

Dynamic Pose-Estimation from the Epipolar Geometry for Visual Servoing of Mobile Robots

H. M. Becerra and C. Sagues

Abstract—In this paper, we propose to exploit the epipolar geometry for dynamic pose-estimation of mobile robots. This is performed using a filtering approach with measurements given by the epipoles. The contribution of the paper is a novel observability analysis using nonlinear and linear tools that leads to achieve an efficient position-based visual servoing (PBVS) approach. Additionally, the visibility constraint problem is solved for this type of approach by using any visual sensor obeying a central projection model. This scheme does not need a target model neither scene reconstruction nor the 3D structure. The effectiveness of the proposed estimation scheme is evaluated via simulations in servoing tasks including obstacles and using omnidirectional vision.

I. INTRODUCTION

In the field of visual servoing [1], most of the efforts have been focused on the enhancement or proposal of image-based (IB) approaches and less attention has been given to the position-based (PB) ones. Thus, the advantages of carrying out a servoing task in the Cartesian space have been wasted, namely, the possibility to define a motion path in accordance to the geometry of the environment and to reduce the dependence on the visual information.

Some works have introduced the notion of dynamic estimation in the control loop for visual servoing (VS) purposes, e.g., [2] and [3]. The former proposes a particular nonlinear observer to track an object using a two-link planar robot. The second paper proposes a filtering approach to estimate the pose in 6 DOF of a robot manipulator. Both of the PB approaches of these papers require a model of the target and use the image point coordinates directly as measurement.

In the context of mobile robots, on one hand, an approach that recovers the robot pose has been proposed using structure from motion in [4]. The authors use the estimated pose in a feedback control law to follow a predefined path, while also the 3D structure of the scene is reconstructed. On the other hand, dynamic estimation has been also used to recover the pose of a mobile robot. The authors of [5] propose a Kalman filtering approach to match a set of landmarks to a prior map and then to estimate the robot pose from these visual observations. Different control tasks can be carried out using this scheme: wall following, leader following and position regulation. The Extended Kalman Filter has been used to recover the pose of a mobile robot and the 3D structure for a homing application in [6]. The effectiveness of applying

a Kalman filtering approach on PBVS has been particularly studied in [7]. In a previous work [8], we have introduced the idea of pose-estimation for mobile robots through the 1D trifocal tensor as measurement, but the observability property of the system has been roughly analyzed.

Given that a VS task with a mobile robot typically implies a significative camera displacement, the use of omnidirectional vision has turned out to be a very good option to keep the target in the field of view [9]. This can be achieved by using catadioptric imaging systems. The imaging process of these systems is represented by a unified model [10], since they have a single center of projection. This has allowed to extend some VS schemes based on a geometric constraint to be used with omnidirectional vision. For instance, researches have exploited the epipolar geometry [11], the homography model [12] and the trifocal tensor [13] in IB approaches.

This paper proposes to exploit the epipolar geometry for dynamic pose-estimation of mobile robots. This is carried out through an EKF-based scheme, which recovers the robot pose (position and orientation) using the kinematic motion model of the onboard camera and the epipoles as measurements. The contribution of the paper is a novel observability analysis using both nonlinear and linear theory, which leads to achieve an efficient PBVS approach. An additional benefit is that our approach solves the visibility constraint problem by using any visual sensor obeying a central projection model. The scheme does not need a target model neither scene reconstruction nor 3D structure. The validity of the estimation is shown through simulations of VS tasks including obstacles.

The paper is organized as follows. Section II describes the kinematic motion model of the camera-robot and introduces the modeling and the epipolar constraint for generic cameras. Section III details the observability analysis from the epipolar geometry and presents the estimation scheme. Section IV shows the performance of the estimation for VS tasks through simulations and Section V states the conclusions.

II. MATHEMATICAL MODELING

A. Camera-Robot Model

This paper focuses on the problem of estimating the state of a differential-drive mobile robot using visual information provided by an onboard camera. Fig. 1 depicts the configuration of the camera-robot system. We assume that the camera is translated a distance ℓ along the longitudinal axis \vec{y}_R . The dynamics of the system's state $\mathbf{x} = [x, y, \phi]^T$ with input vector $\mathbf{u} = [v, \omega]^T$ and output γ can be written as the affine system

This work was supported by project DPI 2009-08126 and grants of Banco Santander-Univ. Zaragoza and Conacyt-México.

H.M. Becerra and C. Sagues are with Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza C/ María de Luna 1, E-50018 Zaragoza, Spain {hector.becerra, csagues}@unizar.es

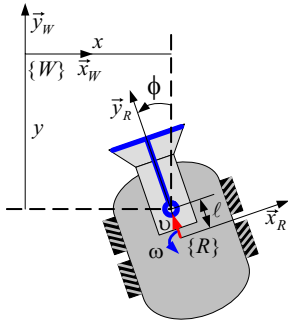


Fig. 1. Kinematic configuration of the robot with an on-board central camera.

$$\dot{\mathbf{x}} = [g_1(\mathbf{x}), g_2(\mathbf{x})] \mathbf{u}, \quad \gamma = h(\mathbf{x}), \quad (1)$$

being $g_1(\mathbf{x}) = [-\sin \phi, \cos \phi, 0]^T$ and $g_2(\mathbf{x}) = [-\ell \cos \phi, -\ell \sin \phi, 1]^T$ smooth input vector fields. The nonlinear function $h(\mathbf{x})$ models a vector of measurements that will be described later. The discrete version of the system (1) is the following

$$\begin{aligned} x_{k+1} &= x_k - \delta (\omega_k \ell \cos \phi_k + v_k \sin \phi_k), \\ y_{k+1} &= y_k - \delta (\omega_k \ell \sin \phi_k - v_k \cos \phi_k), \\ \phi_{k+1} &= \phi_k + \delta \omega_k \end{aligned} \quad (2)$$

where δ is the sampling time. In the sequel, we use the notation $s\phi = \sin \phi$, $c\phi = \cos \phi$. The discrete system (2) can be expressed as follows

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{m}_k, \quad \gamma_k = h(\mathbf{x}_k) + \mathbf{n}_k \quad (3)$$

where the nonlinear function f is the smooth vector field given by the right hand terms of (2). It is assumed that the robot state and the measurements are affected by Gaussian noises \mathbf{m}_k and \mathbf{n}_k , respectively. These noises accomplish $\mathbf{m}_k \sim N(0, \mathbf{M}_k)$, $\mathbf{n}_k \sim N(0, \mathbf{N}_k)$ and $E[\mathbf{m}_{k,i} \mathbf{n}_{k,j}^T] = 0$, with \mathbf{M}_k the state noise covariance and \mathbf{N}_k the measurement noise covariance.

B. Epipolar Geometry (EG) for Generic Cameras

A generic camera has a single center of projection and its image formation process can be modeled as a composition of two central projections [10]. The first is a central projection of a 3D point onto a virtual unitary sphere and the second is a perspective projection onto the image plane. In this work, we assume the use of generic calibrated cameras, which allows us to exploit the representation of the points on the unit sphere. Regarding to Fig. 2(a), let us denote a 3D point as \mathbb{X} , and its corresponding coordinates as \mathbf{X} . Thus, point coordinates on the sphere \mathbf{X}_c can be computed from point coordinates on the normalized image plane \mathbf{x} and the sensor parameter ξ as follows

$$\mathbf{X}_c = (\eta^{-1} + \xi) \bar{\mathbf{x}}, \quad \text{with } \bar{\mathbf{x}} = \left[\mathbf{x}^T, \frac{1}{1+\xi\eta} \right]^T, \quad (4)$$

where $\eta = \frac{-\gamma - \xi(x^2 + y^2)}{\xi^2(x^2 + y^2) - 1}$, $\sigma = \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}$.

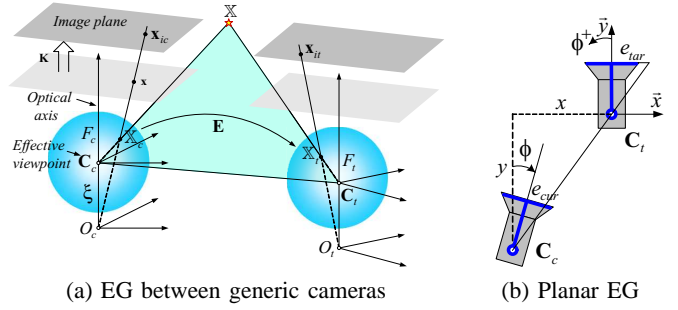


Fig. 2. Generic model of the image formation and epipolar geometry (EG) between generic central cameras.

Let \mathbf{X}_c and \mathbf{X}_t be the coordinates of the 3D point \mathbb{X} projected onto the unit spheres of the current F_c and target frame F_t . The epipolar plane contains the effective viewpoints of the imaging systems \mathbf{C}_c and \mathbf{C}_t , the 3D point \mathbb{X} and the points \mathbb{X}_c and \mathbb{X}_t . The coplanarity of these points leads to the well known epipolar constraint

$$\mathbf{X}_c^T \mathbf{E} \mathbf{X}_t = 0, \quad (5)$$

being \mathbf{E} the essential matrix relating the pair of normalized virtual cameras. Normalized means that the effect of the known calibration matrix has been removed and eventually, the cameras can be represented as virtual perspective. Thus, the epipoles are computed as the right null space of the essential matrix.

Fig. 2(b) shows the configuration of a pair of virtual perspective cameras constrained to planar motion with center of projection \mathbf{C}_c and \mathbf{C}_t respectively. A global reference frame centered in the origin $\mathbf{C}_t = (0, 0, 0)$ of the target viewpoint is defined, so that the current camera location is $\mathbf{C}_c = (x, y, \phi)$. The x -coordinate of the epipoles relating the current and target views can be written as a function of the camera-robot state as follows

$$\begin{aligned} e_{cur} &= \alpha_x \frac{xc\phi + ys\phi}{yc\phi - xs\phi} = \alpha_x \frac{e_{cn}}{e_{cd}}, \\ e_{tar} &= \alpha_x \frac{x}{y} \end{aligned} \quad (6)$$

where α_x is the focal length of the camera in terms of pixel dimensions in the x -direction. For the case of normalized cameras $\alpha_x = 1$. Henceforth, we define the following vector of measurements

$$h(\mathbf{x}) = [h_1 = e_{cur}, h_2 = e_{tar}]^T. \quad (7)$$

III. OBSERVABILITY WITH THE EPIPOLES AS MEASUREMENTS

Observability is a structural property of a system that may affect the convergence of an estimation scheme. This property specifies if two states are distinguishable by measuring the output, i.e., $\mathbf{x}_1 \neq \mathbf{x}_2 \implies h(\mathbf{x}_1) \neq h(\mathbf{x}_2)$. In this section, we aim for showing that distinguishability of the state of the system (1) can be achieved by using the measurement model (7).

There are few works concerning about the state observability of mobile robots. On one hand, some of them take advantage of linearized models to analyze the observability

of the SLAM problem [14]. On the other hand, there also exist some contributions where a nonlinear observability analysis is carried out for localization [15] or SLAM [16]. Some basic results on the observability from visual measurement provided by a geometric constraint (the 1D TT) are reported in [8]. In that work, a linear analysis is used and observability is ensured with three elements of the tensor. In the subsequent, we present a novel and comprehensive observability analysis with a minimum set of measurements and appropriate nonlinear tools.

A. Nonlinear Observability

Firstly, the nonlinear theory for the analysis of continuous systems introduced in [17] is used. According to this theory, the following *observability rank condition* can be enunciated for the case under analysis.

Definition 1. A continuous-time nonlinear system of the form $\dot{\mathbf{x}} = [g_1(\mathbf{x}), g_2(\mathbf{x})] \mathbf{u}$ with a measurement vector $h(\mathbf{x})$ is locally weakly observable if the observability matrix with rows

$$\mathbf{O} \triangleq \left[\nabla L_{g_i g_j}^q h_p(\mathbf{x}) \mid i, j, p = 1, 2; q \in \mathbb{N} \right]^T \quad (8)$$

is of full rank n .

The expression $L_{g_i}^q h_p(\mathbf{x})$ denotes the q th order Lie derivative of the scalar function h_p along the vector field g_i . Thus, the matrix (8) is formed by the gradient vectors $\nabla L_{g_i}^q h_p(\mathbf{x})$ that span a space containing all the possible Lie derivatives. Although the matrix (8) could have infinite number of rows, it suffices to find a set of rows linearly independent in order to fulfill the rank condition. Locally weak observability is a concept stronger than observability, which states that one can instantaneously distinguish each point of the state space from its neighbors, without necessity to travel a considerable distance, as admitted by the observability concept. Next, this definition is used to verify the following lemma.

Lemma 1. The continuous camera-robot system (1) with both epipoles as measurements (7) is a locally weakly observable system. Moreover, this property is maintained even by using only the target epipole as measurement.

Proof: This proof is done by finding the space spanned by all possible Lie derivatives and verifying its dimension. This space is given as

$$\Omega = (h_p, L_{g_1}^1 h_p, L_{g_2}^1 h_p, L_{g_1}^2 h_p, L_{g_2}^2 h_p, \dots)^T, \quad p = 1, 2$$

First, the Lie derivatives given by the current epipole as measurement ($h_1 = e_{cur}$) are presented. As a good approach, the search of functions in the Lie group is constrained for $n - 1$, where $n = 3$ in our case.

$$\begin{aligned} L_{g_1}^1 h_1 &= \nabla h_1 \cdot g_1 = \frac{\alpha_x}{e_{cd}^2} [y, -x, x^2 + y^2] \cdot g_1 = -\alpha_x \frac{e_{cn}}{e_{cd}^2} \\ L_{g_2}^1 h_1 &= \nabla h_1 \cdot g_2 = \alpha_x \frac{x^2 + y^2 - \ell e_{cd}}{e_{cd}^2} \\ L_{g_1}^2 h_1 &= \nabla L_{g_1}^1 h_1 \cdot g_1 = 2\alpha_x \frac{e_{cn}}{e_{cd}^3} \\ L_{g_2}^2 h_1 &= \nabla L_{g_2}^1 h_1 \cdot g_2 = \alpha_x \frac{e_{cn} (2(x^2 + y^2) - 3\ell e_{cd})}{e_{cd}^3} \end{aligned}$$

To verify the dimension of the space spanned by these functions, the gradient operator is applied to obtain the matrix \mathbf{O}_{cur} (9). Given the complexity of the entries of this matrix, only four rows are shown, however, it can be verified that the complete matrix is of rank two. It is required that the gradient of the Lie derivatives obtained from the target epipole as measurement ($h_2 = e_{tar}$) provide one additional row linearly independent to achieve observability. These new Lie derivatives are

$$\begin{aligned} L_{g_1}^1 h_2 &= \nabla h_2 \cdot g_1 = \frac{\alpha_x}{y^2} [y, -x, 0] \cdot g_1 = -\frac{\alpha_x}{y^2} e_{cn} \\ L_{g_2}^1 h_2 &= \nabla h_2 \cdot g_2 = \frac{\alpha_x}{y^2} [y, -x, 0] \cdot g_2 = -\frac{\alpha_x \ell}{y^2} e_{cd} \\ L_{g_1}^2 h_2 &= \nabla L_{g_1}^1 h_2 \cdot g_1 = \frac{2\alpha_x}{y^3} c\phi e_{cn} \\ L_{g_2}^2 h_2 &= \nabla L_{g_2}^1 h_2 \cdot g_2 = \frac{\alpha_x \ell}{y^3} (y e_{cn} - 2\ell s\phi e_{cd}) \end{aligned}$$

By applying the gradient operator, the matrix \mathbf{O}_{tar} (10) is obtained, which effectively provides an additional linearly independent row to the matrix $\mathbf{O} = \begin{bmatrix} \mathbf{O}_{cur}^T & \mathbf{O}_{tar}^T \end{bmatrix}^T$. Indeed, the matrix (10) is full rank by itself, which means that the rank condition of Definition 1 is satisfied by using the target epipole as unique measurement. In summary, the camera-robot system (1) with both epipoles as measurements is locally weakly observable and this property is achieved even by using only the target epipole as measurement. Therefore, the three state variables constituting the camera-robot pose can be estimated from these two measurements. ■

The previous proof implicitly considers the action of both velocities, however, we can analyze the effect for each one of them. For simplicity, this is done using the current epipole as measurement. On one hand, it can be shown from (10) that $\det \left([\nabla h_2^T, \nabla L_{g_1}^1 h_2^T, \nabla L_{g_1}^2 h_2^T]^T \right) = -2\alpha_x y^2 e_{cn}$, which means that, when only a translational velocity is being applied, the matrix loses rank if $e_{cur} = 0$. In other words, observability is lost if the robot is moving forward along the line joining the projection center of the cameras because e_{tar} remains unchanged. Otherwise, observability is guaranteed by a translational velocity different than zero. On the other hand, $\det \left([\nabla h_2^T, \nabla L_{g_2}^1 h_2^T, \nabla L_{g_2}^2 h_2^T]^T \right) = \alpha_x \ell^2 y^2 d(\mathbf{x})$, with $d(\mathbf{x}) \neq 0$ for all $\mathbf{x} \neq 0$. This means that the rotational velocity provides observability iff the camera is shifted from the axis of rotation ($\ell \neq 0$, as assumed), given that in this situation, e_{tar} changes as the robot rotates. Thus, the control strategy should provide the appropriate excitation, at least non-null rotational velocity, in order to ensure observability for any condition.

B. Dynamic Pose-Estimation Scheme

A discrete Kalman filtering approach is proposed to be used in order to estimate the robot pose $\hat{\mathbf{x}}_k = [\hat{x}_k, \hat{y}_k, \hat{\phi}_k]^T$ from the epipoles. An Extended Kalman Filter (EKF) is an effective way to solve this nonlinear estimation problem. The EKF has been applied previously in the VS problem [3], [5], [7] without further analysis. This approach provides

$$\mathbf{O}_{cur} = \begin{bmatrix} \nabla h_1 \\ \nabla L_{g_1}^1 h_1 \\ \nabla L_{g_2}^1 h_1 \\ \nabla L_{g_1}^2 h_1 \end{bmatrix} = \frac{\alpha_x}{e_{cd}^2} \begin{bmatrix} y & -x & x^2 + y^2 \\ -(y + s\phi e_{cn})/e_{cd} & (x + c\phi e_{cn})/e_{cd} & -(x^2 + y^2 + e_{cn}^2)/e_{cd} \\ -(\ell s\phi e_{cd} - 2ye_{cn})/e_{cd} & (\ell c\phi e_{cd} - 2xe_{cn})/e_{cd} & e_{cn}(2(x^2 + y^2) - \ell e_{cd})/e_{cd} \\ (2c\phi e_{cd} + 6s\phi e_{cn})/e_{cd} & (2s\phi e_{cd} - 6c\phi e_{cn})/e_{cd} & (2(y c\phi - x s\phi)^2 + 6e_{cn}^2)/e_{cd} \end{bmatrix} \quad (9)$$

$$\mathbf{O}_{tar} = \begin{bmatrix} \nabla h_2 \\ \nabla L_{g_1}^1 h_2 \\ \nabla L_{g_2}^1 h_2 \\ \nabla L_{g_1}^2 h_2 \\ \nabla L_{g_2}^2 h_2 \end{bmatrix} = \frac{\alpha_x}{y^4} \begin{bmatrix} y^3 & -y^2 x & 0 \\ -y^2 c\phi & y(y s\phi + 2x c\phi) & -y^2(y c\phi - x s\phi) \\ \ell y^2 s\phi & \ell y(y c\phi - 2x s\phi) & \ell y^2(x c\phi + y s\phi) \\ 2y c\phi^2 & -2c\phi(3x c\phi + 2y s\phi) & 2y(y(c\phi^2 - s\phi^2) - 2x s\phi c\phi) \\ -\ell y(y c\phi + 2\ell s^2\phi) & -\ell(y^2 s\phi + y c\phi(2x - 4\ell s\phi) + 6\ell x s^2\phi) & \ell y(y - 2\ell c\phi)e_{cd} + 2\ell s\phi e_{cn} \end{bmatrix} \quad (10)$$

generality to the proposed scheme in comparison to nonlinear observers designed for a particular system. So, the well-known basic form of the EKF is used [20]. The required matrices from the linearization of the camera-robot model (2) and measurement (7) are as follows

$$\mathbf{F}_k = \left. \frac{\partial f}{\partial \mathbf{x}_k} \right|_{\substack{\mathbf{x}_k = \hat{\mathbf{x}}_k^+ \\ \mathbf{m}_k = 0}} = \begin{bmatrix} 1 & 0 & \Delta_{y,k} \\ 0 & 1 & -\Delta_{x,k} \\ 0 & 0 & 1 \end{bmatrix}_{\phi_k = \hat{\phi}_k^+},$$

$$\mathbf{G}_k = \left. \frac{\partial f}{\partial \mathbf{u}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k^+} = \begin{bmatrix} -s\phi_k & -\ell c\phi_k \\ c\phi_k & -\ell s\phi_k \\ 0 & 1 \end{bmatrix}_{\phi_k = \hat{\phi}_k^+},$$

$$\mathbf{H}_k = \left. \frac{\partial h}{\partial \mathbf{x}_k} \right|_{\substack{\mathbf{x}_k = \hat{\mathbf{x}}_k^- \\ \mathbf{n}_k = 0}} = \begin{bmatrix} \frac{\alpha_x}{e_{cd,k}^2} [y_k, -x_k, x_k^2 + y_k^2] \\ \frac{\alpha_x}{y_k} [y_k, -x_k, 0] \end{bmatrix}_{\mathbf{x}_k = \hat{\mathbf{x}}_k^-}$$

where $\Delta_{x,k} = \delta(\omega_k \ell c\phi_k + v_k s\phi_k)$, $\Delta_{y,k} = \delta(\omega_k \ell s\phi_k - v_k c\phi_k)$ and $e_{cd,k} = y_k c\phi_k - x_k s\phi_k$. Up to now, we have proved the observability of the camera-robot system with the epipoles as measurement from a nonlinear point of view. However, since the EKF is based on the previous linearization, an appropriate observability analysis is presented in the following lemma for the linear approximation ($\mathbf{F}_k, \mathbf{G}_k, \mathbf{H}_k$).

Lemma 2. *The linear approximation ($\mathbf{F}_k, \mathbf{G}_k, \mathbf{H}_k$) of the discrete nonlinear system (2) and measurements (7) as required for the EKF-based estimation scheme is an observable system. Moreover, observability is achieved by using only the target epipole as measurement.*

Proof: Firstly, we verify the property of local observability, which is given by the typical observability matrix

$$\mathbf{O}_k = \begin{bmatrix} \mathbf{H}_k^T & (\mathbf{H}_k \mathbf{F}_k)^T & \cdots & (\mathbf{H}_k \mathbf{F}_k^{n-1})^T \end{bmatrix}^T.$$

This is a 6×3 matrix built by stacking the following local observability matrices (LOM) for each measurement

$$\mathbf{O}_{cur,k} = \frac{\alpha_x}{e_{cd,k}^2} \begin{bmatrix} y_k & -x_k & x_k^2 + y_k^2 \\ y_k & -x_k & \Sigma_k + x_k^2 + y_k^2 \\ y_k & -x_k & 2\Sigma_k + x_k^2 + y_k^2 \end{bmatrix},$$

$$\mathbf{O}_{tar,k} = \frac{\alpha_x}{y_k} \begin{bmatrix} y_k & -x_k & 0 \\ y_k & -x_k & \Sigma_k \\ y_k & -x_k & 2\Sigma_k \end{bmatrix}$$

where $\Sigma_k = y_k \Delta_{y,k} + x_k \Delta_{x,k}$. It can be seen that the matrix $\mathbf{O}_k = [\mathbf{O}_{cur,k}^T \ \mathbf{O}_{tar,k}^T]^T$ is of rank 2 and the linear approximation is not observable for each instant time.

The linearization can be seen as a piece-wise constant system (PWCS) for each instant time k . The observability of PWCS can be studied using the so-called stripped observability matrix (SOM) for a number r of instant times, as introduced in [18]. The SOM is defined from the local observability matrices \mathbf{Q}_k for different instant times as follows

$$\mathbf{O}_{SOM,r} = [\mathbf{O}_k^T \ \mathbf{O}_{k+1}^T \ \cdots \ \mathbf{O}_{k+r}^T]^T.$$

According to this theory, when it is satisfied that $\mathbf{F}_k \mathbf{x}_k = \mathbf{x}_k \ \forall \ \mathbf{x}_k \in \text{NULL}(\mathbf{O}_k)$, then the discrete PWCS is completely observable iff $\mathbf{O}_{SOM,r}$ is of rank n . This claims that observability can be gained in a number of steps r even if local observability is not ensured, as in this case. It can be verified that the null space basis of the matrix \mathbf{O}_k is any state $\mathbf{x}_k = \lambda [x_k \ y_k \ 0]^T$, where $\lambda \in R$. This subset of the state space satisfies $\mathbf{F}_k \mathbf{x}_k = \mathbf{x}_k$, so that, the observability can be determined through $\mathbf{O}_{SOM,r}$ for some r .

In order to get a smaller SOM, we use the LOM obtained from the target epipole ($\mathbf{O}_{tar,k}$). This LOM for the next instant time is

$$\mathbf{O}_{tar,k+1} = \frac{\alpha_x}{y_{k+1}^2} \begin{bmatrix} y_{k+1} & -x_{k+1} & 0 \\ y_{k+1} & -x_{k+1} & \Sigma_{k+1} \\ y_{k+1} & -x_{k+1} & 2\Sigma_{k+1} \end{bmatrix}.$$

The two-step stripped observability matrix $\mathbf{O}_{SOM,1} = [\mathbf{O}_{tar,k}^T \ \mathbf{O}_{tar,k+1}^T]^T$ can be reduced by Gaussian elimination to a 3×3 triangular matrix whose determinant is $-2x_k^2 \Delta_{x,k} \Delta_{y,k} + 2x_k y_k \Delta_{x,k}^2 - 2x_k y_k \Delta_{y,k}^2 + 2y_k^2 \Delta_{x,k} \Delta_{y,k}$. Therefore, under the assumption of sampling time different than zero, this matrix is full rank and the linear approximation ($\mathbf{F}_k, \mathbf{G}_k, \mathbf{H}_k$) is observable iff non-null velocities are applied at each instant time. Moreover, a rotational velocity different than zero is enough to achieve observability iff $\ell \neq 0$, which agrees with the comments after Lema 1. ■

It is worth emphasizing that both previous Lemmas are valid for any pair of images, which allows us to exploit the benefit of changing the measurements online that is provided by the Kalman filtering approach. In order to drive a mobile robot to a desired target, we propose to use the Cartesian controller introduced in our previous paper [8], with feedback of the estimated pose $\hat{\mathbf{x}}_k = [\hat{x}_k, \hat{y}_k, \hat{\phi}_k]^T$ given by the scheme of the previous section. This controller is able to control the robot position tracking the following generic parabolic path

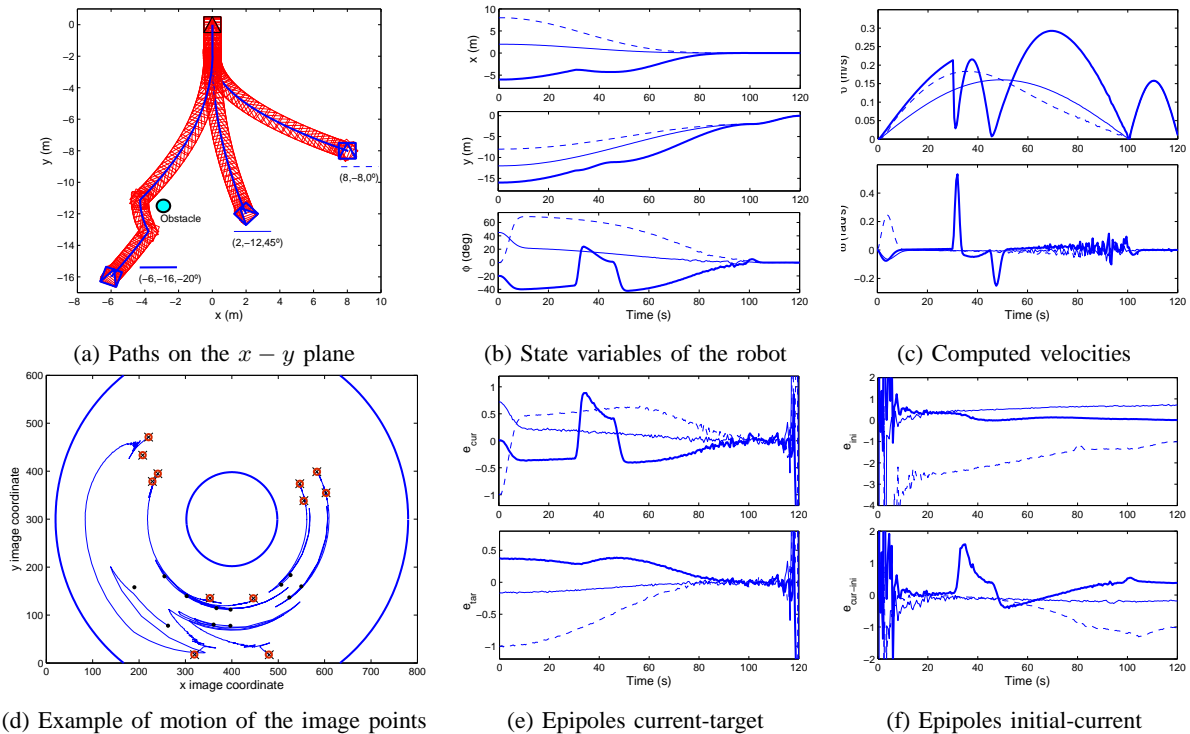


Fig. 3. Simulation results for some visual servoing tasks using feedback of the estimated pose. The motion of the robot starts from three different initial locations and for one of them, an obstacle appears and its avoidance is carried out.

$$\begin{aligned}
 y_k^d &= \frac{y^i - y^f}{2} \left(1 + \cos \left(\frac{\pi - k\delta}{\tau_{st}} \right) \right) + y^f, \\
 x_k^d &= \frac{x^i - x^f}{(y^i - y^f)^2} (y_k^d - y^f)^2 + x^f
 \end{aligned} \quad (12)$$

where (x^i, y^i) and (x^f, y^f) are the initial position and the desired final position of the stage, respectively, which is carried out in τ_{st} seconds. By using this path for the robot position, we can define an intermediate aligned goal in order to avoid the problem of short baseline when the robot is reaching the target. In a second stage, the pose is estimated from the epipoles relating the initial and the current images, which behave adequately. In this stage only the y -coordinate obeys the sinusoidal reference and $x_k^d = 0$. Notice that, similarly to how the intermediate aligned location is defined, we are able to set any other goal through (12). So, this provides the possibility to avoid an obstacle detected over the path toward the target, as shown in the next section.

IV. EVALUATION OF RESULTS

The validity of the proposed estimation scheme for servoing tasks is shown via simulations performed in Matlab with a sampling period of 0.5 s. The epipoles are estimated using an eight-point algorithm from synthetic omnidirectional images of size 800×600 , which are generated through the generic camera model [10]. We use our controller proposed in [8], with $\ell = 8$ cm and adequate control gains. Related to the Kalman filtering, we use small standard deviations in N_k and M_k , so that good confidence is given to the measurements. Image noise of standard deviation 0.5 pixel

has been added. We set $\mathbf{P}_0 = \text{diag}(5^2 \text{ cm}^2, 5^2 \text{ cm}^2, 2^2 \text{ deg}^2)$. For efficiency, the simulations are carried out using only one epipole as measurement.

A. Visual servoing from the estimated pose

Fig. 3(a) shows an upper view of the robot motion on the plane for three different initial locations. In any case the robot is successfully driven to the target in 120 s. In one case, a fixed obstacle is avoided by defining an adequate subgoal using (12). We assume that the obstacle detection is provided accordingly. In Fig. 3(b), it can be seen that in a first stage the intermediate aligned location $(0, -2, 0^\circ)$ is reached at 100 s. After that, the measurements are changed to avoid the short baseline problem. The two stages can be appreciated in the velocities of Fig. 3(c). The same translational velocity is computed for the final rectilinear motion. Note that the velocities excite the system adequately ensuring observability. As an example, Fig. 3(d) shows the motion of the image points for the case with obstacle for a hypercatadioptric camera. Similar overall results can be obtained with paracatadioptric. The epipoles computed from twelve image points along the sequence are shown in Fig. 3(e)-(f) for each case. The epipole e_{tar} is used as measurement during 100 s and after that, when it becomes unstable, the e_{ini} is used for 20 s. During this time e_{ini} changes as the robot moves and observability is achieved given that the translational velocity is non-null.

Table I shows the final error of the camera-robot pose obtained as the average of the final pose from 100 Monte Carlo runs. According to this, the VS task is accomplished with good accuracy.

TABLE I

FINAL LOCATION REACHING THE TARGET (0,0,0°) FOR PATHS IN FIG. 3.

	(8m,-8m,0°)	(2m,-12m,45°)	(-6m,-16m,-20°)
x_{end} (cm)	0.63 ($\sigma = 1.11$)	0.90 ($\sigma = 1.03$)	-0.90 ($\sigma = 0.74$)
y_{end} (cm)	-0.16 ($\sigma = 1.20$)	-0.54 ($\sigma = 0.18$)	-0.37 ($\sigma = 0.27$)
ϕ_{end} (°)	0.10 ($\sigma = 0.51$)	-0.10 ($\sigma = 0.46$)	0.05 ($\sigma = 0.12$)

B. Performance of the Estimation Scheme

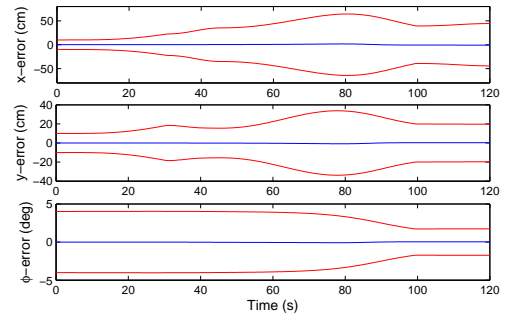
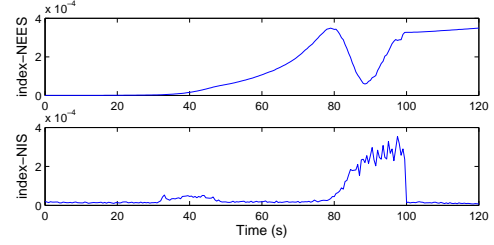
We evaluate the performance of the estimation for the task with obstacle avoidance, however, similar results are obtained in general. Fig. 4(a) presents the average state estimation errors over all 100 Monte Carlo runs for each time step. The computed error is the difference between the truth state given by the robot model and the estimated state. It can be seen that each one of the three estimation errors are maintained within the 2σ confidence bounds. Additionally, a consistency test is carried out to determine if the computed covariances match the actual estimation errors. Two consistency indexes are used: the Normalized Estimate Error Square (NEES) and the Normalized Innovation Squared (NIS) [20]. In any case, if the index is less than the unity the estimation is consistent, otherwise it is optimistic or inconsistent and the estimation may diverge. Fig. 4(b) shows the average indexes over the same 100 Monte Carlo runs. According to this, the EKF is always consistent in spite of the nonlinearities of the state model and measurement model.

V. CONCLUSIONS

In this paper, we have introduced the epipolar geometry for dynamic pose-estimation of mobile robots. The robot position and orientation is recovered using the epipoles as measurements, with the benefit of a temporal filtering, in comparison with a static approach, as the essential matrix decomposition. As main interest of the paper, we present a comprehensive observability analysis using nonlinear and linear tools, which leads to propose an efficient position-based control. Additionally, this is a generic approach that solves the visibility constraint problem by using any visual sensor obeying a central projection model. The scheme does not need a target model neither scene reconstruction nor depth information. Simulations have proved the good performance of the proposed estimation scheme in visual servoing tasks even when maneuvering for obstacle avoidance.

REFERENCES

- [1] F. Chaumette and S. Hutchinson. Visual servo control Part I: Basic approaches. *IEEE Robotics and Autom. Mag.*, 13(14):82–90, 2006.
- [2] K. Hashimoto and H. Kimura. Visual servoing with nonlinear observer. In *IEEE Int. Conf. on Robotics and Automation*, pages 484–489, 1995.
- [3] W. J. Wilson, C. C. W. Hulls, and G. S. Bell. Relative end-effector control using cartesian position-based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5):684–696, 1996.
- [4] E. Royer, M. Lhuillier, M. Dhome, and J. M. Lavest. Monocular vision for mobile robot localization and autonomous navigation. *Int. Journal of Computer Vision*, 74(3):237–260, 2007.
- [5] A. K. Das, R. Fierro, V. Kumar, B. Southall, J. Spletzer, and C. J. Taylor. Real-time vision-based control of a nonholonomic mobile robot. In *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, pages 1714–1718, 2001.

(a) State estimation errors and 2σ uncertainty bounds

(b) Consistency of the estimation

Fig. 4. Average values of the estimation errors and consistency indexes over 100 Monte Carlo runs for the initial location $(-6, -16, -20^\circ)$.

- [6] T. Goedeme, T. Tuytelaars, L. V. Gool, G. Vanacker, and M. Nuttin. Feature based omnidirectional sparse visual path following. In *IEEE/RSJ Int. Conf. on Intell. Rob. and Sys.*, pages 1806–1811, 2005.
- [7] A. Shademan and F. Janabi-Sharifi. Sensitivity analysis of EKF and iterated EKF pose estimation for position-based visual servoing. In *IEEE Conf. on Control Applications*, pages 755–760, 2005.
- [8] H. M. Becerra and C. Sagües. Pose-estimation-based visual servoing for differential-drive robots using the 1D trifocal tensor. In *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, pages 5942–5947, 2009.
- [9] H. H. Abdalkader, Y. Mezouar, N. Andreff, and P. Martinet. Image-based control of mobile robot with central catadioptric cameras. In *IEEE Int. Conf. on Robotics and Automation*, pages 3522–3527, 2005.
- [10] C. Geyer and K. Daniilidis. An unifying theory for central panoramic systems and practical implications. In *European Conf. on Computer Vision*, pages 445–461, 2000.
- [11] G. L. Mariottini, D. Prattichizzo, and G. Oriolo. Image-based visual servoing for nonholonomic mobile robots with central catadioptric camera. In *IEEE Int. Conf. on Rob. and Aut.*, pages 538–544, 2006.
- [12] S. Benhimane and E. Malis. A new approach to vision-based robot control with omni-directional cameras. In *IEEE Int. Conf. on Robotics and Automation*, pages 526–531, 2006.
- [13] H. M. Becerra, G. López-Nicolás, and C. Sagües. Omnidirectional visual control of mobile robots based on the 1D trifocal tensor. *Robotics and Autonomous Systems*, 58(6):796–808, 2010.
- [14] T. Vidal-Calleja, M. Bryson, S. Sukkarieh, A. Sanfeliu, and J. Andrade-Cetto. On the observability of bearing-only SLAM. In *IEEE Int. Conf. on Robotics and Automation*, pages 4114–4119, 2007.
- [15] A. Martinelli and R. Siegwart. Observability analysis for mobile robots localization. In *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, pages 1471–1476, 2005.
- [16] K. W. Lee, W. S. Wijesoma, and J. Ibanez-Guzman. On the observability and observability analysis of SLAM. In *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, pages 3569–3574, 2006.
- [17] R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Trans. on Automatic Control*, 22(5):728–740, 1977.
- [18] D. Goshen-Meskin and I. Y. Bar-Itzhack. Observability analysis of piece-wise constant systems - Part I: Theory. *IEEE Trans. on Aerospace and Electronic Systems*, 28(4):1056–1067, 1992.
- [19] G. López-Nicolás, C. Sagües, J.J. Guerrero, D. Kragic, and P. Jensfelt. Switching visual control based on epipoles for mobile robots. *Robotics and Autonomous Systems*, 56(7):592–603, 2008.
- [20] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley and Sons, New York, 2001.