

New Design Techniques for Globally Convergent Simultaneous Localization and Mapping: Analysis and Implementation

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

Abstract This chapter presents an overview of algorithms deeply rooted in a sensor-based approach to the SLAM problem that provide global convergence guarantees and allow for the use of partially observable landmarks. The presented algorithms address the more usual range-and-bearing SLAM problem, either in 2-D using a LiDAR or in 3-D using an RGB-D camera, as well as the range-only and bearing-only SLAM problems. For each of these formulations a nonlinear system is designed, for which state and output transformations are considered together with augmented dynamics, in such a way that the underlying system structure can be regarded as linear time-varying for observability analysis and filter design purposes. This naturally allows for the design of Kalman filters with, at least, globally asymptotically stable error dynamics, for which several experimental and simulated trials are presented to highlight the performance and consistency of the obtained filters.

C. Silvestre on leave from the Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

P. Lourenço · B. Guerreiro · P. Batista
Laboratory for Robotics and Engineering Systems, Institute for Systems and Robotics,
Lisbon, Portugal
e-mail: plourenco@isr.tecnico.ulisboa.pt

B. Guerreiro
e-mail: bguerreiro@isr.tecnico.ulisboa.pt

P. Batista (✉)
Department of Electrical and Computer Engineering, Instituto Superior Técnico,
Universidade de Lisboa, Lisbon, Portugal
e-mail: pbatista@isr.tecnico.ulisboa.pt

P. Oliveira
Department of Mechanical Engineering, Instituto Superior Técnico,
Universidade de Lisboa, Lisbon, Portugal
e-mail: pjcro@isr.tecnico.ulisboa.pt

C. Silvestre
Department of Electrical and Computer Engineering of the Faculty of Science
and Technology, University of Macau, Macao, China
e-mail: csilvestre@umac.mo

1 Introduction

When navigating in an unknown environment, the mapping of that environment and the localization within that map have been shown to be dependent on each other, and a more intricate solution than the traditional navigation strategies has to be considered: the simultaneous localization and mapping (SLAM). The research community has devoted significant effort to the study of this problem, for which the seminal works that established the statistical foundations for describing the relationships between landmarks and their correlations include [16, 34, 35]. Further research showed that a full and consistent solution to this problem would require all the vehicle pose and map variables to be considered together, which renders the problem intrinsically nonlinear. Among the many technical solutions emerging from this challenge are the extended Kalman filter (EKF) [11], the use of Rao-Blackwellized particle filters as in FastSLAM [31], or the use of information filters [37]. A detailed survey on most of the used techniques, sensors, applications, and challenges can be found in [9], wherein several other more specialized surveys are referenced.

An important component of most SLAM algorithms is the association between the landmarks measurements and the state landmarks, or when it is necessary to close a loop, being one of the major sources of inconsistency in SLAM algorithms. Several strategies are widely used, such as the simplistic nearest neighbor (NN) validation gating, the joint compatibility branch and bound (JCBB) [33], and the combined constrained data association (CCDA) [3], while other strategies such as those in [2] use sensors that provide unique characteristics of each measured landmark.

The SLAM problem can also be characterized by the fundamental type of measurements available for filtering, usually referred to as landmarks. When the landmark measurements have a lower dimension than the considered mapping space (a single noise-free observation provides only a line or surface as an estimate for the relative position of the landmark), the resulting subproblems are usually divided into range-only SLAM (RO-SLAM) and bearing-only SLAM (BO-SLAM), while the more usual SLAM problem is sometimes referred to as range-and-bearing SLAM (RB-SLAM) to underline the case where all the relative coordinates of measured landmarks are readily available.

A fundamental aspect of the RO-SLAM problem is the absence of association errors, as the information carried by the ranging signals allows the unambiguous association of measurements and the corresponding states at all times, which also enables error-free loop closing. Conversely, the initialization of a RO-SLAM strategy may represent a challenge, and most RO-SLAM solutions rely on some form of initializing procedure in order to create a new landmark in the state, such as the trilateration (in 2-D) with ranges from different instants [1]. As the RO-SLAM problem bears resemblance to the sensor networks problem, in the sense that an agent receives signals from a network of sensors, the two ideas have been used in conjunction in works such as [14, 26].

The BO-SLAM problem is more challenging than the RO-SLAM, because an observation of the former corresponds to an unbounded set. Besides triangulation,

more advanced probabilistic approaches can be used to address this issue, such as a sum of Gaussians [24], deterring the initialization until the obtained estimate is well approximated by a Gaussian [4], or using multiple hypothesis [36]. Nowadays, BO-SLAM is most associated with monocular vision [21], as there has been an intense research effort in this particular application. One of the most relevant developments is presented in [12], being the first real-time SLAM algorithm using only a camera as data source. Other interesting approaches include [15] which introduces a closed-form pose-chain optimization algorithm that uses sparse graphs as well as appearance-based loop detection, and ORB-SLAM [32], that uses ORB features, which are rotation invariant and have faster extraction than SURF features [7].

In general, the SLAM problem has a nonlinear nature which can be tackled using EKF-based solutions, as well as other similar filters that usually imply the linearization of the dynamics or measurement models, resulting in lack of consistency and convergence guarantees [20, 22]. To address these issues, several authors have analyzed the unobservable space of the error system obtained after linearization, showing that it has smaller dimensionality than that of the underlying nonlinear error system [19]. This yields erroneous gains of information, and a possible solution proposed by the same authors is to select the linearization points of the EKF that ensure the same dimensions between the mentioned unobservable spaces. While focusing on automatic calibration, [23] addresses the observability properties of nonlinear SLAM based on differential geometry, as it is a necessary (but not sufficient) condition for the convergence of any filtering algorithm. Another approach to this issue is to use the so-called robocentric or sensor-based approach, as firstly proposed in the robocentric map joining algorithm [10]. This algorithm can improve the consistency of the regular EKF, yet, as it still considers the estimation of the (unobservable) incremental pose in the filter state, it cannot provide guarantees of convergence. Other algorithms that provide some guarantees of convergence include methods that usually assume that the linearized system matrices are evaluated at the ideal values of the state variables [13, 20], or other that resort to stochastic stability concepts assuming access to both linear and angular velocity measurements [8]. Nevertheless, a formal theoretical result on global convergence for EKF-based SLAM algorithms is still absent from the literature.

This chapter presents an overview of algorithms that are rooted in a sensor-based approach to the SLAM problem that can be used in aerial robots. The usual SLAM approach requires a single filter to maintain estimates of the map and vehicle pose along with the respective covariances and cross-covariances. However, it is possible to use an alternate formulation that uses a Kalman filter (KF) that explores the linear time-varying (LTV) nature of the sensor-based SLAM system and analyzing its observability. This formulation is deeply rooted in the theory of sensor-based control, exploring the fact that vehicle-fixed sensors provide measurements that in this approach do not need to be transformed to an Earth-fixed frame [38]. The first of these purely sensor-based SLAM filters was proposed for the two-dimensional RB-SLAM problem [18], suppressing pose representation in the state and therefore avoiding singularities and nonlinearities, subsequently extended for 3-D in [29] using an RGB-D camera. These works have provided necessary and sufficient conditions

for observability of the nonlinear error system, for which a KF can be designed that yields global asymptotic stability (GAS).

Regarding the less straightforward formulations of the SLAM problem, the work presented in [27] introduces a novel RO-SLAM algorithm that eliminates the landmark initialization problem through the establishment of global convergence results. As in the previous algorithms, the proposed 3-D sensor-based formulation avoids the representation of the pose of the vehicle in the state, allowing the direct use of odometry-like information that is usually expressed in body-fixed coordinates. Finally, [25] proposes a 3-D BO-SLAM algorithm with exponentially fast global convergence and allows for undelayed initialization at any depth. Building on the previous approaches, this algorithm uses a state augmentation and an output transformation that lead to the design of an LTV system whose observability conditions are given in a constructive analysis with clear physical insight. These two solutions are influenced by the source localization algorithms proposed in [5, 6], as a similar state augmentation is used to achieve the global convergence results.

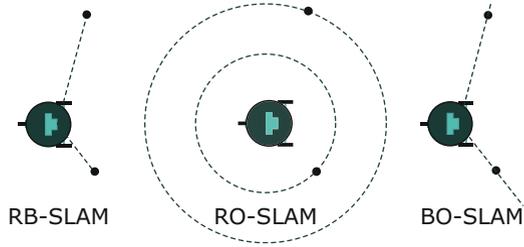
Building on the results mentioned above, the main contributions of this chapter include: (i) the consolidation and definition of a class of sensor-based SLAM problems such as 2-D and 3-D RB-SLAM, RO-SLAM, and BO-SLAM; (ii) a collection of physically intuitive and constructive observability results; (iii) the filter implementation details that ensure a global asymptotic stability of the respective error dynamics; (iv) an alternative method to obtain the Earth-fixed quantities from the results of the sensor-based filters; (v) a collection of experimental and simulation results that validate and illustrate the main properties and performance of the proposed filters.

The remaining of the chapter is organized as follows. Section 2 introduces the sensor-based SLAM problems while their observability analysis is detailed in Sect. 3. The sensor-based filter implementation details are provided in Sect. 4 along with the Earth-fixed trajectory and map algorithm. Finally, the main results stemming from the proposed algorithms are depicted in Sect. 5 and some concluding remarks are given in Sect. 6.

2 Sensor-Based SLAM

Following the previous discussion, this section details the design of dynamical systems as part of the sensor-based simultaneous localization and mapping filters using only one source of external environment perception, capable of either measuring relative positions, ranges, or bearings, apart from vehicle motion information. Let $\{B\}$ denote the body-fixed frame and $\{E\}$ denote the inertial/Earth-fixed frame, whereas $(\mathbf{R}(t), {}^E\mathbf{p}(t))$ represents the transformation from $\{B\}$ to $\{E\}$ and, therefore, the pose of the vehicle. The attitude is given by the rotation matrix $\mathbf{R}(t) \in SO(n)$ and the position is given by ${}^E\mathbf{p}(t) \in \mathbb{R}^n$, with $n = 2, 3$. The former satisfies $\dot{\mathbf{R}}(t) = \mathbf{R}(t)\mathbf{S}(\boldsymbol{\omega}(t))$, where $\boldsymbol{\omega}(t) \in \mathbb{R}^{n_\omega}$, $n_\omega = 1, 3$, is the angular velocity of the vehicle expressed in body-fixed coordinates. The environment is characterized by point landmarks that may be

Fig. 1 For RB-SLAM, sensors measure the position of a landmark relative to the vehicle, for RO-SLAM, the distance to a landmark, and for BO-SLAM, the relative direction to a landmark



naturally extracted or artificially placed. These N landmarks constitute the set \mathcal{M} and are denoted by ${}^E \mathbf{p}_i(t) \in \mathbb{R}^n$ or $\mathbf{p}_i(t) \in \mathbb{R}^n$, respectively describing the landmark location in frame $\{E\}$ or in frame $\{B\}$. As in the Earth-fixed frame landmarks are assumed static, considering the motion of the landmark in $\{B\}$, it is possible to write

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

where $\mathbf{S}(\cdot)$ is a skew-symmetric matrix that encodes the cross product for $n_\omega = 3$, $\mathbf{S}(\boldsymbol{\omega}) = \begin{bmatrix} 0 & -\boldsymbol{\omega} \\ \boldsymbol{\omega} & 0 \end{bmatrix}$ for $n_\omega = 1$, and $\mathbf{v}(t) \in \mathbb{R}^n$ is the linear velocity of the vehicle expressed in $\{B\}$. Typically, $\boldsymbol{\omega}(t)$ and $\mathbf{v}(t)$ are both known inputs and need not be estimated. However, this situation may change, depending on the information provided by the system outputs.

Consider now that the N landmarks are divided in two different sets depending on their visibility status: $\mathcal{M}_O := \{1, \dots, N_O\}$ containing the N_O observed or visible landmarks and $\mathcal{M}_U := \{N_O + 1, \dots, N\}$ containing the N_U unobserved, or non-visible, ones. Landmarks belonging to \mathcal{M}_O will have some kind of system output associated, which leads to the definition of

$$\mathbf{y}_i(t) = \mathbf{f}(\mathbf{p}_i(t)), \quad i \in \mathcal{M}_O \tag{2}$$

where $\mathbf{y}_i(t)$ can be equal to $\mathbf{p}_i(t)$, $\|\mathbf{p}_i(t)\|$, or $\frac{\mathbf{p}_i(t)}{\|\mathbf{p}_i(t)\|}$ according to the version of sensor-based SLAM to be designed (see Fig. 1). Combining this information it is now possible to write the generic nonlinear system

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t)) \mathbf{p}_i(t) - \mathbf{v}(t) & i \in \mathcal{M} \\ \mathbf{y}_j(t) = \mathbf{f}(\mathbf{p}_j(t)) & j \in \mathcal{M}_O \end{cases} \tag{3}$$

If $\boldsymbol{\omega}(t)$ is an input and $\mathbf{v}(t)$ is either added as a state with constant dynamics or kept as an input, then the first equation in (3) can be considered linear for observability analysis purposes. The main problem rests with the output equation that may be nonlinear. In that case, further action is necessary to obtain a linear-like system: state augmentation and/or output transformation. This subject will be addressed in the sequel.

2.1 Range-and-Bearing SLAM

Range-and-bearing sensors provide the most information possible in terms of point-based maps, and, therefore, may be exploited in order to estimate more quantities. Following that line of reasoning, consider that the measured angular velocity is corrupted with a constant bias, i.e.,

$$\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_\omega(t). \quad (4)$$

Further consider that the linear velocity is not directly measured, and, as such, needs to be estimated. Then, the generic nonlinear system (3) becomes

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}_m(t)) \mathbf{p}_i(t) - \mathbf{g}(\mathbf{p}_i(t), \mathbf{b}_\omega(t)) - \mathbf{v}(t) & i \in \mathcal{M} \\ \dot{\mathbf{v}}(t) = \mathbf{0} \\ \dot{\mathbf{b}}_\omega(t) = \mathbf{0} \\ \mathbf{y}_j(t) = \mathbf{p}_j(t) & j \in \mathcal{M}_O \end{cases}, \quad (5)$$

where

$$\mathbf{g}(\mathbf{p}_i(t), \mathbf{b}_\omega(t)) := \begin{cases} -\mathbf{S}(1) \mathbf{p}_i(t) b_\omega(t), & n = 2 \\ \mathbf{S}(\mathbf{p}_i(t)) \mathbf{b}_\omega(t), & n = 3 \end{cases} \quad (6)$$

For the visible landmarks, the nonlinear term $\mathbf{g}(\mathbf{p}_i(t), \mathbf{b}_\omega(t))$ can be written as $\mathbf{g}(\mathbf{y}_i(t), \mathbf{b}_\omega(t))$, which, even though it is still nonlinear it may be considered as linear time-varying for observability purposes as $\mathbf{p}_i(t)$ is available. For the remaining landmarks, this term is still nonlinear.

It should be mentioned that each 2-D landmark could be accompanied by additional features in the form of directional landmarks. That specific case is addressed in [18], where more information on both the system and its observability analysis can be found.

2.2 Range-only SLAM

In the case of range-only external perception, it is necessary to have some measure of the linear movement, and therefore the linear velocity is here also used as an output (it must be measured). Considering the measurement model, $y_i(t) = \|\mathbf{p}_i(t)\|$ for all $i \in \mathcal{M}_O$, which in this case is nonlinear, a state augmentation is proposed, yielding the new state

$$\begin{cases} \mathbf{x}_{L_i}(t) := \mathbf{p}_i(t) \\ \mathbf{x}_V(t) := \mathbf{v}(t) \\ x_{R_i}(t) := \|\mathbf{x}_{L_i}(t)\| \end{cases}. \quad (7)$$

It is a simple matter of computation to then write the full system dynamics for this extended state, resulting in the new system

$$\begin{cases} \dot{\mathbf{x}}_{L_i}(t) = -\mathbf{S}(\boldsymbol{\omega}(t)) \mathbf{x}_{L_i}(t) - \mathbf{x}_V(t) & i \in \mathcal{M} \\ \dot{\mathbf{x}}_V(t) = \mathbf{0} \\ \dot{x}_{R_i}(t) = -\frac{\mathbf{y}_V^T(t)}{\mathbf{x}_{R_i}(t)} \mathbf{x}_{L_i}(t) & i \in \mathcal{M} \\ \mathbf{y}_V(t) = \mathbf{x}_V(t) \\ y_j(t) = \mathbf{p}_j(t) & j \in \mathcal{M}_O \end{cases} \quad (8)$$

As in range-and-bearing SLAM, there are still nonlinear terms in the dynamics that need to be taken care of. In this case, the term $\frac{\mathbf{y}_V^T(t)}{\mathbf{x}_{R_i}(t)} \mathbf{x}_{L_i}(t)$ can be rewritten for the visible landmarks, as the denominator is one of the system outputs, i.e., $\frac{\mathbf{y}_V^T(t)}{\mathbf{x}_{R_i}(t)} \mathbf{x}_{L_i}(t) = \frac{\mathbf{y}_V^T(t)}{\mathbf{y}_i(t)} \mathbf{x}_{L_i}(t)$ for all $i \in \mathcal{M}_O$. With this change, the system pertaining to this subset of the landmarks is now linear time-varying for observability analysis purposes, as the dynamics only depend on known system inputs and outputs.

2.3 Bearing-only SLAM

As in the previous situation, bearing-only measurements require the linear velocity to be measured. In this case, it will be introduced as an input, accompanying the angular velocity as the inputs of the dynamical system. As for the measurement model, the output of the nonlinear system is now $\mathbf{y}_i(t) = \frac{\mathbf{p}_i(t)}{\|\mathbf{p}_i(t)\|} := \mathbf{b}_i(t)$. As this output is nonlinear, the simple output transformation $\mathbf{p}_i(t) - \mathbf{b}_i(t)\|\mathbf{p}_i(t)\| = \mathbf{0}$ is considered together with the state augmentation,

$$\begin{cases} \mathbf{x}_{L_i}(t) := \mathbf{p}_i(t) \\ x_{R_i}(t) := \|\mathbf{x}_{L_i}(t)\| \end{cases}, \quad (9)$$

making it possible to avoid the nonlinearity in the output. This yields the new system

$$\begin{cases} \dot{\mathbf{x}}_{L_i}(t) = -\mathbf{S}(\boldsymbol{\omega}(t)) \mathbf{x}_{L_i}(t) - \mathbf{v}(t) & i \in \mathcal{M} \\ \dot{x}_{R_i}(t) = -\frac{\mathbf{x}_{L_i}^T(t)}{\mathbf{x}_{R_i}(t)} \mathbf{v}(t) & i \in \mathcal{M} \\ \mathbf{0} = \mathbf{x}_{L_j} - \mathbf{b}_j(t)\mathbf{x}_{R_j}(t) & j \in \mathcal{M}_O \end{cases} \quad (10)$$

which is still nonlinear in the dynamics. Notice that here $\boldsymbol{\omega}(t)$ and $\mathbf{v}(t)$ are both inputs, and that $\mathbf{b}_i(t)$ is a measurement. With this in mind, it is possible to replace the nonlinear term in the dynamics of the state $x_{R_i}(t)$ with information that is measured, thus yielding a linear time-varying structure to the system part that relates to the

visible landmarks. This is performed by noting that $\frac{\mathbf{x}_{L_i}^T(t)}{\mathbf{x}_{R_i}(t)} \mathbf{v}(t) = \mathbf{b}_i^T(t) \mathbf{v}(t)$ for all $i \in \mathcal{M}_O$.

With these manipulations, it is possible to obtain dynamical systems for sensor-based SLAM that, when looking only at the visible landmarks, resemble linear time-varying systems that mimic the original nonlinear systems. The following section deals with their observability. It must be noted, however, that the state augmentations are not enforced in any way, i.e., the relation $x_{R_i}(t) = \|\mathbf{x}_L(t)\|$ is not explicitly used in the dynamics.

3 Observability

The systems designed in the previous section resemble LTV systems. However, due to the presence of non-visible landmarks in all of them, there are still nonlinear terms that prevent the use of linear tools for analysis. The quantities associated with the non-visible landmarks are, by definition, not observable. Therefore, it is reasonable to discard them from the state when analyzing observability (see [18, 27, 29] for previous successful applications of this approach by the authors). This yields a reduced system of the form

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}(t, \mathbf{y}(t), \mathbf{u}(t))\mathbf{x}(t) + \mathbf{B}(t, \mathbf{y}(t), \mathbf{u}(t))\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}(t, \mathbf{y}(t), \mathbf{u}(t))\mathbf{x}(t) \end{cases}, \quad (11)$$

whose dynamics, even though still nonlinear, do not depend on the state itself but on inputs and outputs. As both the inputs and outputs are known functions of time, then, for observability analysis and observer design purposes, the system (11) can in fact be considered as linear time-varying.

Having reached this stage in the system design and analysis, it is possible to use [5, Lemma 1] to ascertain whether and in what conditions the system is observable by studying its observability Gramian. Figure 2 summarizes the whole process presented in the previous section and further explained here. After analyzing the observability of the LTV system, the next step is to study the observability of the original nonlinear system, starting with a comparison between the state of the LTV system and that of the nonlinear system. This includes investigating the conditions in which the state augmentation relations become naturally imposed by the dynamics. It is in that case that the states of the two systems become equivalent, which validates the augmentation approach in the sense that a state observer with uniformly globally asymptotically/exponentially stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, and the estimation error converges asymptotically/exponentially fast for all initial conditions.

The final step before proceeding to observer design is to study the uniform complete observability (UCO) of the LTV system. This is a stronger form of observability that is necessary to guarantee the global asymptotical/exponential stability of the

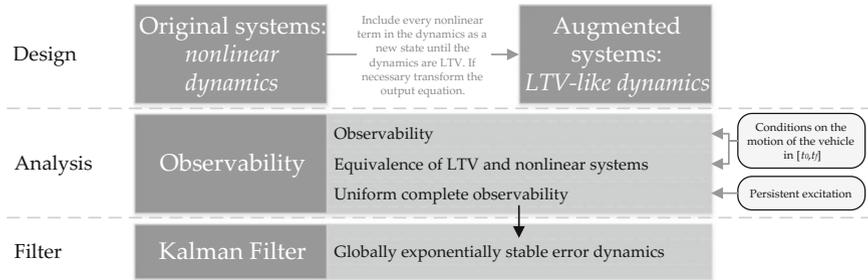


Fig. 2 Schematic description of the process of designing a globally convergent sensor-based SLAM filter

Kalman filter. For this purpose, uniform bounds on the observability Gramian calculated on a moving time interval are investigated. Due to this uniformity, the resulting conditions are more demanding, which can be regarded as persistent excitation-like conditions.

In range-and-bearing SLAM, the conditions for observability depend on the number of landmarks observed, due to the fact that the nonmeasured quantities (linear velocity and rate-gyro bias) affect all the landmarks. On the other hand, the conditions for the observability of both bearing-only and range-only SLAM do not depend on the number of landmarks, as they are independent of each other. There is another important distinction between the two classes of problems pertaining the quantity of information made available with each measurement. In range-and-bearing SLAM one single measurement provides all the necessary information to estimate the position of a landmark, even though several landmarks are needed for immediate full state recovery, whereas in range-only and bearing-only SLAM measurements from several viewpoints have to be acquired to allow for landmark estimation. This is also the case for range-and-bearing SLAM when the number of available landmarks is not enough to guarantee observability without motion.

The remainder of this section summarizes the theoretical results corresponding to each version of SLAM designed previously and the associated necessary and sufficient conditions.

2-D range-and-bearing SLAM [18, Theorems 2-4 and 6]

- Observability and Equivalence UCO* Two landmarks are visible or one landmark is visible and there is an instant when its derivative is nonzero.
- UCO* Two landmarks are visible or one landmark is visible and its derivative is sufficiently away from zero, uniformly in time.

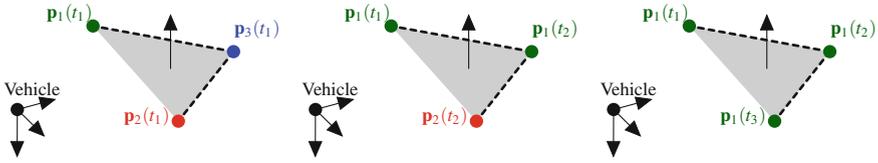


Fig. 3 Geometrical interpretation of the observability conditions for range-and-bearing SLAM in 3-D

3-D range-and-bearing SLAM [29, Theorems 1-4]

Observability and Equivalence Three landmarks form a plane in one observation, two observations of two landmarks form a plane or three observations of one landmark form a plane (see Fig. 3).

UCO The vectors defined by the three landmarks that form a plane (regardless of the observation moment) are sufficiently away from collinearity, uniformly in time.

3-D Range-only SLAM [27, Theorems 1-4]

Observability and Equivalence The linear velocity in three observation moments spans \mathbb{R}^3 (see left side of Fig. 4 for an example of the bidimensional case).

UCO The vectors defined by the three velocity measurements are sufficiently away from co-planarity so that the spanned space does not degenerate in time.

3-D Bearing-only SLAM [25, Theorems 1-3]

Observability and Equivalence Two different absolute bearings to one landmark are measured (see right side of Fig. 4).

UCO The variation in the bearing measurement is sufficiently away from zero to not degenerate in time.

These results allow for the design and implementation of a Kalman filter with globally asymptotically/exponentially stable error dynamics.

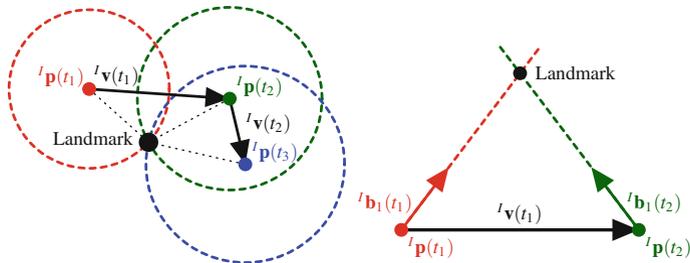


Fig. 4 Examples of trilateration (*left*) and triangulation (*right*) for positioning a landmark in 2-D. These demonstrate the geometric interpretation of the derived observability conditions, i.e., the importance of proper vehicle motion for univocal determination of the coordinates of a landmark

4 Filter Design and Implementation

Considering a discrete time implementation of the filter, let T_s denote the sampling period of the synchronized array of sensors used in each solution, noting that a multi-rate implementation can be devised. The system is discretized using the forward Euler discretization, with special care when there is a rotation of a landmark from one instant to the following, where it is considered that the angular velocity is constant over each sampling interval, i.e., $\mathbf{R}_{k+1}^T \mathbf{R}_k = \exp(-\mathbf{S}(\boldsymbol{\omega}_k) T_s)$. Considering additive disturbances, the generic discretized system is given by

$$\begin{cases} \mathbf{x}_{F_{k+1}} = \mathbf{F}_{F_k} \mathbf{x}_{F_k} + \mathbf{G}_{F_k} \mathbf{v}_k + \boldsymbol{\xi}_k \\ \mathbf{y}_{k+1} = \mathbf{H}_{F_{k+1}} \mathbf{x}_{F_{k+1}} + \boldsymbol{\theta}_{k+1} \end{cases}, \tag{12}$$

where the dynamics matrices have the structure

$$\mathbf{F}_{F_k} = \begin{bmatrix} \mathbf{F}_k & \mathbf{0}_{n_O \times n_U} \\ \mathbf{F}_{UO_k} & \mathbf{F}_{U_k} \end{bmatrix}, \quad \mathbf{G}_{F_k} = \begin{bmatrix} T_s \mathbf{B}_k \\ T_s \mathbf{B}_{U_k} \end{bmatrix}, \quad \mathbf{H}_{F_k} = [\mathbf{C}_k \quad \mathbf{0}_{n_O \times n_U}], \tag{13}$$

and where \mathbf{F}_k is the discretized version of matrix \mathbf{A} in (11), which accounts for the observable part of the system dynamics, whereas matrices with subscripts $(\cdot)_U$ and $(\cdot)_{UO}$ denote the unobservable states and cross terms, respectively. Also, the vectors $\boldsymbol{\xi}_k$ and $\boldsymbol{\theta}_k$ represent the model disturbance and measurement noise, respectively, assumed to be zero-mean discrete white Gaussian noise with covariances $\boldsymbol{\Xi}_k$ and $\boldsymbol{\Theta}_k$. These will depend on the actual formulation of SLAM in question. However, in all the formulations described in this chapter, the dynamics depends on the actual inputs and outputs, which means that the noise characterization is not exact. As a general rule of thumb, these noise parameters can be calibrated a priori with a Monte Carlo analysis and with actual measurements to better cope with this issue.

In particular, recalling the three presented SLAM problems, both the RB-SLAM and BO-SLAM have similar structure for the system matrix, defined as

$$\hat{\mathbf{F}}_{F_k} = \begin{bmatrix} \mathbf{F}_{L_k} & \mathbf{0}_{n_L \times (n_V + n_R)} \\ \mathbf{0}_{(n_V + n_R) \times n_L} & \mathbf{I}_{(n_V + n_R)} \end{bmatrix}, \quad (14)$$

although the state vectors have diverse variables and dimensions. On the other hand, the RO-SLAM has a more intricate structure, defined as

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{F}_{L_k} & T_s \mathbf{A}_{LV} & \mathbf{0}_{n_L \times n_R} \\ \mathbf{0}_{n_V \times n_L} & \mathbf{I}_3 & \mathbf{0}_{n_L \times n_R} \\ T_s \mathbf{A}_{RL_k} & \mathbf{0}_{n_R \times n_V} & \mathbf{I}_{n_R} \end{bmatrix}, \quad (15)$$

with $\mathbf{F}_{L_k} = \text{diag}(\mathbf{R}_{k+1}^T \mathbf{R}_k, \dots, \mathbf{R}_{k+1}^T \mathbf{R}_k)$, see [25, 27] for the remaining matrices.

From these discrete LTV systems, the filter prediction and update steps are computed using the standard equations of the Kalman filter for LTV systems [17], with the detail that the non-visible landmarks must be propagated in open loop. Nevertheless, particularly in RB-SLAM and BO-SLAM, prior to the update step it might be necessary to associate the landmark measurements with the state landmarks, either for a simple update step or for a more intricate loop closing procedure.

To complement the sensor-based filter, the authors proposed a strategy to obtain estimates of the pose of the vehicle and the Earth-fixed landmark map, with uncertainty characterization, denoted as Earth-fixed trajectory and map (ETM) estimation algorithm. Considering the relation between landmarks expressed in the two working frames and noting that ${}^E \mathbf{p}_{i_k}$ is constant, it is possible to write the error function

$${}^E \mathbf{e}_{i_k} = {}^E \hat{\mathbf{p}}_{i_{k-1}} - \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} - {}^E \hat{\mathbf{p}}_k, \quad (16)$$

which can be minimized as in the optimization problem presented in [30, Sect. 3]. This yields the optimal rotation and translation given a map expressed in the Earth-fixed frame and in the body-fixed frame and the combined uncertainty of (16). With this information, the new Earth-fixed map can be computed using

$${}^E \hat{\mathbf{p}}_{i_k} = \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} + {}^E \hat{\mathbf{p}}_k. \quad (17)$$

An important step in this algorithm is the initialization of the Earth-fixed map, which can be computed directly from the sensor-based map in the first instant by assuming that the transformation between Earth-fixed and sensor frames is known at time k_0 , a traditional assumption in most SLAM strategies.

The pose estimates provided by this strategy are accompanied with uncertainty characterizations, following a perturbation theory approach as described in [30]. The same reasoning can be employed to obtain estimates with uncertainty description for the Earth-fixed map (see [28]).

5 Practical Examples

This section aims to provide several examples of practical implementations of the sensor-based algorithms detailed along this chapter. In particular, the results presented in this section are obtained using four sensor-based SLAM algorithms: (i) 2-D Range-and-Bearing SLAM, (ii) 3-D Range-and-Bearing SLAM, (iii) Range-Only SLAM, and (iv) Bearing-Only SLAM. It should also be stressed that each of these experiments was originally designed as a proof of concept, and as such, alternative sensors or processes for obtaining measurements can be employed. Table 1 summarizes the typical sensors used for each one of the SLAM variants discussed here, pinpointing those employed in the first three examples shown below (further details can be found in [18, 25, 27, 29] and references therein). The BO-SLAM example consists of simulation results.

The Earth-fixed estimates of the maps and vehicle trajectories are also provided for the range-and-bearing algorithms, which are obviously dependent on the performance of the ETM algorithm, 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. As the sensor-based SLAM filter does not depend on the ETM algorithm, it is argued that using this separate approach it may be possible to obtain a less uncertain Earth-fixed trajectory and landmark map. As such, 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 estimates.

5.1 Range-and-Bearing SLAM

This subsection presents experimental results from two different implementations of the sensor-based range-and-bearing SLAM filter coupled with the ETM algorithm, one in two dimensions using a LiDAR [18] and other in three dimensions [29] using an RGB-D camera.

Table 1 Measurements and their respective sensors

| Quantities | Sensors |
|-------------------|---|
| Landmark position | LiDAR (i) / RGB-D camera (ii) / Stereo or trinocular camera |
| Landmark range | Radio/acoustic transceivers (iii) |
| Landmark bearing | Radio/acoustic transceivers / Single camera |
| Linear velocity | Odometry / Optical flow (iii) |
| Angular velocity | IMU (i-iii) |

When analyzing the convergence properties of any navigation filter, one of the main goals is to observe a decreasing uncertainty in all variables. This can be seen in Fig. 9 for the 3-D RB-SLAM filter, where the uncertainty of all the vehicle related variables decreases over time, whereas the uncertainty of each landmark decreases whenever visible and increases otherwise, as shown in Fig. 5 for the 2-D RB-SLAM filter. In addition, both the sensor-based map and the result of the ETM algorithm are presented in Fig. 6, featuring the final results of the 2-D RB-SLAM experimental trials. It can be seen that both the sensor-based and ETM maps are consistent, noting also the small sensor-based uncertainty for the landmarks that were recently visible, and the large uncertainty for the landmarks that are not visible for a long time, as the vehicle progresses along environment.

As can be inferred from Fig. 5, there are several loop closure procedures during the trials, adding relevant information about the consistency of the proposed algorithms.

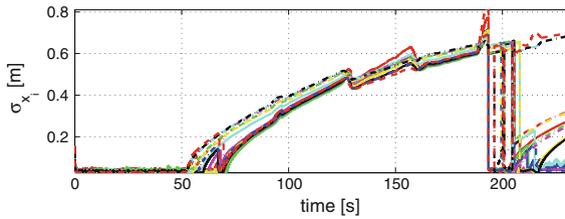


Fig. 5 2-D RB-SLAM: Uncertainty convergence in the sensor-based filter. STD of first 15 landmark positions. ©2013 IEEE. Reprinted, with permission, from [18]

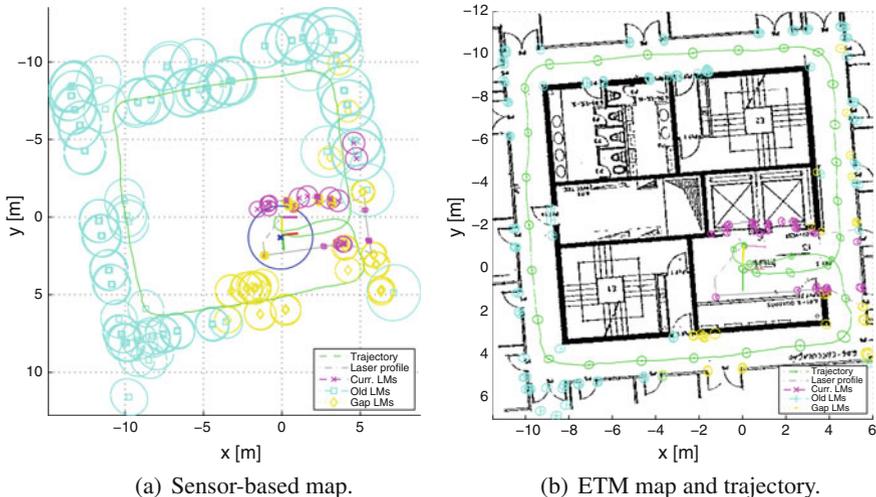


Fig. 6 2-D RB-SLAM: Map of environment in the sensor and Earth frame. This figure shows the current laser profile (in gray), the current/old/older landmarks in magenta/yellow/light blue, along with their 95% confidence bounds. ©2013 IEEE. Reprinted, with permission, from [18]

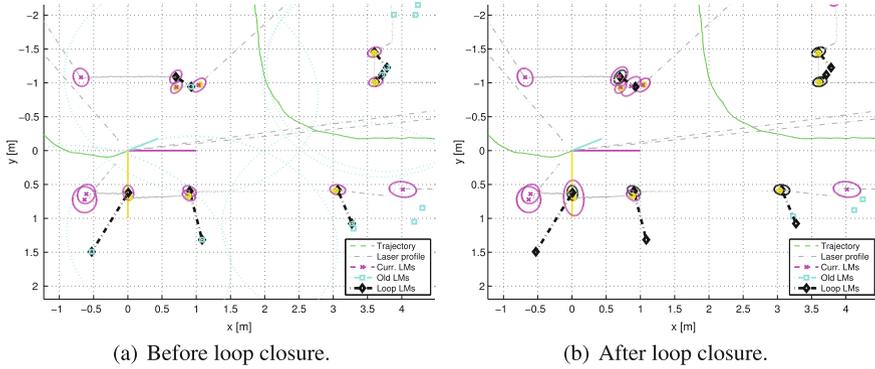


Fig. 7 2-D RB-SLAM: illustration of a loop closure. ©2013 IEEE. Reprinted, with permission, from [18]

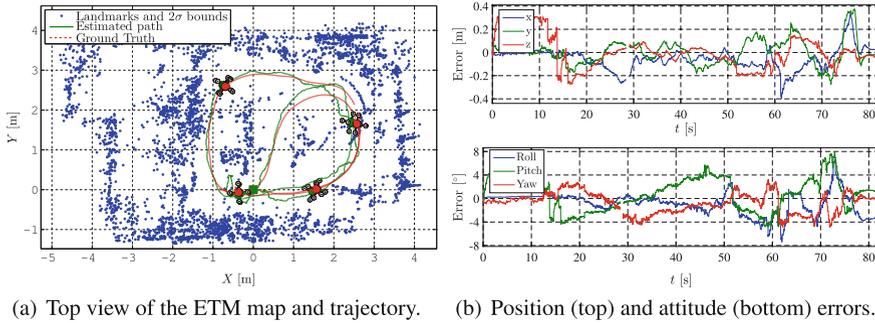


Fig. 8 3-D RB-SLAM: Earth-fixed estimates with ground truth. Reprinted from [29] with permission with permission of Springer

In one of the occasions, the moments just before and right after the loop closure are captured in Fig. 7, where a detailed version of the map of Fig. 6a at the relevant time instant is presented. The landmark associations between the current and older landmarks are shown in solid black and the fused landmarks positions and uncertainty bounds obtained after the loop closure are also depicted in solid black.

In Fig. 8a, 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 a *VICON* motion tracking system. The colored 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).

The evolution of the position estimation error is depicted in the top of Fig. 8b. It is noticeable that, even though the horizontal estimates are quite accurate, the vertical ones are worse. These results are not unexpected since there was no motion in the

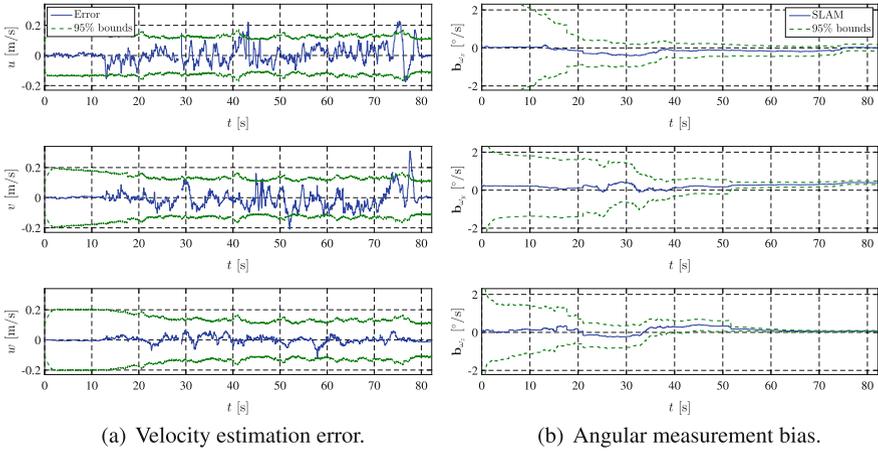


Fig. 9 3-D RB-SLAM: Time evolution of the sensor-based estimates with 2σ bounds. Reprinted from [29] with permission of Springer

vertical direction, and the horizontal angle-of-view of the *Kinect* is greater than the vertical one, thus limiting the vertical separation of landmarks and consequently the information extractable from that axis. Figure 8b confirms this assertion, where the Euler angles are presented, estimated with small error, except for the pitch angle (θ). The results of the sensor-based filter can be evaluated in Fig. 9, which depicts the body-fixed velocity estimation errors and the angular rate measurement bias estimates. The velocity estimation error is depicted alongside the 95% uncertainty bounds, and, even though the velocity is modeled 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.

5.2 Range-only SLAM

In opposition to the range-and-bearing and bearing-only examples presented here, the landmarks used in the range-only experiments are artificially placed beacons in the environment. Therefore, and taking into account the observability requirements, the trajectory of the vehicle was intended to maximize the exposure to each of the beacons, as well as to provide sufficient excitation to the filter.

As explained in Sect. 1, the initialization of the landmarks is one of the more challenging issues in RO-SLAM procedures. In this work, however, the global convergence results imply that this issue is solved as whichever the initial guess the

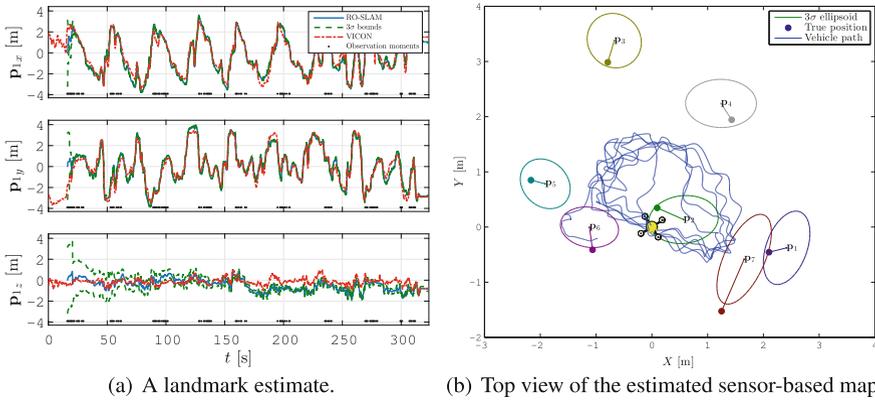


Fig. 10 RO-SLAM: One estimated landmark through time against the ground truth and the observation moments (*left*). The *top* view of the sensor-based map at the end of the run. Both include 3σ uncertainty bounds. Reprinted from [27] with permission from Elsevier

filter will converge. Figure 10a depicts exactly this, showing the estimated position (solid blue) with 3σ uncertainty bounds of one landmark against the ground truth provided by *VICON* (dashed red), and it can be seen that the convergence is very fast in the horizontal plane. Moreover, after converging, the estimates are very close to ground truth. However, in the vertical axis, the estimation is much worse, and the convergence is also slower, which is due to the less rich trajectory in that axis. The optical flow procedure employed is somewhat noisy, and as its measurements of the linear velocity are directly used in the dynamics matrix as if they were the true value, the noise can make that direction appear observable, even if the information is sparse.

Finally, an example of the estimated map in the body-fixed frame is presented in Fig. 10b. The top view of the sensor-based map is shown along with the true landmark positions and the vehicle path rotated and translated to the body-fixed frame. The colored ellipses represent orthogonal cross sections of 3σ uncertainty ellipsoids, i.e., the estimation uncertainty, and the small circles mark the true landmark positions. These experiments show the good performance of the proposed algorithm in realistic conditions, especially in the horizontal variables. The filter has some problems in the vertical coordinates due to the less rich velocity profile and noisy optical flow measurements, although with a proper trajectory the algorithm was shown to behave well [27]. Therefore, these experiments underpin the need for appropriate trajectories.

5.3 Bearing-only SLAM

This example of a sensor-based bearing-only application is performed on a simulated scenario with known association between measurements and landmarks. Figure 11

Fig. 11 BO-SLAM: The estimated map at the end of the experiment, along with the true path. ©2015 IEEE. Reprinted, with permission, from [25]

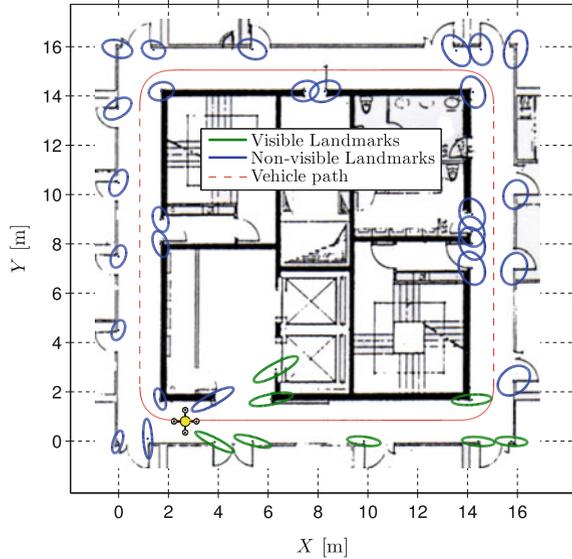
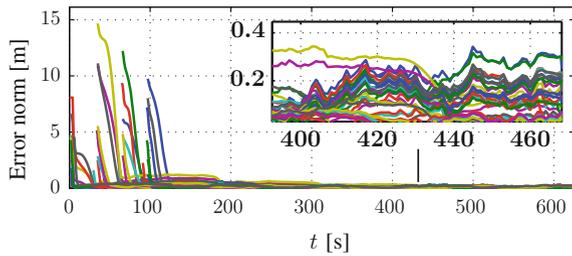


Fig. 12 BO-SLAM: The evolution of the norm of the estimation error for all the 36 landmarks. ©2015 IEEE. Reprinted, with permission, from [25]



depicts a top view of the estimated sensor-based map at the end of the run with the 95% uncertainty ellipses in green and blue depending whether they are observed in that instant or not, including the real trajectory of the vehicle in dashed red, and the pose of the vehicle at that moment, that is represented by the yellow quadrotor. Note that the ellipses surround the true values, as they should in a consistent filter. Furthermore, it can be seen that the recently or currently (re)-observed landmarks have much tighter uncertainty ellipses than older ones, demonstrating the uncertainty convergence when the observability conditions are satisfied. Finally, in Fig. 12 the estimation error for all the 36 landmarks is shown. It is noticeable that even though the initial estimate may be far off, the error will converge until after 2 laps, it is under 40 centimeters depending on how long each landmark is observed.

In an experimental application using natural landmarks extracted from the environment, it is not straightforward to guarantee observability while moving as was attempted for the range-only counterpart and many features will not have enough time to converge. However, preliminary results have shown that the landmark state

still converges without being initialized with any special care, while also showing that loop closures occur naturally throughout the runs. Hence, the algorithm recognizes previously visited places, even without a specially tailored procedure, providing a good measure of the consistency and validity of the algorithm.

6 Concluding Remarks

This chapter aimed at a broad presentation of the fundamentals behind a class of sensor-based simultaneous localization and mapping filters with global convergence guarantees, providing the necessary and sufficient conditions for observability, and thus convergence, in a constructive and physical intuitive manner. Several experimental examples of practical implementations were provided, illustrating the performance and consistency of the proposed sensor-based SLAM strategies, together with the ETM algorithm that also provides the vehicle trajectory and map in Earth-fixed coordinates.

Acknowledgements 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 and B. Guerreiro were supported respectively by the Ph.D. Student Grant SFRH/BD/89337/2012 and by the Post-doc Grant SFRH/BPD/110416/2015 from FCT.

References

1. Ahmad, A., Huang, S., Wang, J.J., Dissanayake, G.: A new state vector for range-only SLAM. In: Proceedings of the 2011 Chinese Control and Decision Conference (CCDC), pp. 3404–3409 (2011)
2. Bacca, B., Salvi, J., Cufi, X.: Long-term mapping and localization using feature stability histograms. *Robot. Auton. Syst.* **61**(12), 1539–1558 (2013)
3. Bailey, T.: Mobile Robot Localisation and Mapping in Extensive Outdoor Environments. Ph.D. thesis, University of Sydney, Australian Center of Field Robotics (2002)
4. Bailey, T.: Constrained initialisation for bearing-only SLAM. In: Proceedings of the 2003 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1966–1971. IEEE (2003)
5. Batista, P., Silvestre, C., Oliveira, P.: Single range aided navigation and source localization: observability and filter design. *Syst. Control Lett.* **60**(8), 665–673 (2011)
6. Batista, P., Silvestre, C., Oliveira, P.: Globally exponentially stable filters for source localization and navigation aided by direction measurements. *Syst. Control Lett.* **62**(11), 1065–1072 (2013)
7. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **110**(3), 346–359 (2008)
8. Bishop, A.N., Jensfelt, P.: A stochastically stable solution to the problem of robocentric mapping. In: IEEE International Conference on Robotics and Automation, 2009. ICRA '09, pp. 1615–1622 (2009)
9. Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J.J.: Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. *IEEE Trans. Robot.* **32**(6), 1309–1332 (2016)

10. Castellanos, J.A., Martinez-Cantin, R., Tardós, J.D., Neira, J.: Robocentric map joining: improving the consistency of EKF-SLAM. *Robot. Auton. Syst.* **55**(1), 21–29 (2007)
11. Csorba, M., Durrant-Whyte, H.F.: A new approach to simultaneous localisation and map building. In: *SPIE Aerosense* (1996)
12. Davison, A.J., Reid, I.D., Molton, N.D., Stasse, O.: MonoSLAM: real-time single camera SLAM. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(6), 1052–1067 (2007)
13. Dissanayake, G., Newman, P., Durrant-Whyte, H.F., Clark, S., Csobra, M.: A solution to the simultaneous localisation and mapping (SLAM) problem. *IEEE Trans. Robot. Autom.* **17**(3), 229–241 (2001)
14. Djugash, J., Singh, S.: Motion-aided network SLAM with range. *Int. J. Robot. Res.* **31**(5), 604–625 (2012)
15. Dubbelman, G., Browning, B.: COP-SLAM: closed-form online pose-chain optimization for visual SLAM. *IEEE Trans. Robot.* **31**(5), 1194–1213 (2015)
16. Durrant-Whyte, H.F.: Uncertain geometry in robotics. *IEEE J. Robot. Autom.* **4**(1), 23–31 (1988)
17. Gelb, A.: *Applied Optimal Estimation*. MIT Press, Cambridge (1974)
18. Guerreiro, B.J., Batista, P., Silvestre, C., Oliveira, P.: Globally asymptotically stable sensor-based simultaneous localization and mapping. *IEEE Trans. Robot.* **29**(6), 1380–1395 (2013)
19. Huang, G., Mourikis, A.I., Roumeliotis, S.I.: Observability-based rules for designing consistent EKF SLAM estimators. *Int. J. Robot. Res.* **29**(5), 502–528 (2010)
20. Huang, S., Dissanayake, G.: Convergence and consistency analysis for extended Kalman filter based SLAM. *IEEE Trans. Robot.* **23**(5), 1036–1049 (2007)
21. Jensfelt, P., Kragic, D., Folkesson, J., Bjorkman, M.: A framework for vision based bearing only 3D SLAM. In: *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pp. 1944–1950 (2006)
22. Julier, S.J., Uhlmann, J.K.: 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)*, Seoul, South Korea, vol. 4, pp. 4238–4243 (2001)
23. Kelly, J., Sukhatme, G.S.: Visual-inertial sensor fusion: localization, mapping and sensor-to-sensor self-calibration. *Int. J. Robot. Res.* **30**(1), 56–79 (2011)
24. Lemaire, T., Lacroix, S., Solà, J.: A practical 3D bearing-only SLAM algorithm. In: *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2449–2454 (2005)
25. Lourenço, P., Batista, P., Oliveira, P., Silvestre, C.: A globally exponentially stable filter for bearing-only simultaneous localization and mapping in 3-D. In: *Proceedings of the 2015 European Control Conference*, Linz, Austria, pp. 2817–2822 (2015)
26. Lourenço, P., Batista, P., Oliveira, P., Silvestre, C.: Simultaneous localization and mapping in sensor networks: a GES sensor-based filter with moving object tracking. In: *Proceedings of the 2015 European Control Conference*, Linz, Austria, pp. 2359–2364 (2015)
27. Lourenço, P., Batista, P., Oliveira, P., Silvestre, C., Philip Chen, C.L.: Sensor-based globally exponentially stable range-only simultaneous localization and mapping. *Robot. Auton. Syst.* **68**, 72–85 (2015)
28. Lourenço, P., Guerreiro, B.J., Batista, P., Oliveira, P., Silvestre, C.: 3-D inertial trajectory and map online estimation: building on a GAS sensor-based SLAM filter. In: *Proceedings of the 2013 European Control Conference*, Zurich, Switzerland, pp. 4214–4219 (2013)
29. Lourenço, P., Guerreiro, B.J., Batista, P., Oliveira, P., Silvestre, C.: Simultaneous localization and mapping for aerial vehicles: a 3-D sensor-based GAS filter. *Auton. Robot.* **40**, 881–902 (2016)
30. Lourenço, P., Guerreiro, B.J., Batista, P., Oliveira, P., Silvestre, C.: Uncertainty characterization of the orthogonal procrustes problem with arbitrary covariance matrices. *Pattern Recognit.* **61**, 210–220 (2017)
31. Montemerlo, M., Thrun, S., Koller, D., Wegbreit, B.: FastSLAM: a factored solution to the simultaneous localization and mapping problem. In: *Eighteenth National Conference on Artificial Intelligence*, pp. 593–598. American Association for Artificial Intelligence (2002)

32. Mur-Artal, R., Montiel, J.M.M., Tardós, J.D.: ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans. Robot.* **31**(5), 1147–1163 (2015)
33. Neira, J., Tardós, J.D.: Data association in stochastic mapping using the joint compatibility test. *IEEE Trans. Robot. Autom.* **17**(6), 890–897 (2001)
34. Smith, R.C., Cheeseman, P.: On the representation and estimation of spatial uncertainty. *Int. J. Robot. Res.* **5**(4), 56–68 (1986)
35. Smith, R.C., Self, M., Cheeseman, P.: Estimating uncertain spatial relationships in robotics. pp. 167–193. Springer, New York (1990)
36. Solà, J., Monin, A., Devy, M., Lemaire, T.: Undelayed initialization in bearing only SLAM. In: *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2499–2504 (2005)
37. Thrun, S., Liu, Y., Koller, D., Ng, A., Durrant-Whyte, H.: Simultaneous localization and mapping with sparse extended information filters. *Int. J. Robot. Res.* **23**(7–8), 693–716 (2004)
38. Weiss, L., Sanderson, A.C., Neuman, C.P.: Dynamic sensor-based control of robots with visual feedback. *IEEE J. Robot. Autom.* **3**(5), 404–417 (1987)