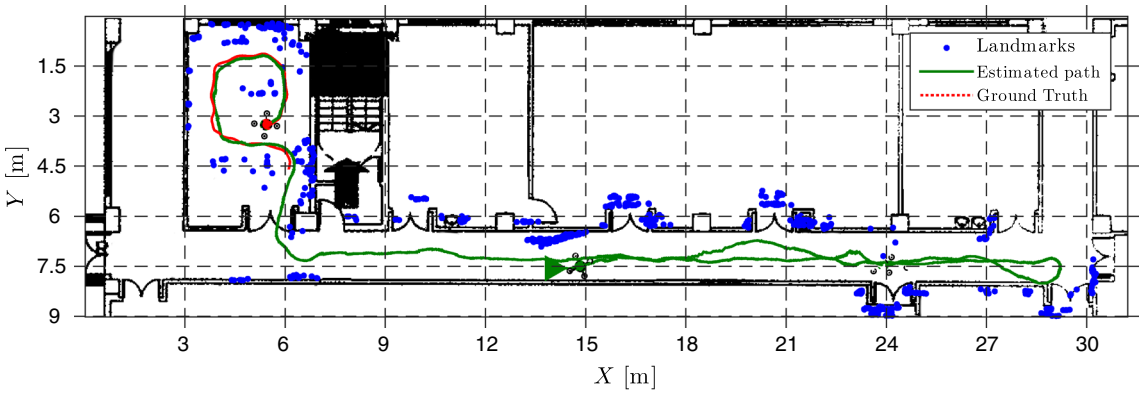


Fig. 7 Run #2—Top view of the Earth-fixed map with vehicle trajectory and ground truth against the floor blueprint

it, indicating that the landmark associations and the overall sensor-based algorithm are consistent.

6.2 Run #2

In the second run, the vehicle starts inside the usable area of the lab and, after a small lap inside the room, it is hand-driven outside the room into the corridor shown in Fig. 7 travelling a total of 70 meters at an average speed of 0.4 m/s. The estimated Earth-fixed map and trajectory are presented in Fig. 7. From the top view depicted in that figure it is possible to recognize the floor plan that is in the background. Please note that the *VICON* system is only available inside the working area of the lab, and that is why ground truth disappears after around 45 s into the run. In this figure (as in Fig. 3b), it is noticeable that a number of landmarks appears inside the lab due to the presence of obstacles not indicated in the blueprint, especially the protection net whose knots are identified as landmarks several times. This is shown in Fig. 8a which depicts the output of the RGB-D camera along with the extracted SURF features in the beginning of the experiment. The green circles represent extracted features that are not matched to the camera pointcloud as they are out of range, the blue circles are the detected features, and the red are used as measurements in the filter. If on one hand the lab has plenty of features, the corridor outside it presents a completely different environment. The most recognizable areas are the ones close to doors, as shown in Fig. 8b. This can be confirmed in Fig. 7 where it can be seen that the greater agglomerations of landmarks occur near doors. In fact, outside those areas, there are situations where the number of detected features is very low (Fig. 8c), which, combined with the limited field-of-view of the camera impairs the performance of the algorithm. On the other hand, the tiled walls of the corridor present an excellent field for feature extraction. However, the environment can become cluttered with features similar to one another (see Fig. 8d), and, although only a reduced set of landmarks from the detected set are actually used as measurements, there can be some difficulty in associating them.

Once again, the horizontal estimates are more accurate than the vertical ones, as can be seen in Fig. 9b, where the Earth-fixed estimates are presented in the same fashion as those of Fig. 4. When compared with the first run, one concludes that the Euler angles are better estimated in this run, although pitch and roll are still erroneous momentarily. The yaw angle is quite accurately estimated once more. Looking at the products of the sensor-based filter, it is easily checked that the linear velocity estimates presented in Fig. 9a are consistent with the ground truth.

In summary, these experiments were intended to allow the practical validation of the algorithm proposed in this paper, assess its performance on real conditions and evaluate two of the main issues on any SLAM algorithm: consistency and convergence of the solution. The comparison with the ground truth enabled the performance evaluation, showing that the sensor-based algorithm performs well. The consistency of the sensor-based SLAM filter is also evaluated by the ability to close a loop after revisiting terrain. Finally, the convergence is demonstrated by the decreasing in the uncertainty of all variables, except when landmarks are not visible, confirming the theoretical results of Sect. 3.

7 Conclusions

This paper presented an algorithm for simultaneous localization and mapping, reporting the design, analysis, and validation stages. The main contribution is the novel sensor-based filter with global stability and convergence guarantees. These stem from the provided formal proofs for the observability of the underlying nonlinear system, through the establishment of both necessary and sufficient conditions. This was accomplished by designing the filtering framework in the space of the sensors, which enabled the construction of a nonlinear system that can be regarded as linear time-

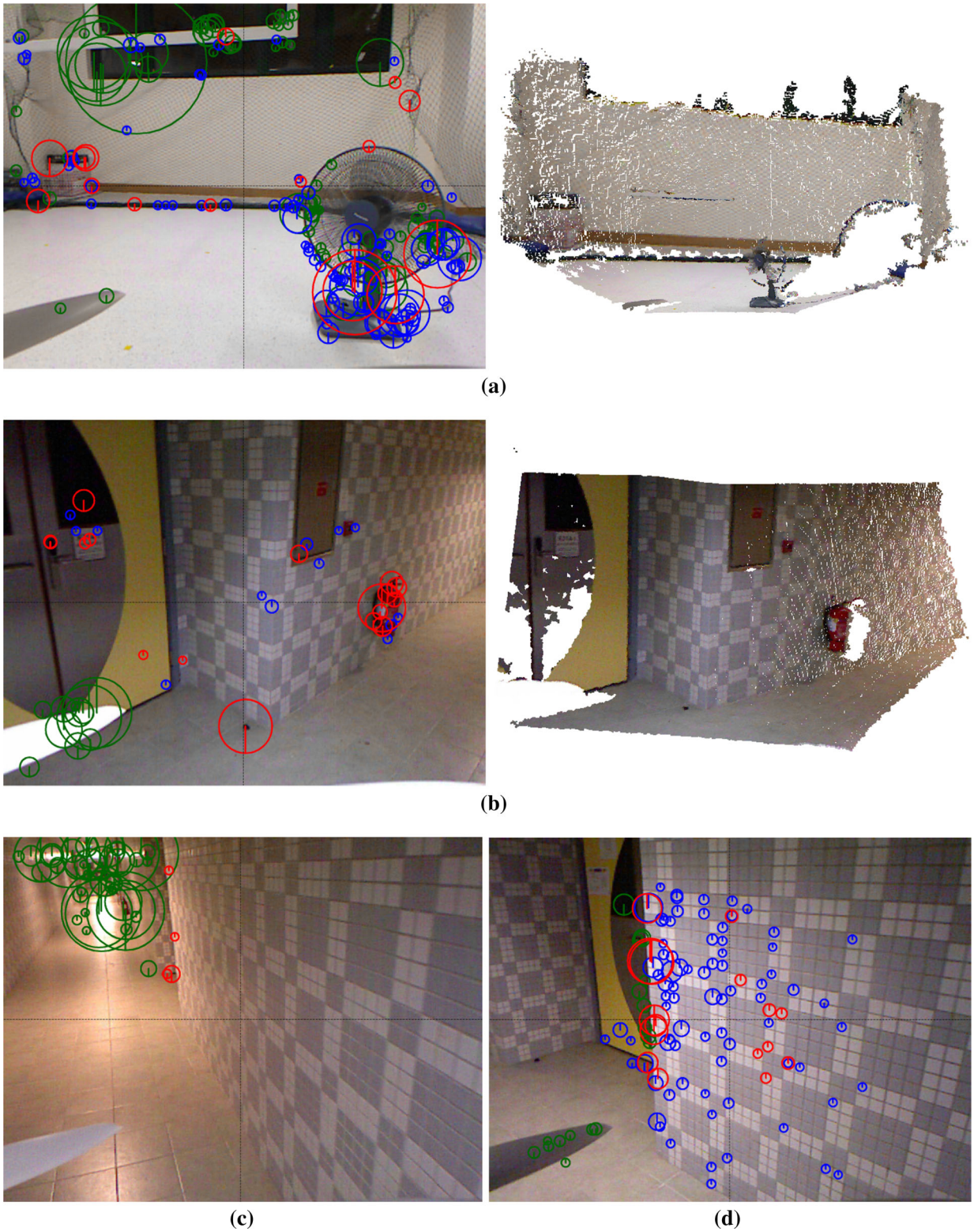


Fig. 8 Run #2—RGB pictures (*left*) with detected features and coloured pointclouds (*right*) obtained by the *Kinect* at the beginning of the experiment (*top*) and while in the corridor (*bottom*). The *size of the circles* indicates the magnitude of the Hessian of each feature and the direction of the vertical radius represents the Laplacian (positive-up, negative-down). *Green circles* represent features extracted but not matched in the point-cloud, *blue circles* are detected landmarks, and red circles are used as measurements. **a** RGB-D output at the beginning of the experiment, **b** RGB-D output near a corridor door, **c** A zone with few relevant landmarks, **d** A zone cluttered with similar features (Color figure online)

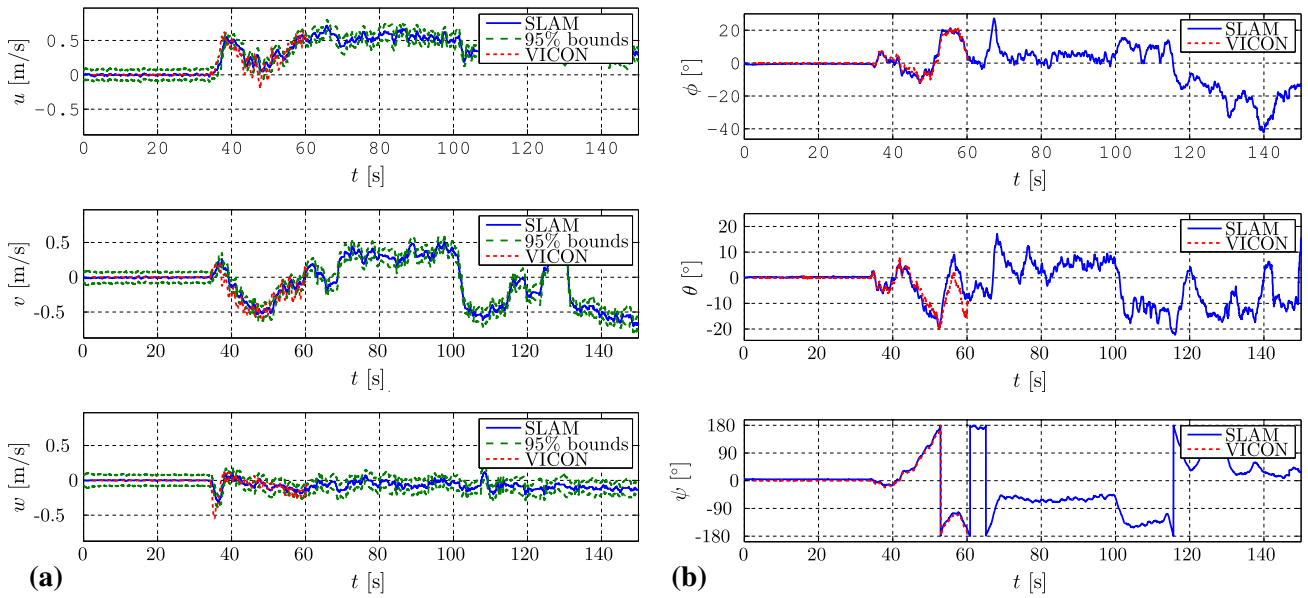


Fig. 9 Run #2—Time evolution of the attitude and body-fixed linear velocity estimates against the ground truth, **a** velocity with 2σ bounds, **b** attitude

varying, thus allowing the use of the well-established Kalman filter theory. The sensor-based approach allows the establishment of purely kinematic systems without the need for any pose representation, as well as the direct use of the measurements in the state, especially as the sensors used in the vehicle directly provide measurements in vehicle-centered coordinates. Even though the pose of the vehicle is deterministically defined by the origin and attitude of the body-fixed frame by construction, an Earth-fixed estimate may be desirable, which can also be obtained by an additional algorithm proposed by the authors.

The theoretical results were verified in real world experiments in an indoor environment, validating the convergence of the error dynamics of the Kalman filter for LTV systems, as predicted by the observability analysis presented in Sect. 3. The performance and consistency evaluation showed a decreasing uncertainty in every variable except the non-visible landmarks, as well as the production of a consistent map. This was complemented by the use of the ETM algorithm working in conjunction with the sensor-based SLAM filter, as the Earth-fixed estimation is also shown to be performing accurately in comparison with the ground truth data.

One direction of future research is the optimization of the algorithm for real-time operation, which is of paramount importance for achieving a truly online filter, thus paving the way for its use in small autonomous unmanned aerial vehicles. Further additions can also be made to the sensor suite to improve performance, such as altitude and accelerometer measurements.

Acknowledgments The authors would like to thank everyone involved in the development of the software and hardware used in the data acquisition for this paper, namely André Oliveira and Bruno Cardeira of the DSOR Lab and David Cabecinhas of the SCORE Lab. This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) through ISR under LARSyS UID/EEA/50009/2013, and through IDMEC, under LAETA UID/EMS/50022/2013 contracts, by the University of Macau Project MYRG2015-00126-FST, and by the Macao Science and Technology Development Fund under Grant FDCT/048/2014/A1. The work of P. Lourenço was supported by the PhD Student 1136 Grant SFRH/BD/89337/2012 from FCT.

Appendix: Proof of Theorem 4

Consider again system (6) and recall that it is related by a Lyapunov transformation to system (4). Thus, the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t, \mathbf{y}(t)), \boldsymbol{\mathcal{C}}(t))$ implies the uniform complete observability of the pair $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$. Hence, the proof will continue with the transformed system.

A pair $(\boldsymbol{\mathcal{A}}(t, \mathbf{y}(t)), \boldsymbol{\mathcal{C}}(t))$ is uniformly completely observable if and only if there exist positive constants δ and α such that for all $t \geq t_0$ and for all unit vectors $\mathbf{c}$ the quadratic form $\mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c}$ is greater than or equal to α, i.e., if and only if

$$\underset{\substack{\delta>0\\ \alpha>0}}{\exists} \quad \underset{t\geq t_0}{\forall} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_z}\\ \|\mathbf{c}\|=1}}{\forall} : \quad \mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} \geq \alpha,$$

meaning that, in contrast with the observability definition used in Theorem 1, the Gramian must have uniform bounds at all times. The proof follows by exhaustion, by analysing the quadratic form in the previous expression for all the possible cases of unit vectors $\mathbf{c}$ for all time. It is recalled from (8) and (9) that $\mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} = \int_t^{t+\delta} \|\mathbf{f}(\tau, t)\|^2 \, d\tau$, with

$$\|\mathbf{f}(\tau,t)\|^2 = \sum_{i=1}^{N_O} \left\|\mathbf{c}_i - \mathbf{f}_{v_i}(\tau,t)\right\|^2, \tag{25}$$

where

$$\mathbf{f}_{v_i}(\tau,t) := \int_t^\tau \mathbf{R}_m(\sigma_i)\left(\mathbf{c}_v + \mathbf{S}[\mathbf{p}_i(\sigma_i)]\mathbf{c}_b\right) d\sigma_i.$$

In order to be able to proceed, Batista et al. (2011a, Proposition 4.2) is needed. It states that if it is possible to find a positive constant β such that $\|\frac{\partial^i}{\partial\tau^i}\mathbf{f}(\tau,t)\| \geq \beta$ then there exists a $\gamma > 0$ such that $\|\mathbf{f}(t,t+\delta)\| \geq \gamma$ as long as $\frac{\partial^j}{\partial\tau^j}\mathbf{f}(\tau,t)|_{\tau=t} = 0$ for all $j < i$ and the norm of the $(i+1)$-th derivative is upper bounded. It is possible to see that this applies to $\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}$, and therefore it states conditions upon which the quadratic form in analysis is lower bounded if (25) is lower bounded, uniformly in time. Then, it suffices to prove that $\|\mathbf{f}(\tau,t)\| \geq \alpha_1$ with $\tau \in \mathcal{T}_\delta$ and $\alpha_1 > 0$ for every possible $\mathbf{c}$, provided that the conditions of Batista et al. (2011a, Proposition 4.2) are satisfied uniformly in time.

Consider the case where there exists at least one $i \in \mathcal{I}_O$ such that $\|\mathbf{c}_i\| \geq \alpha_p$ for some $0 < \alpha_p < 1$ and no restriction is imposed on $\mathbf{c}_v$ and $\mathbf{c}_b$. Then,

$$\|\mathbf{f}(t,t)\|^2 = \sum_{j=1}^{N_O} \|\mathbf{c}_j\|^2 \geq \|\mathbf{c}_i\|^2 \geq \alpha_p^2$$

for all t. Consider now that $\|\mathbf{c}_i\| < \alpha_p$, for all $i \in \{1,\ldots,N_O\}$ and for some $0 < \alpha_p < 1$. For that purpose, let $\mathbf{c}_p$ and $\mathbf{f}_v(\tau,t)$ be the stacking of all the $\mathbf{c}_i$ and $\mathbf{f}_{v_i}(\tau,t)$, respectively, and note that $\mathbf{f}(\tau,t) = \mathbf{c}_p - \mathbf{f}_v(\tau,t)$. Then, it is possible to write

$$\begin{aligned}\|\mathbf{f}(\tau,t)\|^2 &\geq \|\mathbf{f}_v(\tau,t)\|\left(\|\mathbf{f}_v(\tau,t)\| - 2\|\mathbf{c}_p\|\right)\\ &\geq \|\mathbf{f}_v(\tau,t)\|\left(\|\mathbf{f}_v(\tau,t)\| - 2\sqrt{N_O}\alpha_p\right).\end{aligned}$$

Assuming that $\|\mathbf{f}_v(\tau,t)\|$ is lower bounded at $\tau = t^*$ in $\mathcal{T}_\delta$ by some $\alpha_2 > 0$ so that α_p can be chosen to satisfy $\alpha_p \leq \frac{\alpha_2}{4\sqrt{N_O}}$, it follows that

$$\left\|\mathbf{f}(t^*,t)\right\|^2 \geq \frac{1}{2}\alpha_2^2.$$

It is necessary then to show that $\|\mathbf{f}_v(t^*,t)\| \geq \alpha_2 \geq 4\sqrt{N_O}\alpha_p$ under the conditions of the theorem. In this situation $\mathbf{f}_v(\tau,t)$ is zero for $\tau = t$, and therefore Batista et al. (2011a, Proposition 4.2) applies once again. Thus, it is enough to show that the norm of the derivative of $\mathbf{f}_v(\tau,t)$, given by

$$\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|^2 = \sum_{i=1}^{N_O} \|\mathbf{c}_v + \mathbf{S}[\mathbf{p}_i(\tau)]\mathbf{c}_b\|^2, \tag{26}$$

is lower bounded by some $\alpha_3 > 0$ for a $\tau \in \mathcal{T}_\delta$. Then, there exists an α_2 that bounds $\|\mathbf{f}_v(\tau,t)\|$ below for some $t^* \in \mathcal{T}_\delta$ and α_p can be chosen accordingly. Under the restriction $\|\mathbf{c}_i\| < \alpha_p$ for all $i \in \mathcal{I}_O$, there exist three possibilities for $\mathbf{c}$, depending on $\mathbf{c}_v$ and $\mathbf{c}_b$. The first case is set by $\|\mathbf{c}_v\| \geq \alpha_v$ and $\|\mathbf{c}_b\| < \alpha_b$ for some α_v and α_b in the interval $(0,1)$. It is possible to write

$$\begin{aligned}\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|^2 &\geq \sum_{i=1}^{N_O} \|\mathbf{c}_v\|\left(\|\mathbf{c}_v\| - 2\|\mathbf{p}_i(\tau)\|\|\mathbf{c}_b\|\right)\\ &\geq N_O\alpha_v\left(\alpha_v - 2P_M\alpha_b\right)\\ &\geq \frac{N_O}{2}\alpha_v^2,\end{aligned}$$

where Assumption 2 was employed, and α_v was chosen so that it satisfies $\alpha_b \leq \frac{\alpha_v}{4P_M}$. The second case, where $\|\mathbf{c}_v\| < \alpha_v$ for some $0 < \alpha_v < 1$ and $\|\mathbf{c}_b\| \geq \alpha_b$ for some $0 < \alpha_b < 1$, and the third case, where $\|\mathbf{c}_v\| \geq \alpha_v$ and $\|\mathbf{c}_b\| \geq \alpha_b$, can be analysed together. Consider then that $\|\mathbf{c}_b\| \geq \alpha_b$ and that $0 \leq \|\mathbf{c}_v\| \leq 1$, in which case the conditions of the theorem must be addressed separately. The first condition of the theorem states that there are at least three landmarks ($N_O \geq 3$) that are sufficiently away from collinearity, i.e., the plane defined by the vectors that unite each pair of landmarks is well-defined uniformly in time and sufficiently away from degenerating into a line. This can be illustrated by taking the cross product between each vector that defines the plane and $\mathbf{c}_b$, summing the norms of the results, and noting that the lowest possible value for this sum occurs when $\mathbf{c}_b$ is collinear with the largest of these vectors, say, $\mathbf{c}_b = \pm\frac{\|\mathbf{c}_b\|}{\|\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1)\|}(\mathbf{p}_1(t_1) - \mathbf{p}_3(t_1))$. This worst case can be substituted in the conditions of the theorem to yield

$$\|\mathbf{S}[\mathbf{p}_1(t_1) - \mathbf{p}_2(t_1)]\mathbf{c}_b\| \geq \frac{\alpha_b\alpha_*}{2P_M}$$

where Assumption 2 was used to show that $\|\mathbf{p}_1(t_1) - \mathbf{p}_3(t_1)\| \leq 2P_M$. This means that

$$\|\mathbf{S}[\mathbf{p}_1(t_1) - \mathbf{p}_2(t_1)]\mathbf{c}_b\| + \|\mathbf{S}[\mathbf{p}_1(t_1) - \mathbf{p}_3(t_1)]\mathbf{c}_b\| \geq \frac{\alpha_b\alpha_*}{2P_M}$$

for any $\mathbf{c}_b$ with norm greater than $0 < \alpha_b < 1$. It is a matter of algebraic manipulation to obtain

$$\begin{aligned}&\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_2(t_1)]\mathbf{c}_b + \mathbf{c}_v)\|\\ &\quad + \|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_3(t_1)]\mathbf{c}_b + \mathbf{c}_v)\| \geq \frac{\alpha_b\alpha_*}{2P_M}\end{aligned}$$

from the previous expression, which can be further manipulated using the triangle inequality to yield

$$2\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_2(t_1)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_3(t_1)]\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{2P_M}. \quad (27)$$

There are two possible conclusions to draw from this inequality depending on the values of $\|\mathbf{S}[\mathbf{p}_2(t_1)]\mathbf{c}_b + \mathbf{c}_v\|$ and $\|\mathbf{S}[\mathbf{p}_3(t_1)]\mathbf{c}_b + \mathbf{c}_v\|$. If either is greater than or equal to some $\alpha_{p_1} > 0$, (26) is lower bounded by α_{p_1} at $\tau = t_1$, as intended. On the other hand, if both are smaller than $\alpha_{p_1} > 0$, then (27) leads to

$$\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\|\Big|_{\tau=t_1} \geq \|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{4P_M} - \alpha_{p_1} \geq \frac{\alpha_b \alpha_*}{8P_M}$$

if α_{p_1} is chosen so that $\alpha_{p_1} \leq \frac{\alpha_b \alpha_*}{8P_M}$. The second condition can be treated in similar terms. Consider, without loss of generality that the landmark set that is considered is $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_1(t_2)\}$. Using the same reasoning used for the analysis of the first conditions, it is possible to write

$$\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_2(t_1)]\mathbf{c}_b + \mathbf{c}_v)\| + \|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_1(t_2)]\mathbf{c}_b + \mathbf{c}_v)\| \geq \frac{\alpha_b \alpha_*}{2P_M},$$

which can also be manipulated to obtain a lower bound to the sum of the norms, yielding

$$2\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_2(t_1)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_1(t_2)]\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{2P_M}. \quad (28)$$

Once again, the analysis depends on the behaviour of $\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\|$ and $\|\mathbf{S}[\mathbf{p}_2(\tau)]\mathbf{c}_b + \mathbf{c}_v\|$. If either is greater than or equal to some $\alpha_{p_1} > 0$, then, for $\tau = t_1$, $\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\| \geq \alpha_{p_1}$. Conversely, if both are smaller than α_{p_1}, then from (28) it is possible to write

$$\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\|\Big|_{\tau=t_1} \geq \|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{4P_M} - \alpha_{p_1} \geq \frac{\alpha_b \alpha_*}{8P_M},$$

for the previous choice of α_{p_1}. Finally, for the third condition, the set of landmarks considered is $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$. In the same line of thought followed previously, one knows that

$$\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_1(t_2)]\mathbf{c}_b + \mathbf{c}_v)\| + \|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}[\mathbf{p}_1(t_3)]\mathbf{c}_b + \mathbf{c}_v)\| \geq \frac{\alpha_b \alpha_*}{2P_M}.$$

Applying the triangle inequality to the previous expression as was done when manipulating the similar expressions for the first two conditions yields

$$2\|\mathbf{S}[\mathbf{p}_1(t_1)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_1(t_2)]\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}[\mathbf{p}_1(t_3)]\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{2P_M}. \quad (29)$$

If $\|\mathbf{S}[\mathbf{p}_1(\tau)]\mathbf{c}_b + \mathbf{c}_v\| \geq \alpha_{p_1}$ for $\tau = t_2, t_3$ or both, then (26) is bounded by α_{p_1} in one or both of these instants. However, if $\|\mathbf{S}[\mathbf{p}_1(\tau)]\mathbf{c}_b + \mathbf{c}_v\| < \alpha_{p_1}$ for all $\tau \in \mathcal{T}_\delta$, $\tau \neq t_1$ and nothing is imposed for $\tau = t_1$, (29) leads to

$$\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\|^2\Big|_{\tau=t_1} \geq \frac{\alpha_b \alpha_*}{4P_M} - \alpha_{p_1} \geq \frac{\alpha_b \alpha_*}{8P_M}$$

where the fact that $\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\| = \|\mathbf{S}[\mathbf{p}_1(\tau)]\mathbf{c}_b + \mathbf{c}_v\|$ was used, and α_{p_1} was chosen so that $\alpha_{p_1} \leq \frac{\alpha_b \alpha_*}{8P_M}$. In these three conditions, positive lower bounds were found for $\left\| \frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t) \right\|$, which, from Batista et al. (2011a, Proposition 4.2) implies that $\mathbf{f}_v(\tau, t)$ is also lower bounded and, therefore, $\mathbf{f}(\tau, t)$ has a lower bound too within the possibilities studied. All possible cases are now enumerated, hence the sufficiency of the conditions of Theorem 4 for the pair $(\mathcal{A}(t, \mathbf{y}(t)), \mathcal{C}(t))$ to be uniformly completely observable is proved. Because (5) is a Lyapunov transformation, the uniform complete observability of the transformed pair implies uniform complete observability of the $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$, and thus the proof of sufficiency of the conditions of the theorem is concluded.

The proof of the necessity of the conditions of the theorem follows by contraposition. The hypothesis that the conditions of the theorem do not hold is considered, and it is shown that this implies that the pair $(\mathcal{A}(t), \mathcal{C}(t))$ is not uniformly completely observable.

Consider then that the conditions of the theorem do not hold. The negation of these conditions means that all the landmarks in the set $\mathcal{I}_O$ are not sufficiently away from collinearity. Consider the alternative definition for a landmark, given by

$$\mathbf{p}_i(t) = p_{d_i}(t)\mathbf{d} + p_{d_i^{\perp 1}}(t)\mathbf{d}^{\perp 1} + p_{d_i^{\perp 2}}(t)\mathbf{d}^{\perp 2} + \mathbf{p}_1(t_0),$$

where $\{\mathbf{d}, \mathbf{d}^{\perp 1}, \mathbf{d}^{\perp 2}\} \in \mathbb{R}^3$ are orthonormal vectors that span the planes defined by all the combinations of three landmarks in $\mathcal{I}_O$, and $p_{d_i}(t)$, $p_{d_i^{\perp 1}}(t)$, $p_{d_i^{\perp 2}}(t) \in \mathbb{R}^3$. The negated conditions imply that, either there are no landmarks which immediately renders the pair unobservable or, considering, without loss of generality, that $\mathbf{d}$ is the predominant direction of landmarks, there are at least two directions $\mathbf{d}^{\perp 1}$ and $\mathbf{d}^{\perp 2}$ along which the vector $\mathbf{p}_i(t) - \mathbf{p}_1(t_0)$ is bounded and as small as wanted for some $t \geq t_0$, i.e.,

$$\underset{\beta>0}{\forall}\ \underset{i\in\mathcal{I}_O}{\forall}\ \underset{t\geq t_0}{\exists}:\quad |p_{d_i^{\perp 1}}(t)|<\beta \wedge |p_{d_i^{\perp 2}}(t)|<\beta. \tag{30}$$

The proof follows by showing that this implies that the pair in analysis is not uniformly completely observable, which can be stated as

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall}\ \underset{t\geq t_0}{\exists}\ \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_z}\\ \|\mathbf{c}\|=1}}{\exists}:\quad \mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}<\epsilon.$$

For that purpose, the expansion of $\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}$ is considered for some unit $\mathbf{c}$, yielding the following expression that can be obtained by substituting (25) into (8),

$$\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}=\int_t^{t+\delta}\sum_{i=1}^{N_O}\left\|\mathbf{c}_i-\int_t^{\tau}\mathbf{R}_m(\sigma_i)\left(\mathbf{c}_v+\mathbf{S}[\mathbf{p}_i(\sigma_i)]\mathbf{c}_b\right)d\sigma_i\right\|^2 d\tau.$$

In the following steps, a particular $\mathbf{c}$ is chosen and the negated condition (30) is used. Considering that $\mathbf{c}_i=\mathbf{0}$ for all $i\in\mathcal{I}_O$, and using the Cauchy–Schwartz inequality in the previous expression leads to

$$\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}\leq\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\|\mathbf{c}_v-\mathbf{S}[\mathbf{c}_b]\mathbf{p}_i(\sigma)\|^2\,d\sigma\,d\tau.$$

Recall that $\mathbf{R}_m(t)$ maintains the norm, which is why it is omitted in the previous expression. Let a be a real constant, and choose $\mathbf{c}_b=a\mathbf{d}$ and $\mathbf{c}_v=a\mathbf{S}[\mathbf{d}]\mathbf{p}_1(t_0)$ such that $\|\mathbf{c}\|=1$. This yields

$$\begin{aligned}
&\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}\\
&\quad\leq a^2\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\left\|\mathbf{S}[\mathbf{d}]\left(p_{d_i^{\perp 1}}(\sigma)\mathbf{d}^{\perp 1}+p_{d_i^{\perp 2}}(\sigma)\mathbf{d}^{\perp 2}\right)\right\|^2 d\sigma\,d\tau\\
&\quad\leq a^2\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\left(|p_{d_i^{\perp 1}}(\sigma)|+|p_{d_i^{\perp 2}}(\sigma)|\right)^2 d\sigma\,d\tau\\
&\quad\leq a^2N_O\int_t^{t+\delta}\int_t^{\tau}(2\beta)^2 d\sigma\,d\tau,
\end{aligned}$$

which can be rewritten as

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall}\ \underset{t\geq t_0}{\exists}\ \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_z}\\ \|\mathbf{c}\|=1}}{\exists}:\quad \mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}<\epsilon.$$

as $\beta:=\frac{1}{a\delta}\sqrt{\frac{\epsilon}{2N_O}}$. With this step it was shown that the transformed system cannot be uniformly completely observable if the conditions of the theorem do not apply, which, along with the fact that the Lyapunov transformation preserves the observability properties of the system, leads to the necessity of the conditions for the uniform complete observability of the nonlinear system (4), regarded as LTV, and thus concluding the proof.

References

Anderson, B. D. O. (1971). Stability properties of Kalman–Bucy filters. *Journal of the Franklin Institute*, *291*(2), 137–144.

Bailey, T., & Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics & Automation Magazine*, *13*(3), 108–117.

Bailey, T., Nieto, J., Guivant, J., Stevens, M., & Nebot, E. (2006, Oct.). Consistency of the EKFSLAM algorithm. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3562–3568).

Batista, P., Silvestre, C., & Oliveira, P. (2010). Optimal position and velocity navigation filters for autonomous vehicles. *Automatica*, *46*(4), 767–774.

Batista, P., Silvestre, C., & Oliveira, P. (2011a). On the observability of linear motion quantities in navigation systems. *Systems & Control Letters*, *60*(2), 101–110.

Batista, P., Silvestre, C., & Oliveira, P. (2011b). Single range aided navigation and source localization: Observability and filter design. *Systems & Control Letters*, *60*(8), 665–673.

Bay, H., Ess, A., Tuytelaars, T., & Gool, L. V. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, *110*(3), 346–359.

Bishop, A., & Jensfelt, P. (2009, May). A stochastically stable solution to the problem of robocentric mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 1615–1622). Kobe.

Brockett, R. (1970). *Finite dimensional linear systems*. New York: Wiley.

Castellanos, J., Martinez-Cantin, R., Tardós, J., & Neira, J. (2007). Robocentric map joining: Improving the consistency of EKF-SLAM. *Robotics and Autonomous Systems*, *55*(1), 21–29.

Davison, A., Reid, I., Molton, N., & Stasse, O. (2007). Monoslam: Real-time single camera slam. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(6), 1052–1067.

Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localisation and mapping (SLAM): Part I the essential algorithms. *IEEE Robotics & Automation Magazine*, *13*(2), 99–110.

Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., & Burgard, W. (2012, May). An evaluation of the RGB-D SLAM system. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1691–1696). St. Paul, MA.

Gelb, A. (1974). *Applied optimal estimation*. Cambridge: MIT Press.

Guerreiro, B. J., Batista, P., Silvestre, C., & Oliveira, P. (2012, June). Sensor-based simultaneous localization and mapping—part II: Online inertial map and trajectory estimation. In *Proceedings of the 2012 American Control Conference* (pp. 6334–6339). Montréal.

Guerreiro, B. J., Batista, P., Silvestre, C., & Oliveira, P. (2013). Globally asymptotically stable sensor-based simultaneous localization and mapping. *IEEE Transactions on Robotics*, *29*(6), 1380–1395.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, *22*(5), 728–740.

Huang, G., Mourikis, A. I., & Roumeliotis, S. I. (2010). Observability-based rules for designing consistent EKF SLAM estimators. *The International Journal of Robotics Research*, *29*(5), 502–528.

Huang, S., & Dissanayake, G. (2007). Convergence and consistency analysis for extended Kalman filter based SLAM. *IEEE Transactions on Robotics*, *23*(5), 1036–1049.

Jones, E., & Soatto, S. (2011). Visual-inertial navigation, mapping and localization: A scalable realtime causal approach. *The International Journal of Robotics Research*, *30*(4), 407–430.

Julier, S., & Uhlmann, J. (2001, May). A counter example to the theory of simultaneous localization and map building. In *Proceedings of the 2001 IEEE International Conference on Robotics and Automation (ICRA)* (Vol. 4, pp. 4238–4243). Seoul

Kelly, J., & Sukhatme, G. S. (2011). Visual-inertial sensor fusion: Localization, mapping and sensor-tosensor self-calibration. *The International Journal of Robotics Research*, *30*(1), 56–79.

Kermorgant, O., & Chaumette, F. (2011, May). Multisensor data fusion in sensor-based control: Application to multi-camera visual servoing. In *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 4518–4523). doi:10.1109/ICRA.2011.5979715

Khalil, H. (2002). *Nonlinear systems* (3rd ed.). Upper Saddle River: Prentice Hall.

Leutenegger, S., Furgale, P., Rabaud, V., Chli, M., Konolige, K., & Siegwart, R. (2013, June). Keyframe-based visual-inertial slam using nonlinear optimization. In *Proceedings of Robotics: Science and Systems*. Berlin.

Lourenço, P., Guerreiro, B. J., Batista, P., Oliveira, P., & Silvestre, C. (2013a, July). 3-D inertial trajectory and map online estimation: Building on a GAS sensor-based SLAM filter. In *Proceedings of the 2013 European Control Conference* (pp. 4214–4219). Zurich.

Lourenço, P., Guerreiro, B. J., Batista, P., Oliveira, P., & Silvestre, C. (2013b, June). Preliminary results on globally asymptotically stable simultaneous localization and mapping in 3-D. In *Proceedings of the (2013) American Control Conference* (pp. 3093–3098). Washington D.C.

May, S., Droeschel, D., Holz, D., Fuchs, S., Malis, E., Nüchter, A., et al. (2009). Three-dimensional mapping with time-of-flight cameras. *Journal of Field Robotics*, *26*(11–12), 934–965.

Schönemann, P. H. (1966). A generalized solution of the orthogonal procrustes problem. *Psychometrika*, *31*(1), 1–10.

Se, S., Lowe, D., & Little, J. (2002). Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *The International Journal of Robotics Research*, *21*(8), 735–758.

Umeyama, S. (1991). Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(4), 376–380.

Victorino, A., Rives, P., & Borrelly, J.-J. (2000). Localization and map building using a sensorbased control strategy. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vol. 2, pp. 937–942). doi:10.1109/IROS.2000.893139

Weiss, L., Sanderson, A. C., & Neuman, C. P. (1987). Dynamic sensor-based control of robots with visual feedback. *IEEE Journal on Robotics and Automation*, *3*(5), 404–417.

Weiss, S., Achtelik, M. W., Lynen, S., Achtelik, M. C., Kneip, L., Chli, M., et al. (2013). Monocular vision for long-term micro aerial vehicle state estimation: A compendium. *Journal of Field Robotics*, *30*(5), 803–831.

Pedro Lourenço received the B.Sc. and M.Sc. degrees in Aerospace Engineering in 2010 and 2012, respectively, both from Instituto Superior Técnico (IST), Lisbon, Portugal, on the course of which he received the Merit Award. He is currently pursuing his Ph.D. degree in Electrical and Computer Engineering with the Institute for Systems and Robotics, Associated Laboratory for Robotics and Systems in Engineering and Science, at IST. His research interests include sensor-based simultaneous localization and mapping with RGB-D, vision, ranging, and bearing sensors.

Bruno J. Guerreiro received the Licenciatura and Ph.D. degrees in Electrical and Computer Engineering in 2004 and 2013, respectively, both from Instituto Superior Técnico (IST) of the University of Lisbon, Portugal, and his Ph.D. thesis received the Best Thesis Award by the Portuguese Robotics Society. He is currently a post-doctoral researcher at the SCORE laboratory of the Faculty of Science and Technology, at the University of Macau, and also an Invited Assistant Professor with the Department of Electrical and Computer Engineering at IST. He has been involved in several research and development projects regarding the use of autonomous vehicles for automatic inspection of buildings, infrastructures and archaeological sites. His research interests include the sensor-based navigation and control of autonomous vehicles, laser calibration methods and 3-D reconstruction, as well as nonlinear and predictive control.

Pedro Batista received the Licenciatura degree in electrical and computer engineering in 2005 and the Ph.D. degree in 2010, both from Instituto Superior Técnico (IST), Lisbon, Portugal. From 2004 to 2006, he was a Monitor with the Department of Mathematics, IST, where he is currently an Invited Assistant Professor with the Department of Electrical and Computer Engineering. His research interests include sensor-based navigation and control of autonomous vehicles. Dr. Batista has received the Diploma de Mérito twice during his graduation and his Ph.D. thesis was awarded the Best Robotics Ph.D. Thesis Award by the Portuguese Society of Robotics.

Paulo Oliveira received the Ph.D. degree in Electrical and Computer Engineering, from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2002. He is an Associate Professor in the Department of Mechanical Engineering of IST and researcher in the Institute of Mechanical Engineering, IST. He also collaborates with the Institute for Systems and Robotics, IST. His areas of scientific activity are Mechatronics with special focus on the fields of Autonomous Vehicles, Robotics, Sensor Fusion, Navigation, Positioning, and Nonlinear Estimation. Over the last 25 years he participated in more than 30 European and Portuguese research projects and he co-authored more than 50 journal and 150 conference papers.

Carlos Silvestre received the Licenciatura degree in Electrical Engineering from the Instituto Superior Tecnico (IST) of Lisbon, Portugal, in 1987 and the M.Sc. degree in Electrical Engineering and the Ph.D. degree in Control Science from the same school in 1991 and 2000, respectively. In 2011 he received the Habilitation in Electrical Engineering and Computers also from IST. Since 2000, he is with the Department of Electrical Engineering of the Instituto Superior Tecnico, where he is currently an Associate Professor of Control and Robotics on leave. Since 2012 he is an Associate Professor of the Department o Electrical and Computers Engineering of the Faculty of Science and Technology of the University of Macau. Over the past years, he has conducted research on the subjects of navigation guidance and control of air and underwater robots. His research interests include linear and nonlinear control theory, coordinated control of multiple vehicles, gain scheduled control, integrated design of guidance and control systems, inertial navigation systems, and mission control and real time architectures for complex autonomous systems with applications to unmanned air and underwater vehicles.