J Intell Robot Syst
DOI 10.1007/s10846-015-0181-7

Design and Validation of an RGB-D Based Localization System - Integration in a Docking System

João Barbosa · Carlos Cardeira · Paulo Oliveira · Pedro Batista · Carlos Silvestre

Received: 21 July 2014 / Accepted: 17 November 2014

Abstract This paper proposes a localization system for a mobile robot based on odometric data and RGB-D (Red, Green, Blue image and Depth map) measurements relative to a landmark, available from sensors installed on board. The localization system is composed of two cascaded estimators: i) a kinematic optimal attitude estimator; and ii) a position estimator designed in body-frame, based on an underlying LPV (Linear Parameter Varying) model, that avoids the need of approximations or linearization. Both underlying models are observable, even considering the presence of angular and linear slippages and the resulting estimators present globally asymptotically stable estimation error dynamics. Experiments to assess the performance of the proposed estimators were carried out resorting to a wheeled differential drive mobile robot in a laboratory instrumented with a Qualysis Motion Tracking System, used for ground-truth validation. An effective real-time localization system is obtained, featuring convergence to zero of the estimated errors, regardless the initial estimate and without requiring the landmark to be always visible, thus validating the system global stability. The results obtained paved the way to the integration of the proposed localization solution in a docking system for the same robot. The docking problem is solved with a smooth, time-invariant, globally asymptotically stable feedback control law, which allows for a very human-like closed-loop steering that drives the robot to a certain goal with a desired attitude and tunable curvature. Simulation and experimental results with the aforementioned robot are also presented, that illustrate the performance of the docking solution based on the proposed localization methods central to this work.

Keywords Mobile robots · Robot kinematics · Lyapunov methods · Nonlinear control systems

This work was partially supported by the program COMPETE/QREN/FEDER under PRODUTECH-PTI (P. 3904) and FCT, through IDMEC, under LAETA Pest-OE/EME/LA0022 and ISR/LARSyS PEst-OE/EEI/LA0009.

J. Barbosa · C. Cardeira (✉) · P. Oliveira
IDMEC/LAETA-Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
e-mail: carlos.cardeira@tecnico.ulisboa.pt

J. Barbosa
e-mail: joao.vitor.barbosa@tecnico.ulisboa.pt

P. Oliveira · P. Batista · C. Silvestre
ISR/LARSyS-Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
e-mail: pjcro@isr.tecnico.ulisboa.pt

P. Batista
e-mail: pbatista@isr.tecnico.ulisboa.pt

C. Silvestre
e-mail: cjs@isr.ist.utl.pt

C. Silvestre
Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, China

Published online: 17 January 2015

1 Introduction

The localization of mobile robots is one important challenge to the scientific community. The robots have to be able to use the sensors on-board, which often consist of optical encoders, mono and stereo cameras, gyroscopes, accelerometers, laser range-finders, among many others [1]. The localization in the environment consists of knowing its position in some global frame or in any local frame of interest. This localization is always needed if the robot is to autonomously plan its motions that go towards the satisfaction of a certain goal. The specific task to be carried out by the mobile robot motivates the choice of the kind of localization needed, ranging from a topological localization [2, 3], often aided by a structured map or other representation of the environment, to a scenario where the robot may have to build its own map of the surroundings while simultaneously localizing itself in it. The latter problem is widely known as SLAM [4, 5] and the solutions are usually based in particle filters or extended Kalman Filters.

The strategy proposed in this paper consists of a sensor-based localization system that estimates the position of a certain feature or landmark in the robot frame, based on measurements from an RGB-D (red, green, blue and depth) camera. Thus an intuitive interaction of the robot with its close surroundings is possible. The purpose of this method is twofold: i) to tackle some problems present in a number of localization solutions, by reducing the consistency and accuracy issues caused by approximations or linearizations, as is the case of the EKF based localization systems [6, 7]; and ii) to complement their functionality, allowing for a modularized approach to a specific mission. This estimator takes advantage of the independence between the attitude and position of a feature relative to the body-fixed frame, thus preventing the attitude errors from augmenting given the position estimation errors. The system kinematics is expressed as an LPV (Linear Parameter Varying) system, thus allowing the design of an estimator with global asymptotically stable error dynamics. This characteristic is due to the fact that the exact kinematics does not need to be linearized for the estimator design process. The models used, augmented by constant or slowly time-varying angular and linear slippages, respectively, are shown to be observable, following the approach introduced in [8, 9].

Localization systems are key enabling elements for the development of docking systems. Industrial automation has experienced huge advances in the last decades [10]. Flexible automation manufacturing cells require the use of automatic handling solutions usually resorting to Automatic Guided Vehicles (AGVs). Nowadays, fleets of AGVs must navigate among warehouses, automated workcells, and charging stations. Thus, the automated docking of mobile robotic platforms, with minimal structuring of the environment is still an active topic of research [11]. The solutions found in the literature to solve this problem are diverse, both in terms of algorithms and sensors. One approach, defined as visual servoing, with an early contribution reported in [12], consists of representing a given task directly by an error relative to a goal image to be captured by the vision system. This approach became popular from 1990 onward, with works such as [13], and the contribute of task function approach [14]. Visual servoing benefits from contributes with out-of-body cameras, i.e., Camera-Space Manipulation (CSM) [15], or Mobile Camera-Space Manipulation (MCSM) which extends the latter with body embedded cameras. More recently, [16] computes the goal configuration using visual landmarks. Other approaches to the docking problem include the computation of feedback control laws by using Lyapunov and backstepping techniques. That line of research lead to the development of an Ultra-Short Baseline (USBL) acoustic positioning system [17] applied on an underwater solution for a similar problem as the one that will be illustrated in this work. Electromagnetic based homing systems [18], optical guidance approaches such as [19], and computing the deceleration needed by a robot, resorting to an estimation of a *time-to-contact* (τ) through optical flow field divergence measurements of an image stream as in [20], are some other alternatives that have been exploited by the community. In [21] a method based on the direction of arrival (DOA) of signals transmitted by RFID transponders is proposed, showing that a robot can dock in a station transmitting through an RFID by using two antennas installed on board of the vehicle. A method proposing the estimation of the position and orientation of a visual landmark is presented in [22] to help on docking and automatic recharging, thus being similar to the work presented herein. Recently, in [23] an indoor RGB-D and WiFi localization system is proposed resorting to a Markov

Localization technique, with an one meter standard deviation error. The solution detailed in this work is based on a preliminary version presented in [24] and clearly present a reduced error relative to this last contribution.

The main contributions of the method proposed in this paper are: i) to tackle some problems present in a number of localization solutions, by reducing the consistency and accuracy issues caused by approximations or linearizations, as is the case of the EKF based localization systems [6, 7]; and ii) to complement their functionality, allowing for a modularized approach to a specific mission. Moreover, this estimator takes advantage of the independence between the attitude and position of a feature relative to the body-fixed frame, thus preventing the attitude errors from augmenting given the position estimation errors.

This work is structured as follows: the architecture is briefly outlined in Section 2, where all the modules that compose the localization system are shown, followed by the robot environment formalization in Section 3. Then, both attitude and position estimator modules are presented, respectively, in Section 4 and Section 5. Finally the experimental results are analyzed in Section 6, where a Qualisys™ Visual Tracking System [25] is used for ground-truth. Results of the proposed localization solution integrated in a docking system are also reported.

2 Architecture

The proposed landmark-based on-board localization system, depicted in Fig. 1, is composed of three modules: i) the landmark detector module that consists of the algorithm that will process the RGB and depth images, in order to obtain measurements of the landmark position and orientation in the robot frame; ii) a sub-optimal position estimator based on a Kalman Filter; iii) an optimal attitude estimator, also based on a Kalman filter.

All the modules of the proposed architecture, depicted in Fig. 1, rely on the sensor package composed by: i) an RGB-D camera, that will provide visual data that the landmark detector uses to compute the relative position $\bar{\mathbf{e}}$ and relative attitude $\overline{\psi}$; and ii) the left and right velocities of the wheels, used as measurements to compute the linear and angular velocities, $\overline{v}$ and $\overline{\omega}$, respectively. The RGB-D camera used is a Microsoft Kinect, which includes an 8 bit RGB image with the VGA resolution (640 × 480 pixels), and a 2D depth sensor with the same resolution, with 11 bits of resolution. The use of this sensor for mobile robots localization could combine the capture of an RGB image and a depth map about the environment, obtaining RGB-D images. Optical encoders attached to the wheels provide angular and linear velocity readings. The velocity readings can be derived from the encoder readings directly or from the velocity commands stored in the controller, assuming perfect tracking.

3 Model Description

The scenario of operation under study in this work is composed of a mobile robot and a static landmark, as depicted in Fig. 2, where the frame $\{I\}$ is an inertial frame that can be considered to be stationary for the purposes of this study. Frame $\{B\}$ is attached to

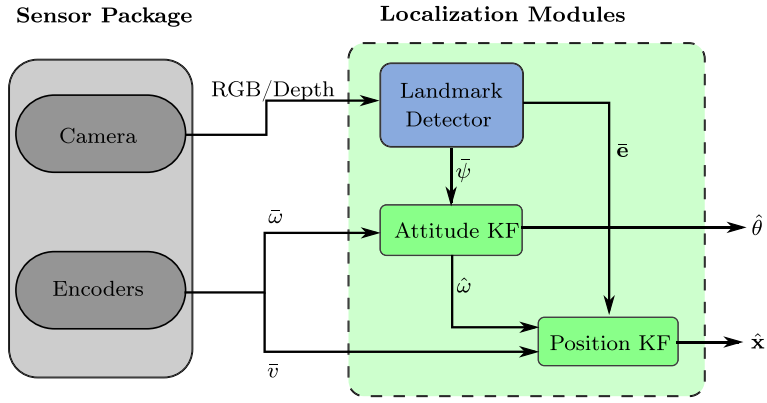


Fig. 1 Estimator Modules

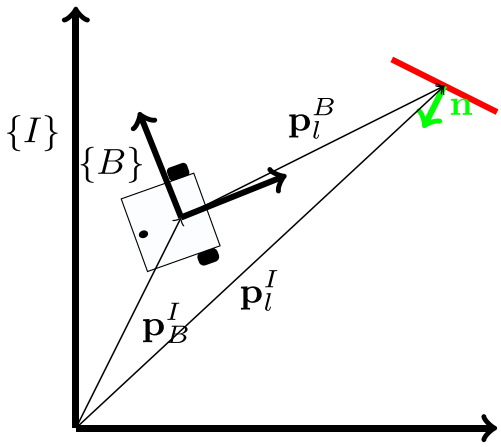


Fig. 2 Schematic of Inertial and Body-Fixed frames

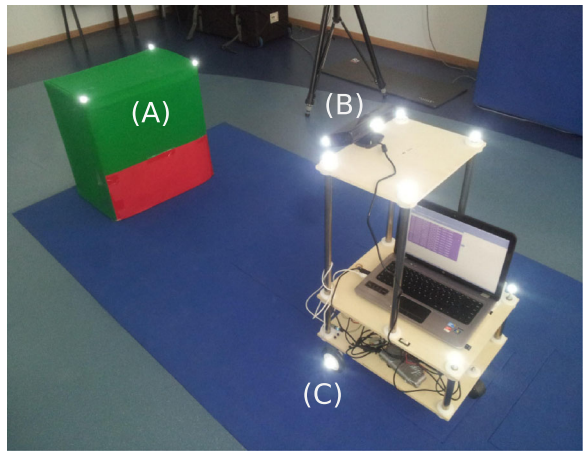


Fig. 3 IST Biomechanics Laboratory. **A** - Landmark, **B** - RGB-D Camera, and **C** - Robot

the mobile robot and is hence designated a body-fixed frame.

In order to transform a position written in $\{B\}$ into one expressed in $\{I\}$, a transformation needs to be executed. The translation is defined by the body-fixed frame position in the inertial frame $\mathbf{p}_B^I$, and so the landmark position in both frames follows

$$\mathbf{p}_B^I(t) + {}^I\mathbf{p}_l^B(t) = \mathbf{p}_l^I(t), \tag{1}$$

where ${}^I\mathbf{p}_l^B(t) \in \mathbb{R}^2$ is the landmark position in $\{B\}$ expressed in $\{I\}$ and $\mathbf{p}_l^I(t)$ is the landmark position in $\{I\}$, expressed in the latter. The landmark position in the body fixed frame $\mathbf{p}_l^B(t) \in \mathbb{R}^2$ is what needs to be derived and will henceforth, for a matter of simplicity, be denoted as $\mathbf{e}(t)$.

The rotation matrix from $\{B\}$ to $\{I\}$ that simplifies (1) is denoted by ${}^I\mathbf{R}_B(t) \in SO(2)$ and respects the kinematics

$${}^I\dot{\mathbf{R}}_B(t) = \mathbf{S}(\omega(t))\mathbf{R}(t), \tag{2}$$

where

$$\mathbf{S}(\omega(t)) = \begin{bmatrix} 0 & -\omega(t) \\ \omega(t) & 0 \end{bmatrix},$$

is a skew symmetric matrix. The angular velocity of the body-fixed frame is given by $\omega(t) \in \mathbb{R}$, that can be obtained from the differential readings of the encoders. The rotation ${}^I\mathbf{R}_B(t)$ will henceforth be denoted as $\mathbf{R}(t)$ for simplicity. It is straightforward to show that the inverse rotation follows a similar expression to Eq. 2 by taking the derivative on both sides of $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ and making the necessary substitutions

$$\dot{\mathbf{R}}^T(t) = -\mathbf{S}(\omega(t))\mathbf{R}^T(t).$$

In this section all quantities were defined in 2D but the generalization to 3D is straightforward.

4 Optimal Attitude and Angular Slippage Estimation

This section will focus on deriving the attitude estimator. Firstly, the kinematic model will be described, followed by a brief observability study that will naturally lead to the definition of the estimator. The proposed kinematic system is appended explicitly by the unavoidable angular slippage, that may occur due to the lack of knowledge on the contact points with the floor. Lack of accuracy in the measurement of each wheel radius, aging, or asymmetries

Table 1 Qualysis Motion Tracking system characteristics

Cameras	14 Qualisys Pro Reflex 1000
Frequency	100 Hz
Markers	19 mm diam. passive retroreflectors
Accuracy	<1 mm after calibration

in mechanical construction, can also be explained by the model adopted. Here the angular slippage $s(t)$ is considered to be constant or slowly time-varying, i.e. $\dot{s} = 0$. The model that describes the attitude system is given by the following dynamic and the output equations

$$\dot{\boldsymbol{\theta}}(t) = \mathbf{A}^{\boldsymbol{\theta}}\boldsymbol{\theta}(t) + \mathbf{B}^{\boldsymbol{\theta}}\omega(t) + \boldsymbol{\nu}(t), \tag{3}$$

and

$$y(t) = \mathbf{C}^{\boldsymbol{\theta}}\boldsymbol{\theta}(t) + \eta(t), \tag{4}$$

respectively, where $\boldsymbol{\theta}(t) = [\psi(t)\ s(t)]^T$, $\mathbf{A}^{\boldsymbol{\theta}} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{B}^{\boldsymbol{\theta}} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\mathbf{C}^{\boldsymbol{\theta}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$, and $\boldsymbol{\nu}(t)$ and $\eta(t)$ are the plant and output noises, respectively. The superscript θ is added to denote the dynamics, input, and output matrices, respectively, of the attitude state space model. Both disturbances are assumed to have Gaussian distributions with zero mean and with covariances $\mathbf{Q}^{\boldsymbol{\theta}}$ and $\mathbf{R}^{\boldsymbol{\theta}}$, respectively, i.e.

$$\boldsymbol{\nu}(t) \sim N(\mathbf{0}, \mathbf{Q}^{\boldsymbol{\theta}})$$
$$\eta(t) \sim N(\mathbf{0}, \mathbf{R}^{\boldsymbol{\theta}}).$$

This model considers the landmark as if it is moving in the body reference frame. Thus, $\psi(t)$ represents the landmark attitude expressed in body frame, which as stated before, is under the assumption that it is possible to define a unique reference system in the landmark used. Moreover, it is assumed that the camera is able to detect the relative orientation. Since the continuous time system is LTI (Linear Time Invariant), as expressed by Eq. 3 and 4, it is sufficient to check if the observability matrix $\mathcal{O}_a$ is full rank,

$$\mathcal{O}_a = \begin{bmatrix} \mathbf{C}^{\boldsymbol{\theta}} \\ \mathbf{C}^{\boldsymbol{\theta}}\mathbf{A}^{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The solution that we will develop along this work will be implemented in discrete time. Thus, assuming that the attitude filter runs fast enough compared to the robot closed loop dynamics bandwidth we can consider constant angular velocity between sampling instants. The state transition equation for this linear time invariant system is

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\Phi}_k^{\boldsymbol{\theta}}\boldsymbol{\theta}_k + \mathbf{G}_k^{\boldsymbol{\theta}}\omega_k + \boldsymbol{\nu}_k,$$

where ω_k is the measured angular velocity obtained using the command sent to the dual-motor driver, $\boldsymbol{\Phi}_k^{\boldsymbol{\theta}} = \exp(A^{\boldsymbol{\theta}}T_k)$, $\mathbf{G}_k^{\boldsymbol{\theta}} = \omega_k \int_{t_{k-1}}^{t_k} \boldsymbol{\Phi}^{\boldsymbol{\theta}}(t_k - \tau)\left(\mathbf{B}^{\boldsymbol{\theta}}\right) d\tau$, and T_k is the time interval between samples $k-1$ and k. From the implementation point of view this quantity can be measured, which is not constant in general.

The attitude kinematics and the outputs are completely defined, so defining the state to be estimated

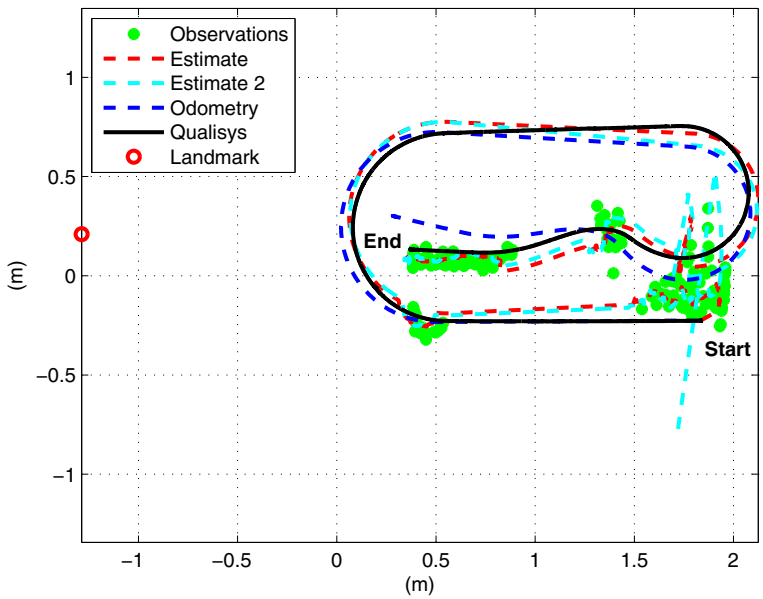


Fig. 4 Trajectory ground truth and estimate results

as $\hat{\boldsymbol{\theta}}(t) = \left[\hat{\psi}(t) \;\; s(\hat{t})\right]^T$, it is now possible to compute the dynamics of the optimal estimator, making use of the celebrated Kalman filter [26]. The state estimate dynamics and the error covariance are updated according respectively by

$$\hat{\boldsymbol{\theta}}_{k+1} = \boldsymbol{\Phi}_k^{\boldsymbol{\theta}} \hat{\boldsymbol{\theta}}_k + \mathbf{G}_k^{\boldsymbol{\theta}} \omega_k + \mathbf{K}_k^{\boldsymbol{\theta}} \left[\bar{\psi}_k - \hat{\psi}_k\right],$$

$$\mathbf{P}_{k|k}^{\boldsymbol{\theta}} = \left(\mathbf{I} - \mathbf{K}_k^{\boldsymbol{\theta}} \mathbf{C}^{\boldsymbol{\theta}}\right) \mathbf{P}_{k|k-1}^{\boldsymbol{\theta}} \left(\mathbf{I} - \mathbf{K}_k^{\boldsymbol{\theta}} \mathbf{C}^{\boldsymbol{\theta}}\right) + \mathbf{K}_k^{\boldsymbol{\theta}} \mathbf{R}_k^{\boldsymbol{\theta}} \mathbf{K}_k^{\boldsymbol{\theta}}.$$

where $\mathbf{K}_k^{\boldsymbol{\theta}}$ is the Kalman gain, at time instant kT, computed as

$$\mathbf{K}_k^{\boldsymbol{\theta}} = \mathbf{P}_{k|k}^{\boldsymbol{\theta}} \mathbf{C}^{\boldsymbol{\theta}} \left[\mathbf{C}^{\boldsymbol{\theta}} \mathbf{P}_{k|k}^{\boldsymbol{\theta}} \mathbf{C}^{\boldsymbol{\theta}}\right]^{-1}.$$

The error covariance prediction can be obtained resorting to

$$\mathbf{P}_{k|k-1}^{\boldsymbol{\theta}} = \boldsymbol{\Phi}(k)^{\boldsymbol{\theta}} \mathbf{P}_{k-1|k-1}^{\boldsymbol{\theta}} \boldsymbol{\Phi}^{\boldsymbol{\theta}}(k)^T + \mathbf{Q}_k^{\boldsymbol{\theta}}. \tag{5}$$

5 Sub-Optimal Position and Linear Slippage Estimation

In order to have a localization system working in the mobile robot, thus expressed in $\{B\}$, as introduced in Section 3, we need to compute the error dynamics of the position of the landmark in body frame $\dot{\mathbf{e}}(t)$. After having completely defined the model of the

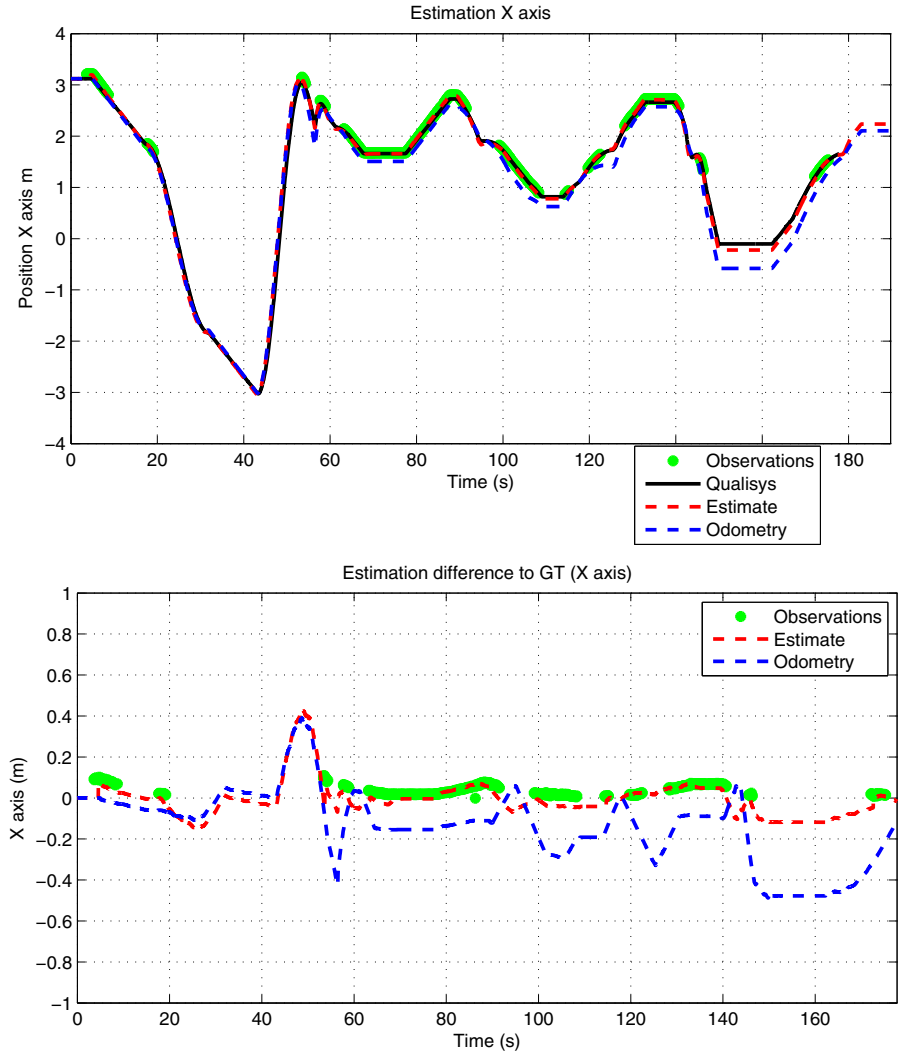


Fig. 5 Trajectory ground truth and estimate results in X axis on top; Estimation errors at the bottom

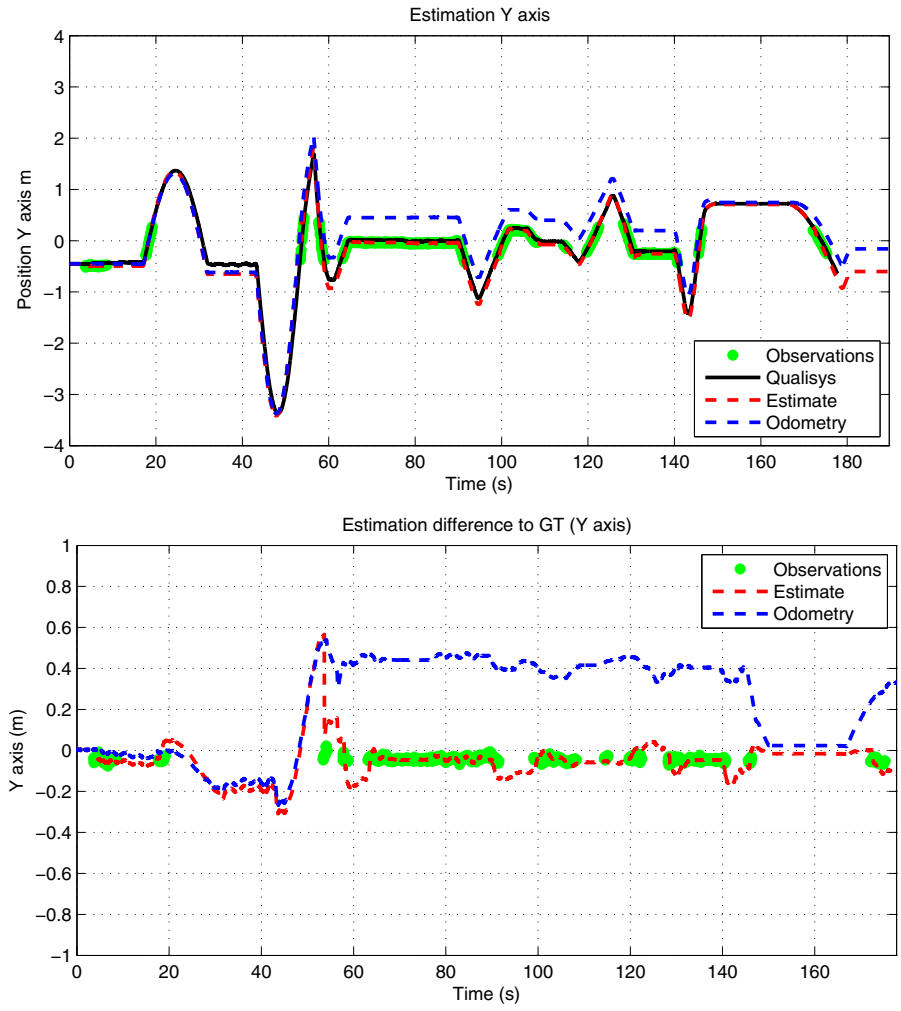


Fig. 6 Trajectory ground truth and estimate results in Y axis on top; Estimation errors at the bottom

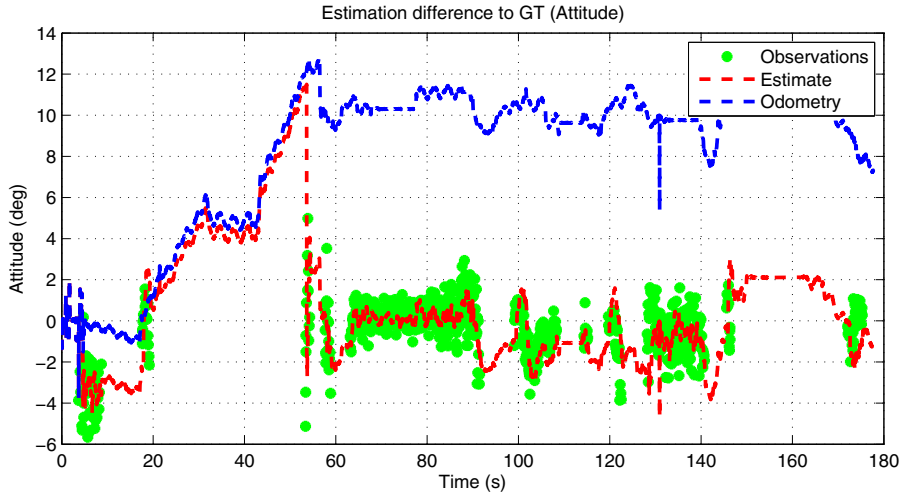


Fig. 7 Attitude estimate error, relative to ground-truth information

robot environment in Section 3, we start by expressing the dynamics of the robot's position in $\{I\}$ as

$$\dot{\mathbf{p}}(t) = {}^{I}\mathbf{R}_{B}(t)\mathbf{u}(t),$$

where $\mathbf{u}(t) = [v(t)\ \ 0]$ and $v(t) \in \mathbb{R}$ is the robot velocity in the body-fixed frame. By expressing the product of Eq. 1 by $\mathbf{R}^T$ we get the position error $\mathbf{e}(t)$ expressed in terms of $\mathbf{p}_l$ and $\mathbf{p}$ each corresponding to the landmark position and $\{B\}$ position in $\{I\}$ (respectively $\mathbf{p}_l^I$ and $\mathbf{p}_B^I$).

$$\mathbf{e}(t) = \mathbf{R}(t)^T(\mathbf{p}_l(t) - \mathbf{p}(t)), \tag{6}$$

which, once the time derivative is taken gives

$$\dot{\mathbf{e}} = \dot{\mathbf{R}}^T(t)(\mathbf{p_l}(t) - \mathbf{p}(t)) + \mathbf{R}^T(t)\,(\dot{\mathbf{p}}_l(t) - \dot{\mathbf{p}}(t))\,. \tag{7}$$

Considering that the landmark will be static in the inertial reference system, the term $\dot{\mathbf{p}}_l(t)$ drops. Using Eq. 6 in Eq. 7 we get

$$\dot{\mathbf{e}}(t) = -\mathbf{S}(\omega)\mathbf{e}(t) - \mathbf{u}(t). \tag{8}$$

It can be further assumed that the common mode velocity $v(t)$ can suffer from a biased measurement due to slippage. The velocity could then be expressed as $v(t) = \bar{v}(t)+b$ where b is a constant or slowly time-varying bias and $\bar{v}(t)$ is the measured linear velocity, while $v(t)$ is the actual linear velocity. Thus, the slippage is only considered along the longitudinal axis of the robot. An extension to consider transversal slippage would be straightforward. Assuming the state vector as $\mathbf{x}(t) = [\mathbf{e}^T(t)\ b(t)]^T$, the matrix expression for the dynamics of the position will be given by

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(\omega(t))\mathbf{x}(t) + \mathbf{B}\bar{v}(t) + \mathbf{v}(t)\\ &= \begin{bmatrix} 0 & \omega & -1 \\ -\omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}\bar{v}(t) + \mathbf{v}(t),\end{aligned}$$

where $\mathbf{v}(t) \in \mathbb{R}^3$ is Gaussian white plant noise caused by model uncertainty and is characterized by

$$\begin{aligned}E[\mathbf{v}(t)] &= 0, \quad \forall t \in \mathbb{R}\\ E[\mathbf{v}(t)\mathbf{v}^T(\tau)] &= \mathbf{Q}^{\mathbf{x}}\delta(t-\tau).\end{aligned}$$

$\mathbf{A}$ and $\mathbf{B}$ denote the dynamics and input matrices, respectively of the position state space model. Since the camera provides measurements of the landmark localization in $\{B\}$, the output equation of this underlying system can be expressed as

$$\mathbf{y}(t) = \mathbf{e}(t) + \mathbf{w}(t),$$

where $\mathbf{w}(t)$ represents the noise generated by the camera sensor as well as the detection algorithm,

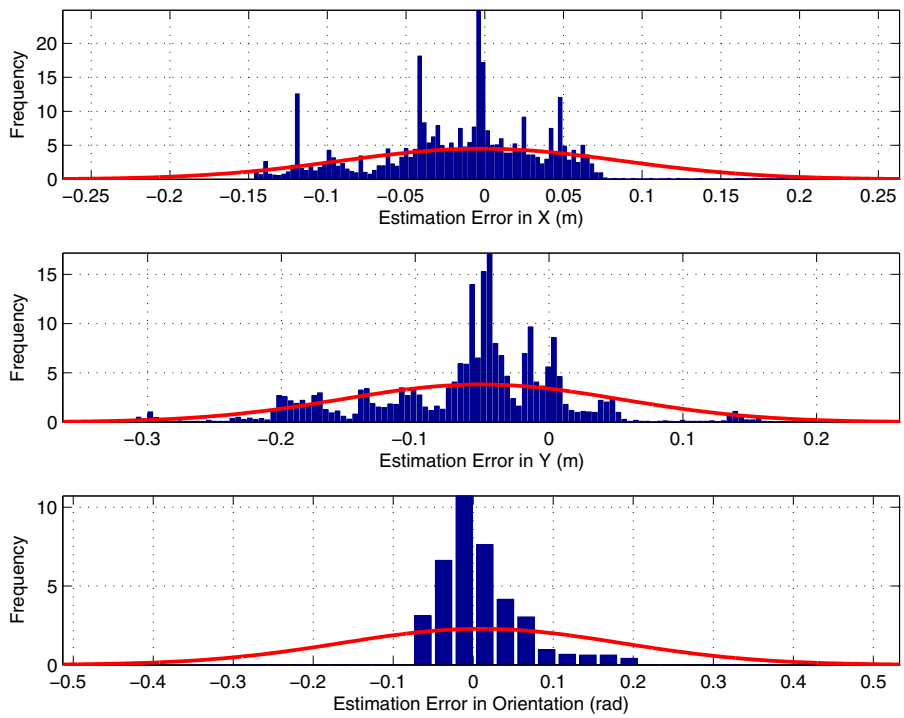


Fig. 8 Histograms of the estimation errors, relative to ground truth

assumed to have properties similar to those of the plant noise

$$E[\mathbf{w}(t)] = 0, \quad \forall t \in \mathbb{R}$$
$$E\left[\mathbf{w}(t)\mathbf{w}^T(\tau)\right] = \mathbf{R}^{\mathbf{x}}\delta(t-\tau), \quad \forall t, \tau \in \mathbb{R}.$$

Also, both the plant and the sensor noise are assumed as uncorrelated, which can be expressed as

$$E[\mathbf{w}(\eta)\mathbf{v}(\tau)] = 0, \quad \forall \eta, \tau \in \mathbb{R}.$$

In order to be possible to estimate the state of the system just introduced, it needs to be observable. This analysis uses the observation matrix $\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, where slippage is not explicitly present. The transition matrix can be computed as $\mathbf{\Phi}(t_1, t_0) = \exp\left(\int_{t_0}^{t_1} \mathbf{A}(\omega(\tau))d\tau\right)$, where $t_1 > t_0$. As the dynamics were written as an LPV system, observability can be studied resorting to the tools proposed in [9]. In general, the observability Gramian $\mathbf{W_O}(t_1, t_0)$ will be full rank rendering the system observable. Even if the linear movement case is considered, i.e. $\omega(t) = 0$, the observability Gramian degenerates in

$$\mathbf{W_O}(t_1, t_0)|_{\omega=0} = \begin{bmatrix} \Delta t & 0 & -\Delta t^2/2 \\ 0 & \Delta t & 0 \\ -\Delta t^2/2 & 0 & \Delta t^3/3 \end{bmatrix}$$

where $\Delta t = t_1 - t_0$ and full observability is preserved.

To obtain a discretized version of this system, we adopted an identical approach to that pursued in Section 4, leading to the following discrete time LPV system, with sampling period T

$$\mathbf{x}_{k+1} = \mathbf{\Phi}(\omega_k)\mathbf{x}_k + \mathbf{G}_k \upsilon_k + \mathbf{v}_k,$$

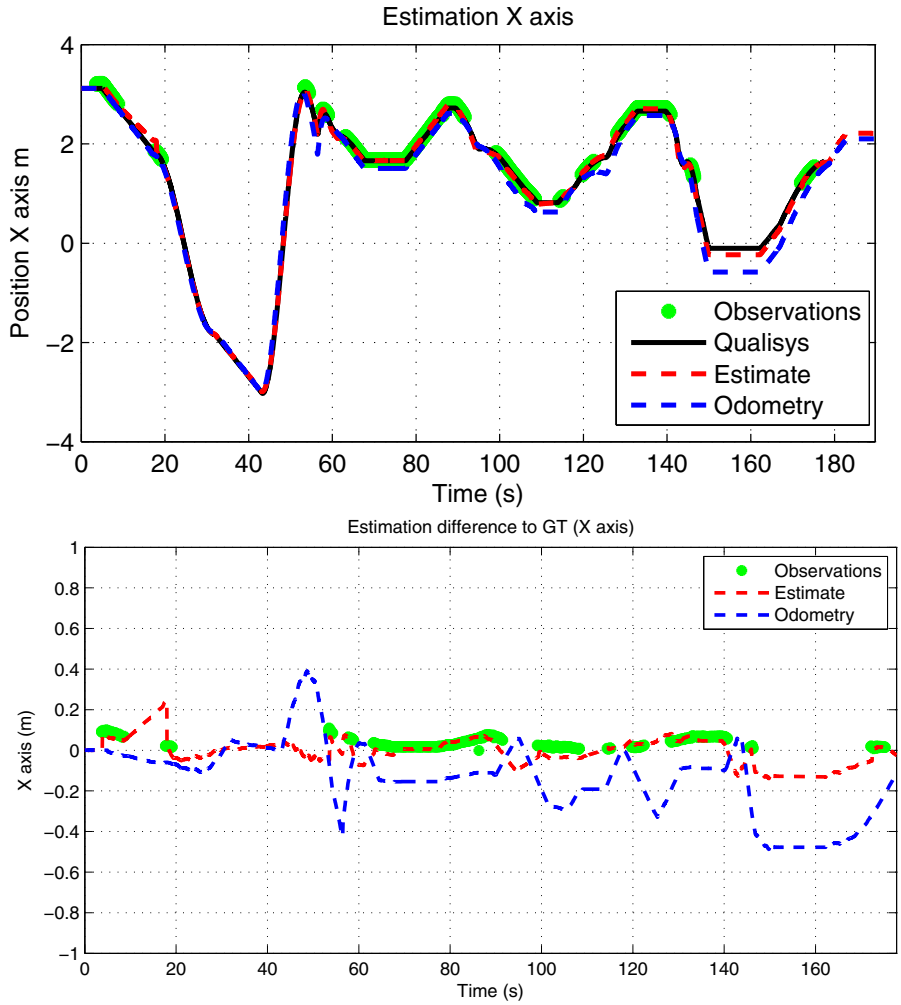


Fig. 9 Trajectory ground truth and estimate results in X axis on top; Estimation errors at the bottom

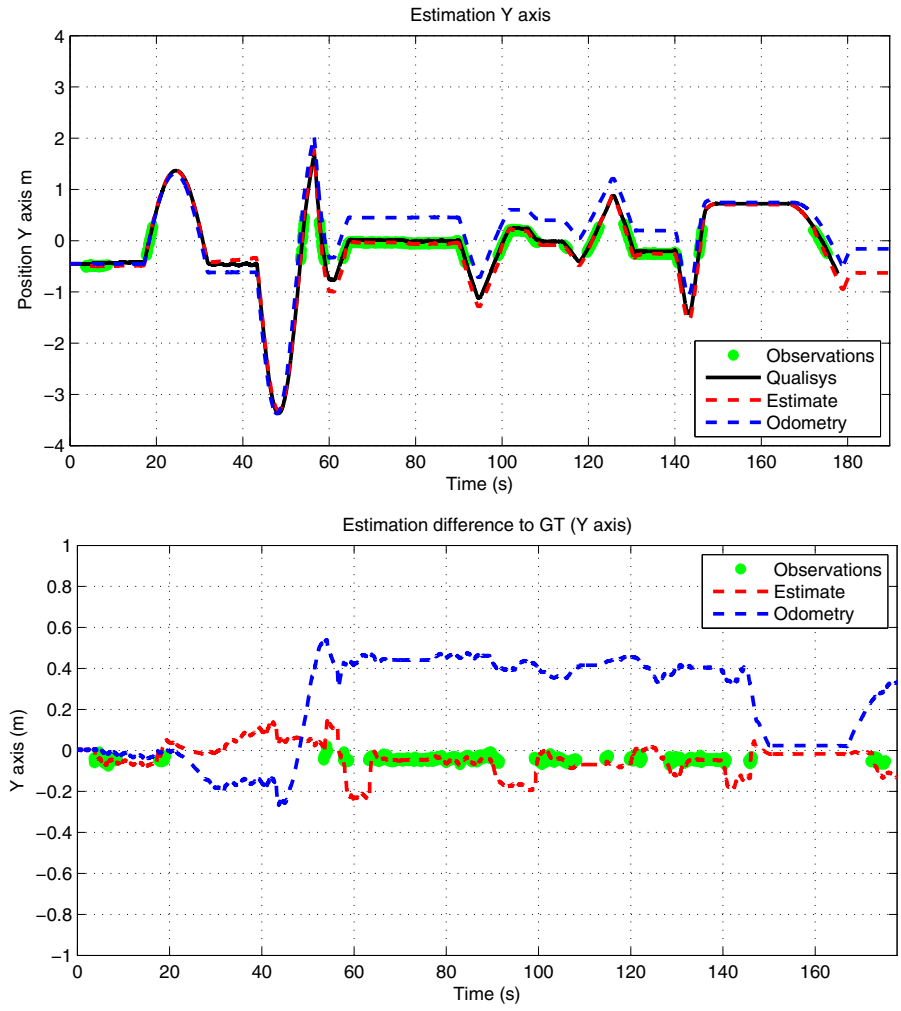


Fig. 10 Trajectory ground truth and estimate results in Y axis on top; Estimation errors at the bottom

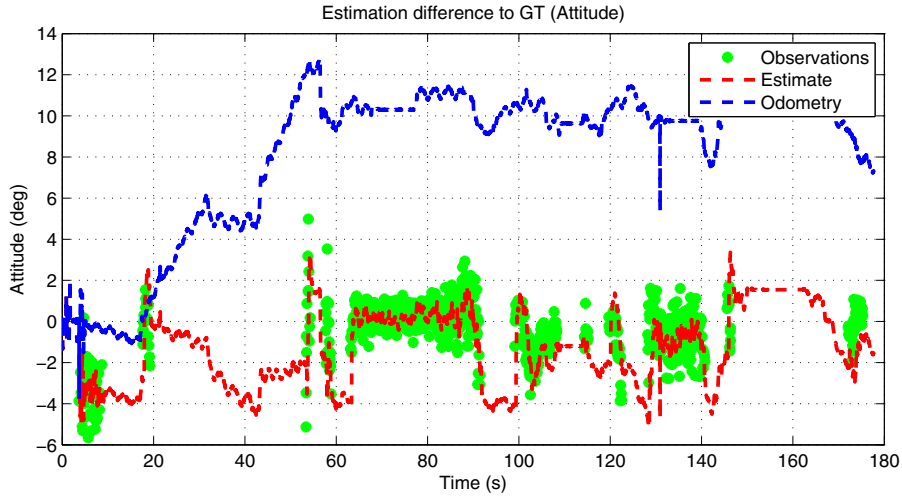


Fig. 11 Attitude Estimate error, relative to ground-truth information

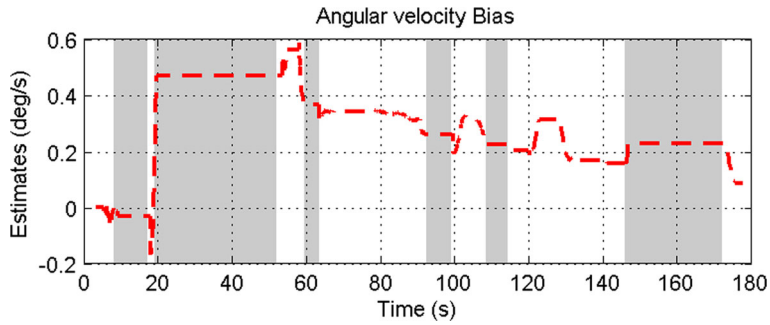


Fig. 12 Angular slippage estimation

where the transition matrix $\mathbf{\Phi}(\omega_k)$ is an abbreviavion for $\mathbf{\Phi}_{\omega_k}(t_k, t_{k-1})$, expressed by

$$\begin{aligned}\mathbf{\Phi}(\omega_k) &= \mathbf{\Phi}_{\omega_k}(t_k, t_{k-1}) \\ &= \exp\left(\int_{t_{k-1}}^{t_k} \mathbf{A}(\omega_\tau) d\tau\right) \\ &= \begin{bmatrix} c(\omega_k T_k) & s(\omega_k T_k) & -\frac{s(\omega_k T_k)}{\omega_k} \\ -s(\omega_k T_k) & c(\omega_k T_k) & \frac{1-c(\omega_k T_k)}{\omega_k} \\ 0 & 0 & 1 \end{bmatrix},\end{aligned}$$

where $s(.)$ and $c(.)$ are abbreviations for the sine and cosine functions, respectively, $T_k = t_k - t_{k-1}$ is the sampling interval, $\mathbf{v}_k$ is the discrete Gaussian white noise, and $\mathbf{G}_k$ is the discrete input matrix expressed as

$$\mathbf{G}_k = \int_{t_{k-1}}^{t_k} \mathbf{\Phi}_{\omega_k}(\tau, t_{k-1}) \mathbf{B} d\tau = \begin{bmatrix} -\frac{s(\omega_k T_k)}{\omega_k} \\ \frac{1-c(\omega_k T_k)}{\omega_k} \\ 0 \end{bmatrix}.$$

The position estimator will perform a sub-optimal estimation since the angular velocity that parametrizes the state transition matrix of this system is meant to take into account the angular slippage whose estimation is described in Section 4. Nevertheless, the equation that describes the estimate dynamics is similar to that used for the attitude system estimate, resulting in

$$\hat{\mathbf{x}}_{k+1} = \mathbf{\Phi}(\hat{\omega}_k)\hat{\mathbf{x}}_k + \mathbf{B}^{\mathbf{x}} v_k + \mathbf{K}_k^{\mathbf{x}} \left[\bar{\mathbf{e}}_k - \hat{\mathbf{e}}_k\right].$$

The Kalman gain for this estimator is calculated in the exact same way as in the attitude estimator, using Eq. 5, only using the appropriate sensor noise covariance matrix $\mathbf{R}^{\mathbf{x}}$ and model noise covariance $\mathbf{Q}^{\mathbf{x}}$.

6 Experimental Results

This section details the experimental setup and the analysis of the estimation errors of the proposed solutions that constitute the localization system.

The Qualisys ™Motion Tracking [25] system, using 14 different cameras to track the position of reflectors placed upon the mobile robot, provides measurements that are used as ground-truth data. The characteristics of the tracking system are listed in Table 1. The robot prototype and landmark setup are shown in Fig. 3. Several passive retroreflectors, which are highlighted by the camera flash, were placed on the robot and landmark to provide redundant data. The landmark

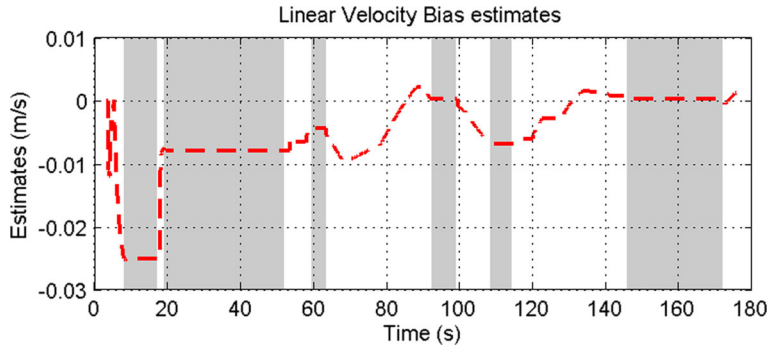


Fig. 13 Linear slippage estimation

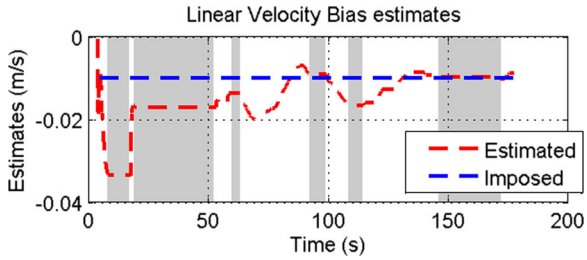


Fig. 14 Linear slippage estimation

used serves only the purpose of validation of the localization method proposed.

A summary of the parameters and initialization of both Kalman Filters is summarized next:

- Sampling period: $0.2s$ for both estimators.
- Camera noise covariance: $\mathbf{R}^{\mathbf{x}} = 1 \times 10^{-2}\mathbf{I}_2$ and $\mathbf{R}^{\theta} = 1 \times 10^{-2}$
- Plant noise covariance: $\mathbf{Q}^{\mathbf{x}} = diag(4.1 \times 10^{-6}\mathbf{I}_2, 1 \times 10^{-8})$ and $\mathbf{Q}^{\theta} = diag(2 \times 10^{-5}, 1 \times 10^{-8})$
- Initial covariance matrix: $\mathbf{P}_0^{\mathbf{x}} = 1\mathbf{I}_3$ and $\mathbf{P}_0^{\theta} = 0.1\mathbf{I}_3$
- Initial conditions: $\hat{\mathbf{e}}$ and $\hat{\theta}$ were set to the approximate position values, and both bias estimates $\hat{b}$ and $\hat{s}$ were set to zero.

The RGB-D camera installed on board is assumed to have a frame $\{C\}$ attached. To express the measurements from this sensor in the body frame, it is required to express the position and attitude of the frame $\{C\}$ relative to the body-fixed frame $\{B\}$. To that purpose, it was identified a translation given by $\mathbf{p}_C^B = [0.090 \quad 0.03 \quad 0.775]^T \, (m)$, and a rotation

$$ {}^{B}\mathbf{R}_C = \begin{bmatrix} c(\theta) & -s(\theta) & 0 \\ s(\theta) & c(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, $$

where $\theta = 0.216 \, rad$.

6.1 Experimental Validation of the Localization System

This section comprises the localization results for two separate trajectories performed in a laboratory equipped with the Qualisys Motion Tracking system. For an easier visualization, only a portion of the first trajectory that was tested is depicted in Fig. 4, where *Estimate* refers to a correct initialization and *Estimate 2* to a wrong one, for global stability validation. In the remaining figures of the document *Observations* correspond to the measurements of the RGB-D camera.

The comparison between ground truth data expressed in $\{B\}$ and the estimate is depicted in

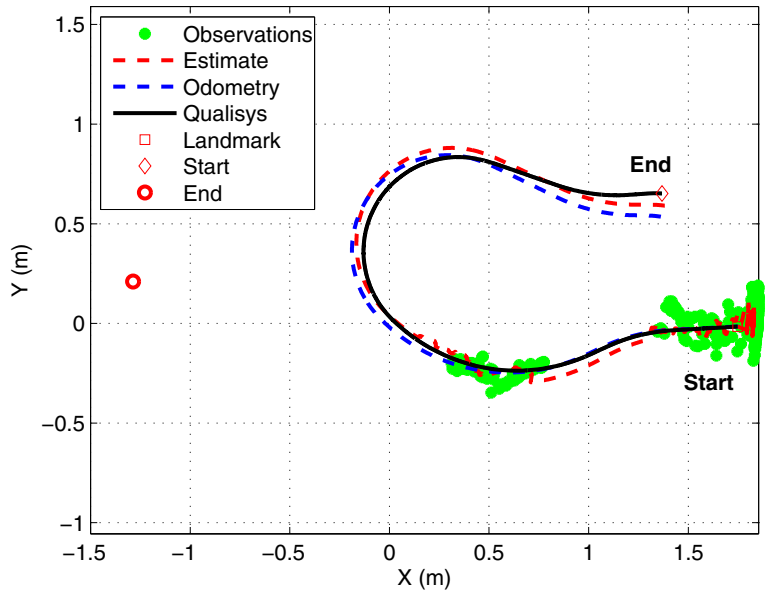


Fig. 15 Trajectory without slippage estimation, at constant linear velocity

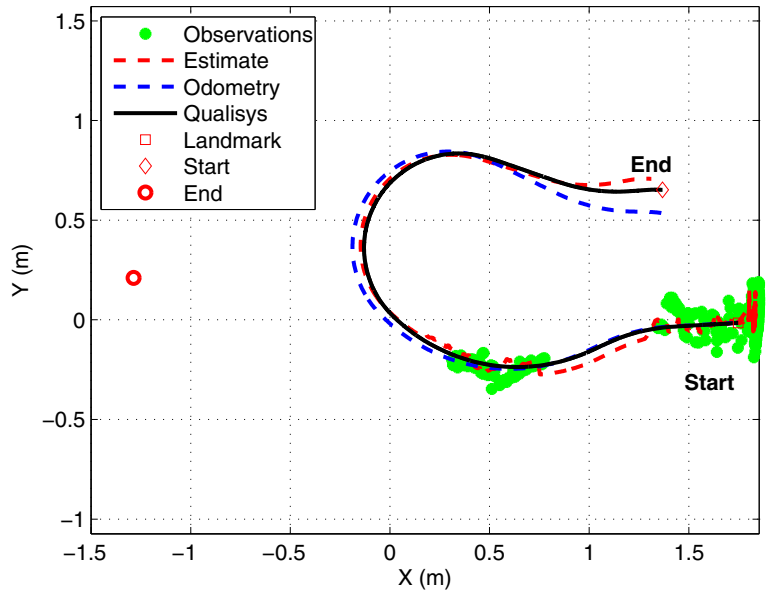


Fig. 16 Trajectory with slippage estimation, at constant linear velocity

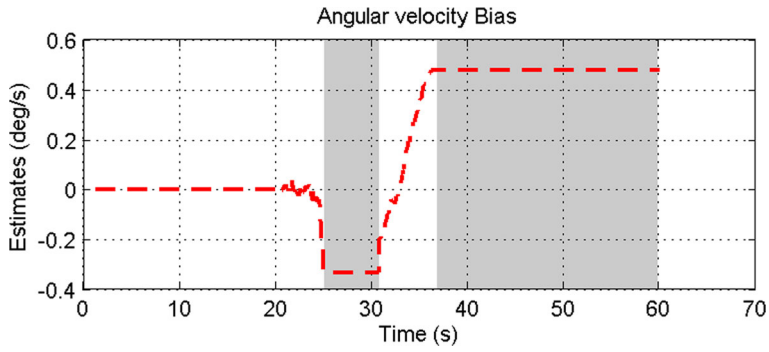


Fig. 17 Angular slippage estimation

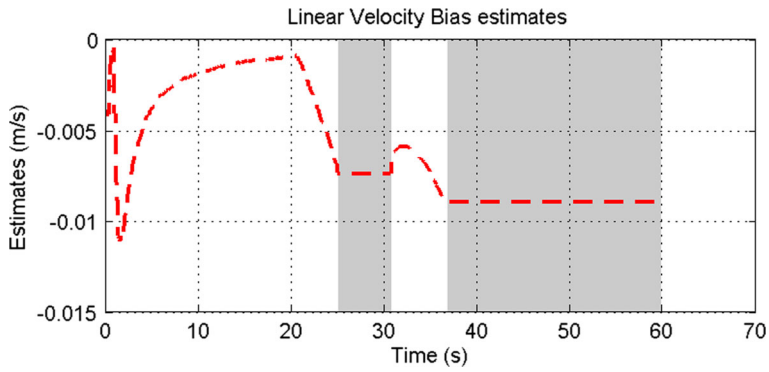


Fig. 18 Linear slippage estimation

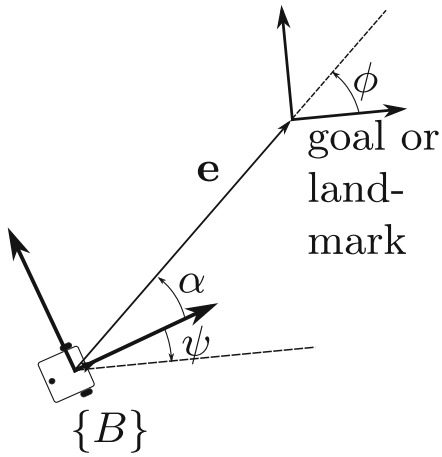


Fig. 19 Mobile robot and docking station definition of frames

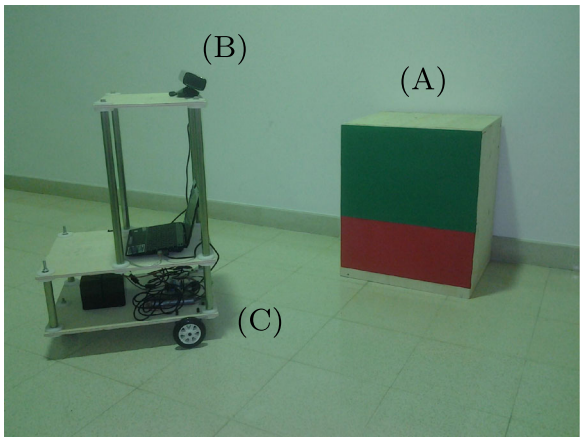


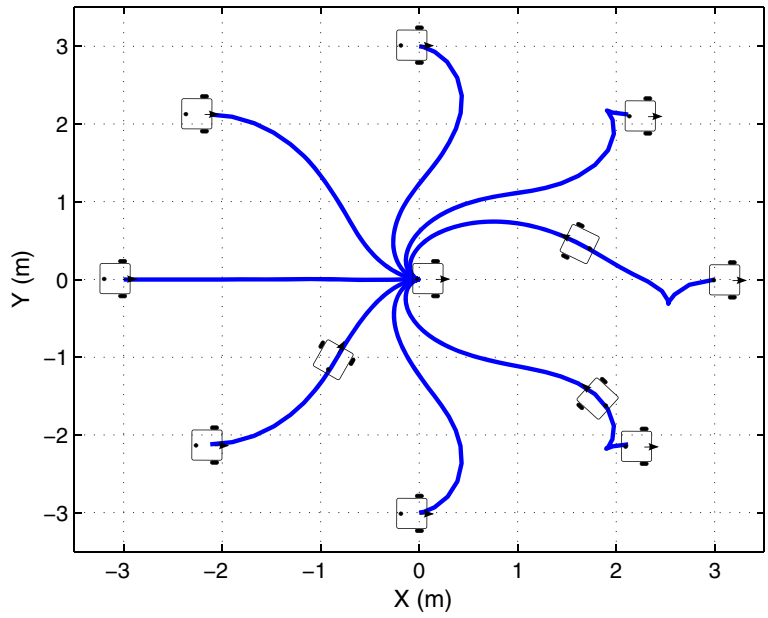
Fig. 21 **A** Docking station, **B** 3D camera and **C** Robot prototype

Figs. 5, 6, 7, for the X, Y and ψ, respectively. In these results, no slippages were estimated, so open-loop trajectory is computed with odometry data alone.

The landmark is not always visible from the robot. Trajectory estimation degrades after a period of 30 seconds where the landmark was not visible and also some corrupted measurements occurred, near the end of the experiment. A statistical representation of these differences can be seen in Fig. 8. When the same trajectory data is processed while also estimating the slippages, some improvement can be noticed, especially in the interval of the trajectory where both angular and linear velocities are maintained. The mentioned data set goes from 20 seconds to 50 seconds from the beginning of the experiment and the slippage estimation effect can be seen when comparing Figs. 5 to 7 with Figs. 9, 10, 11.

The robot kept both velocities nearly constant, the slippage estimation that took place until the 20 second mark was suitable until the 50 second mark, allowing for a reduced open-loop estimation error. The slippage estimates are depicted in Figs. 12 and 13, where the shaded areas correspond to the time intervals of landmark unavailability.

Fig. 20 Trajectories performed with $e(0) = 3$ and $\psi(0) = 0$



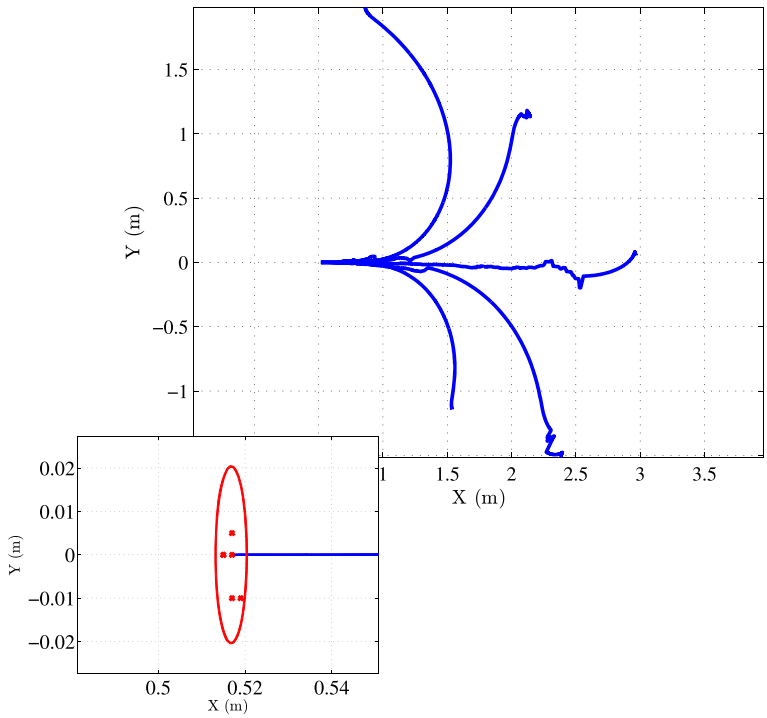


Fig. 22 Docking manoeuvers: 3σ error confidence and final position ground truth (zoom)

With the same data, a different test was conducted, this time forcing the linear slippage $b = -0.01\ (m/s)$ with results as depicted in Fig. 14.

A more intuitive demonstration of the slippage estimation effect in the path estimation is depicted in Figs. 15 and 16 which show that the data relative to a second simpler trajectory, where the robot maintained its linear velocity throughout the whole experiment. The slippage estimation both in the linear and angular velocities, respective to the trajectory in Fig. 15 are depicted in Figs. 17 and 18.

The slippage estimation both in the linear and angular velocities, respective to the trajectory in Fig. 15 are depicted in Figs. 17 and 18.

6.2 Integration of the Localization Solution in a Docking System

Given the promising results obtained, the integration of the proposed localization solution in a docking system was tested. The docking problem is solved with a smooth, time-invariant, globally asymptotically stable feedback control law, which allows for a very human-like closed-loop steering that drives the mobile robot to a certain goal with a desired attitude and a tunable curvature, based on the by now classic work [27]. For that purpose, suppose that the state of the robot is described by $\mathbf{z} = \left[e_x\ e_y\ \psi\right]^T \in \mathbb{R}^3$ composed by the quantities depicted in Fig. 19, and that were introduced in the design of the estimators above.

Representing the docking station position $\mathbf{e}(t) \in \mathbb{R}^2$ and attitude $\psi(t) \in \mathbb{R}$ in the body frame allows for the derivation of linear kinematics and output equations.

Following the work [27], for the state defined as $\mathbf{z} = \left[e_x\ e_y\ \psi\right]^T$, using the new coordinates

$$\begin{cases} e = \|\mathbf{e}\| \\ \alpha = atan(e_y/e_x) \\ \phi = atan(e_y/e_x) - \psi \end{cases},$$

Table 2 Error in docking manoeuvers

	μ	σ
$e_x\ [cm]$	3.5	0.12
$e_y\ [cm]$	−0.5	0.61
$\psi\ [^\circ]$	0.08	0.285

allows the system to be represented, using $\mathbf{Z} = [e\ \alpha\ \phi]^T$, as

$$\begin{cases} \dot{e} = -v\cos\alpha \\ \dot{\alpha} = -\omega + v\frac{\sin\alpha}{e} \\ \dot{\phi} = v\frac{\sin\alpha}{e} \end{cases}, \tag{9}$$

where an isomorphism $g : \mathbb{R}^3 \setminus \{\mathbf{0}\} \mapsto \mathbb{R}^3 \setminus \{\mathbf{0}\}$ is used. Due to the singularity at the origin, Brockett's theorem no longer applies, since the regularity assumptions do not hold, and so the asymptotic stabilization of Eq. 9 is possible, see [28] for details. Following the solution proposed in [27], the control inputs for the linear and angular velocities are given by

$$v = \gamma e\cos\alpha, \quad \gamma > 0,$$

and

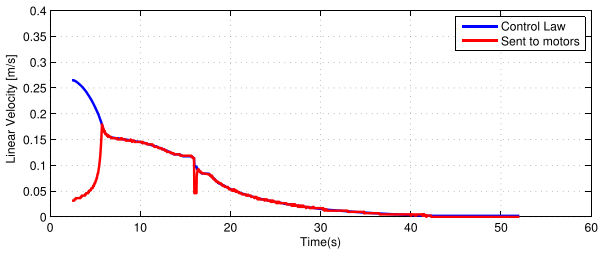
$$\omega = k\alpha + \gamma\frac{\cos\alpha\sin\alpha(\alpha + h\phi)}{\alpha},$$

respectively. For this solution, resorting to Lyapunov theory, it is possible to prove that the origin is globally asymptotically stable.

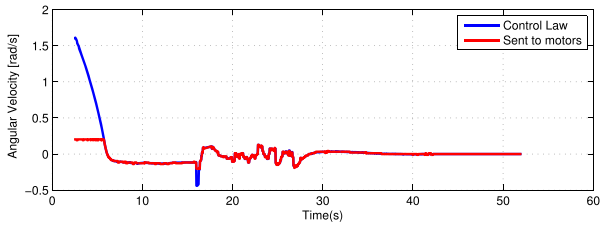
Note that the objective is to dock the vehicle in a certain station with positive linear velocity. It is possible to obtain different trajectories by simply changing the goal objective (to for instance $\mathbf{Z_g} = [0, \pm\pi, \pm\pi]$). The trajectories obtained in simulation are depicted in Fig. 20, where the controller parameters were set to: $\gamma = 3$, $h = 1$, and $k = 6$. It is important to remark that, as intended, the vehicle always arrives at the target location facing the landmark, which goes accordingly with the state vector converging to the origin. The results described above paved the way for an experimental validation with the mobile robot previously used. The robot prototype and landmark setup are shown in Fig. 21. The architecture of the localization system, central to this work, that provides the necessary measurements is the same as represented in Fig. 1.

Without loss of generality, for the tests presented in this subsection, the landmark or docking station is considered to be the origin of the inertial frame. The goal of every experiment presented next was set to 0.5 m in front of the real landmark object used. Also, in every experiment, unless stated otherwise, the initial estimate of the position was set to a value similar to the real position of the mobile robot. Moreover, a saturation of $v_{max} = 0.2\,ms^{-1}$ and $\omega_{max} = 0.2\,ms^{-1}$ was imposed. Figure 22 depicts the localization estimate and true final position for several runs carried out in the laboratory environment. In each one of the six runs, as represented in the zoom of Fig. 22, the prototype was able to perform a successful docking manoeuver, even when the initial estimate had a slight error. The average μ and the standard deviation σ of the final position errors, in the original coordinates, are summarized in Table 2.

In Fig. 23 the commands in a particular experiment are shown, together with the effect of the correction of the estimate in them while Fig. 24 depicts the state variables converging to zero. Notice that, due to the discrete nature of the commands, the vehicle never reaches the goal completely. Thus, the error distance to the target will depend on the value of γ and the error in the camera to body position and rotation calibration values.



(a) Linear velocity command



(b) Angular velocity command

Fig. 23 Docking controller actuation in one of the experiments

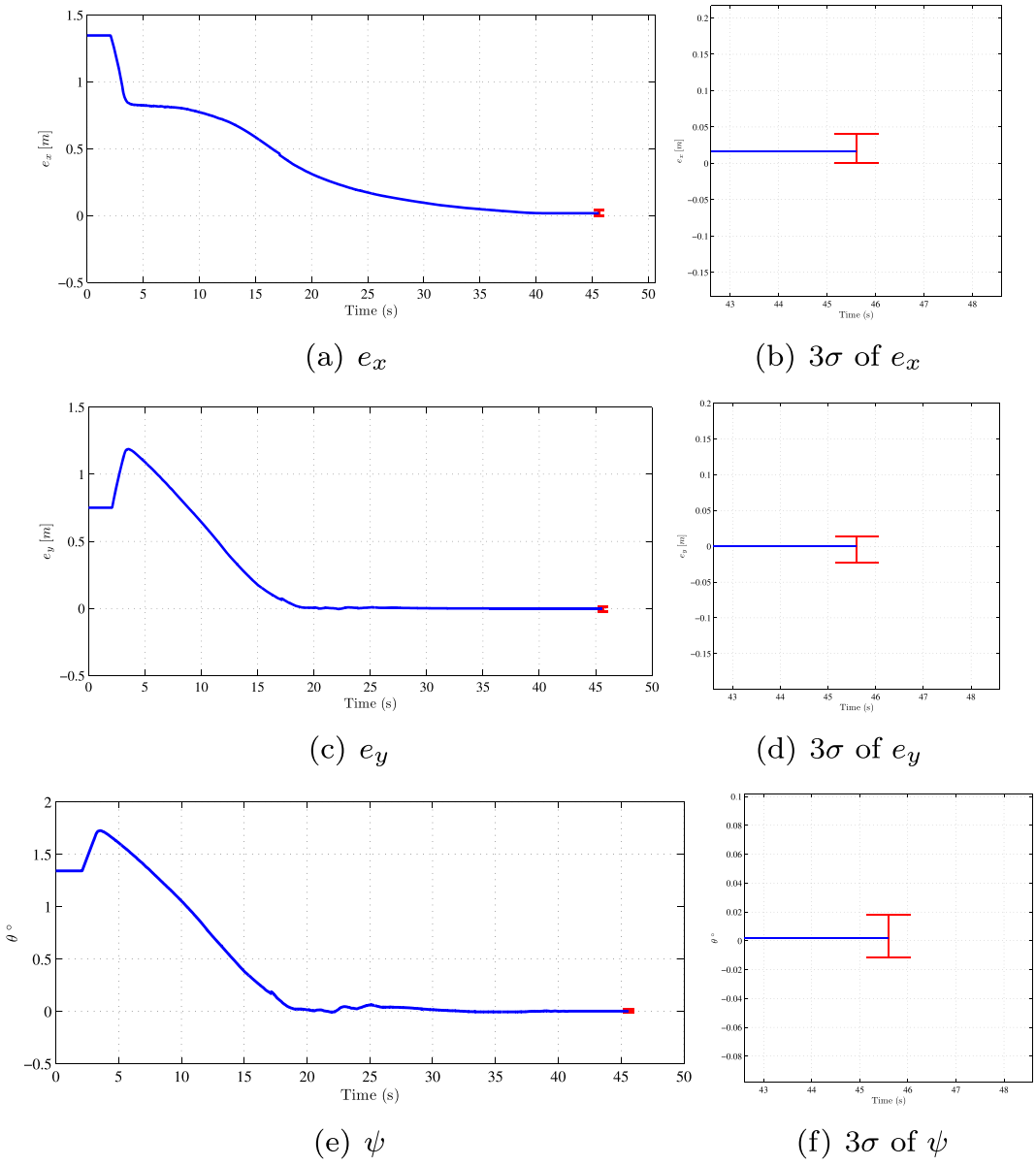


Fig. 24 Time progression of state variables and the 3σ interval for final position

7 Conclusions

This paper proposed a localization system for a mobile robot based on odometric data and RGB-D measurements relative to a landmark, available on board. The localization system is composed of two cascaded estimators, both developed for dynamic models that are observable even in presence of angular and linear slippage. Moreover, the estimation errors of the localization system that was proposed features globally asymptotically stable dynamics. Experiments to assess the performance of the proposed estimators were reported, resorting to a wheeled differential drive mobile robot in an instrumented laboratory. The

integration of the proposed localization solution in a docking system for the same robot was also reported. Simulations and experimental results with the aforementioned robot validate the proposed localization solution, and the performance of the docking solution based on the proposed localization methods central to this work is also illustrated.

Acknowledgments The authors would like to thank Prof. Miguel Tavares da Silva for the use of the IST Biomechanics Laboratory for the ground-truth validation.

References

1. Siegwart, R., Nourbakhsh, I., Scaramuzza, D. Introduction to autonomous mobile robots, 2nd edn. MIT press, Cambridge (2011)
2. Blaer, P., Allen, P.: Topological mobile robot localization using fast vision techniques. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 1, pp. 1031–1036, IEEE (2002)
3. Andreasson, H., Duckett, T.: Topological localization for mobile robots using omni-directional vision and local features. In: Proceedings of IAV, pp. 53–58. IFAC, Lisbon (2004)
4. Bailey, T., Durrant-Whyte, H.: Simultaneous localization and mapping (slam): part ii. IEEE Robot. Autom. Mag. **13**(3), 108–117 (2006)
5. Durrant-Whyte, H., Bailey, T.: Simultaneous localization and mapping: part i. IEEE Robot. Autom. Mag. **13**(2), 99–110 (2006)
6. Bailey, T., Nieto, J., Guivant, J., Stevens, M., Nebot, E.: Consistency of the ekf-slam algorithm. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3562–3568, IEEE (2006)
7. Huang, S., Dissanayake, G.: Convergence and consistency analysis for extended kalman filter based slam. IEEE Trans. Robot. **23**(5), 1036–1049 (2007)
8. Batista, P., Silvestre, C., Oliveira, P.: Optimal position and velocity navigation filters for autonomous vehicles. Automatica **46**(4), 767–774 (2010)
9. Batista, P., Silvestre, C., Oliveira, P.: On the observability of linear motion quantities in navigation systems. Syst. Control Lett. **60**(2), 101–110 (2011)
10. Groover, M. Automation, production systems, and computer-integrated Manufacturing, 3rd edn. Prentice Hall Press, Upper Saddle River, NJ (2007)
11. Hada, Y., Yuta, S.: A first-stage experiment of long term activity of autonomous mobile robot - result of repetitive base-docking over a week. In: Experimental Robotics VII, pp. 229–238. Springer, Berlin Heidelberg New York (2001)
12. Agin, G.J.: Real time control of a robot with a mobile camera. SRI International (1979)
13. Espiau, B., Chaumette, F., Rives, P.: A new approach to visual servoing in robotics. IEEE Trans. Robot. Autom. **8**(3), 313–326 (1992)
14. Samson, C., Espiau, B., Borgne, M.: Robot control: the task function approach. Oxford University Press, Oxford, UK (1991)
15. Skaar, S., Yalda-Mooshabad, I., Brockman, W.: Nonholonomic camera-space manipulation. IEEE Trans. Robot. Autom. **8**(4), 464–479 (1992)
16. Lefebvre, O., Lamiraux, F.: Docking task for nonholonomic mobile robots. In: Proceedings of the 2006 IEEE International Conference on Robotics and Automation, ICRA 2006, pp. 3736–3741, IEEE (2006)
17. Batista, P., Silvestre, C., Oliveira, P.: A two-step control strategy for docking of autonomous underwater vehicles. In: American Control Conference (ACC), pp. 5395–5400, IEEE (2012)
18. Feezor, M.D., Sorrell, F., Blankinship, P., Bellingham, J.: Autonomous underwater vehicle homing/docking via electromagnetic guidance. IEEE J. Ocean. Eng. **26**(4), 515–521 (2001)
19. Park, J., Jun, B., Lee, P., Lee, F., Oh, J.: Experiment on underwater docking of an autonomous underwater vehicleisimi'using optical terminal guidance. In: Oceans 2007-Europe, pp. 1–6, IEEE (2007)
20. McCarthy, C., Barnes, N., Mahony, R.: A robust docking strategy for a mobile robot using flow field divergence. IEEE Trans. Robot. **24**(4), 832–842 (2008)
21. Kim, M., Chong, N.: Direction sensing rfid reader for mobile robot navigation. IEEE Trans. Autom. Sci. Eng. **6**(1), 44–54 (2009)
22. Luo, R., Liao, C., Su, K., Lin, K.: Automatic docking and recharging system for autonomous security robot. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS 2005), pp. 2953–2958, IEEE (2005)
23. Ito, S., Endres, F., Kuderer, M., Diego Tipaldi, G., Stachniss, C., Burgard, W.: W-rgb-d: floor-plan-based indoor global localization using a depth camera and wifi. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 417–422, IEEE (2014)
24. Barbosa, J., Cardeira, C., Oliveira, P., Batista, P., Silvestre, C.: Design and validation of a sensor-based localization algorithm. In: IEEE International Conference on Autonomous Robot Systems and Competitions, pp. 98–103, IEEE (2014)
25. Qualysis motion capture system. http://www.qualisys.com. Accessed: 2013-07-10
26. Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng. **82**(1), 35–45 (1960)
27. Aicardi, M., Casalino, G., Bicchi, A., Balestrino, A.: Closed loop steering of unicycle like vehicles via lyapunov techniques. IEEE Robot. Autom. Mag. **2**(1), 27–35 (1995)
28. Brockett, R.: Asymptotic stability and feedback stabilization. In: Differential Geometric Control Theory, pp. 181–191. Defense Technical Information Center (1983)