The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Pose-Estimation-Based Visual Servoing for Differential-Drive Robots using the 1D Trifocal Tensor

H. M. Becerra and C. Sagues

Abstract— **A pose-estimation-based approach to perform visual control for differential-drive robots is presented in this paper. Our scheme recovers the camera location (position and orientation) using an Extended Kalman Filter (EKF) algorithm with the 1D trifocal tensor (TT) as measurement model. This new visual servoing scheme allows knowing the real world path performed by the robot without the computational load introduced by position-based approaches. A state-estimated feedback control law is designed to solve a tracking problem for the lateral and longitudinal robot coordinates. The desired trajectories to be tracked ensure total correction of both position and orientation using a single control law, even though the orientation is a DOF in the control system. The effectiveness of our approach is tested via simulations.**

I. INTRODUCTION

A promising sensor-based control of mobile robots can be found on the basis of visual servoing (VS). Vision as the main source of information on a robotic platform has allowed to improve its navigation capabilities in a single robot task or in robot coordination tasks. Recently, in order to extract matched features from images in a robust way, some geometric constraints relating two views have been applied. Two of them have been well exploited to control mobile robots, epipolar geometry (for instance [1], [2]) and the homography model ([3], [4]). Nevertheless, these geometric constraints have both serious drawbacks. Epipolar geometry is ill-conditioned with short baseline and with planar scenes, while the homography model is not well defined if there are no dominant planes in the scene or with large baselines.

In this field of geometric constraint-based control a less explored is the trifocal tensor (TT). We propose to exploit its properties in order to overcome the drawbacks of the typical geometric constraints. The TT describes all the geometric relations between three views and is independent of the observed scene [5]. In the general case, it is a $3 \times 3 \times 3$ TT, however, when the motion is constraint to be planar, the TT can be expressed with eight elements. The TT has proved its effectiveness to recover the robot location in [6] and [7]. In the first work conventional cameras and artificial landmarks are used, while in the second one both conventional and omnidirectional cameras are used to analyze linear approaches of estimating the TT. Both of these works propose the TT to be used for initialization of bearing-only SLAM algorithms. In [8], a visual control for mobile robots based on the elements of a 2D TT constrained to a planar motion is presented. It shows good performance reaching the target location, however, the stability properties of the controller should be studied.

In image-based VS approaches there exist an intrinsic problem because the real world path that is performed by the robot is unknown. Whereas position-based approaches overcome this issue, the control loop is very slow. A possibility that has not been explored for mobile robots is to estimate the location using an observer scheme from machine vision. There are two papers in which visual servoing with observer is introduced for applications with manipulators ([9], [10]). A visual tracking scheme that includes a kinematic model for the object to be tracked is used. An approach that introduces estimation in a visual servoing scheme for mobile robots is presented in [11]. It proposes a method to obtain just depth estimation for point features using a nonlinear observer. Another work that is related to camera-motion estimation is [12], which presents an EKF algorithm to tackle the vision-based pose-tracking problem for augmented reality applications. A constant velocity motion model is used as dynamic system and the TT constraint is incorporated into the measurement model.

Using the classical teach-by-showing strategy, this paper proposes a visual control scheme for differential-drive robots based on the recovery of on-line camera motion with an EKF algorithm. The kinematic motion model of the camera mounted on the robot is used as dynamic system and the TT as measurement model. We concern for the observability analysis of the discrete linear approximation of the system. Observability can be assured by selecting a suitable set of measurements from the TT elements. The kinematic motion model allows us to design a static state feedback control law to track desired trajectories for the lateral and longitudinal robot coordinates. Thus, the proposed visual control solves the problem of not knowing the real world path of image-based schemes, but without the computational load introduced by position-based approaches. Moreover, the defined path allows to correct position and orientation simultaneously using smooth control inputs, regardless of the orientation is a degree of freedom in the control system.

The paper is organized as follows. Sect. II describes the kinematic motion model of the robot-camera and its relationship with the TT. Sect. III details the EKF implementation and issues on observability. Sect. IV presents the synthesis of the control law. In Sect. V the stability is analyzed. Sect. VI shows the control system performance through simulations, and finally, Sect. VII provides conclusions.

This work was supported by projects DPI 2006-07928/DPI 2009-08126, IST-1-045062-URUS-STP and grants of Banco Santander-Univ. Zaragoza and Conacyt-México.

H.M. Becerra and C. Sagues are with DIIS-I3A, Universidad de Zaragoza C/ María de Luna 1, E-50018 Zaragoza, Spain {hector.becerra, csagues}@unizar.es

II. MATHEMATICAL MODELING

A. Robot-Camera Model

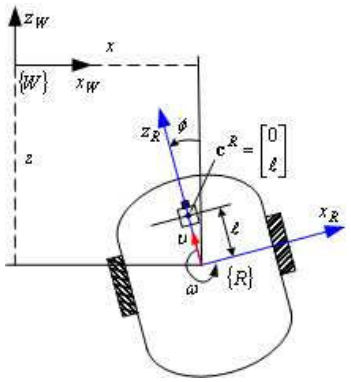


Fig. 1. Kinematic configuration of the robot.

This work focuses on controlling a differential-drive robot under the framework that is depicted in Fig. 1. The velocity vector in the robot reference frame $\{R\}$ is $\mathbf{s}^R = \begin{bmatrix} 0 & \upsilon \end{bmatrix}^T$. This velocity in the world reference frame $\{W\}$ is given by $\mathbf{s}^W = \begin{bmatrix} \dot{x} & \dot{z} \end{bmatrix}^T = \mathbf{R}\mathbf{s}^R$, where $\mathbf{R}$ is the rotation matrix relating both reference frames $\mathbf{R} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$. The decomposed linear velocity together with the angular velocity gives the relationships $\dot{x} = -\upsilon\sin\phi$, $\dot{z} = \upsilon\cos\phi$ and $\dot{\phi} = \omega$. These equations represent the dynamics of the robot reference frame $\{R\}$ with respect to the world frame. Now we are interested in knowing the dynamics of a point translated a distance ℓ along the z_R axis, i.e. the point $\mathbf{c}^R$, where a camera is fixed to the robot. Using a general transformation between frames, which is given as $\mathbf{c}^W = \mathbf{R}\mathbf{c}^R + \mathbf{t}$, its derivative is $\dot{\mathbf{c}}^W = \mathbf{R}\dot{\mathbf{c}}^R + \dot{\mathbf{R}}\mathbf{c}^R + \dot{\mathbf{t}}$. Applying this expression to the point $\mathbf{c}^R = \begin{bmatrix} 0 & \ell \end{bmatrix}^T$, which has no relative velocity with respect to the robot reference frame $\dot{\mathbf{c}}^R = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$ and knowing that

$$\dot{\mathbf{t}} = \begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\upsilon\sin\phi \\ \upsilon\cos\phi \end{bmatrix} \text{ and } \dot{\phi} = \omega,$$

$$\dot{\mathbf{R}} = \begin{bmatrix} -\omega\sin\phi & -\omega\cos\phi \\ \omega\cos\phi & -\omega\sin\phi \end{bmatrix},$$

it results in the following system

$$\begin{aligned} \dot{x} &= -\omega\ell\cos\phi - \upsilon\sin\phi, \\ \dot{z} &= -\omega\ell\sin\phi + \upsilon\cos\phi, \\ \dot{\phi} &= \omega. \end{aligned} \tag{1}$$

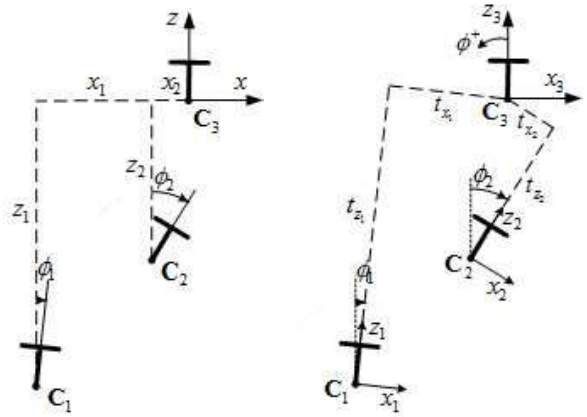
It is important to keep in mind that such system represents the dynamics of the point where the camera is fixed to the robot with respect to the world frame, and thus, any subscript is avoided from now on. Applying an Euler approximation (forward difference) for the continuous derivative, we obtain the following discrete time nonlinear system.

$$\begin{aligned} x_{k+1} &= x_k - \delta\omega_k\ell\cos\phi_k - \delta\upsilon_k\sin\phi_k, \\ z_{k+1} &= z_k - \delta\omega_k\ell\sin\phi_k + \delta\upsilon_k\cos\phi_k, \\ \phi_{k+1} &= \phi_k + \delta\omega_k. \end{aligned} \tag{2}$$

We can write the state vector as $\mathbf{x}_k = \begin{bmatrix} x_k & z_k & \phi_k \end{bmatrix}^T$ and the input vector as $\mathbf{u}_k = \begin{bmatrix} \upsilon_k & \omega_k \end{bmatrix}^T$. In general, an increment of input is given by $\delta\mathbf{u}_k = T_s\mathbf{u}_k$, where T_s is the sampling time. In the sequel, we use the notation $s\beta = \sin\beta$, $c\beta = \cos\beta$.

B. The 1D Trifocal Tensor

The TT relates geometrically three views. It only depends on the relative locations of the observed scene in the three images. Let us define a global (world) reference frame as depicted in Fig. 2(a) with the origin in the third camera. Then, the camera locations with respect to that global reference are $\mathbf{C}_1 = (x_1, z_1, \phi_1)$, $\mathbf{C}_2 = (x_2, z_2, \phi_2)$ and $\mathbf{C}_3 = (0, 0, 0)$. We assume that the motion is constrained to be planar. The relative locations between cameras is defined by a local reference frame in each camera as is shown in Fig. 2(b).



(a) (b)

Fig. 2. (a) Global reference definition, (b) Relative location between cameras.

The expression of the tensor as it is obtained from metric information of the three views is

$$\begin{aligned} T_1 &= \begin{bmatrix} T_{111} & T_{112} \\ T_{121} & T_{122} \end{bmatrix} \\ &= \begin{bmatrix} t_{z_1}s\phi_2 - t_{z_2}s\phi_1 & -t_{z_1}c\phi_2 + t_{z_2}c\phi_1 \\ t_{z_1}c\phi_2 + t_{x_2}s\phi_1 & t_{z_1}s\phi_2 - t_{x_2}c\phi_1 \end{bmatrix}, \\ T_2 &= \begin{bmatrix} T_{211} & T_{212} \\ T_{221} & T_{222} \end{bmatrix} \\ &= \begin{bmatrix} -t_{x_1}s\phi_2 - t_{z_2}c\phi_1 & t_{x_1}c\phi_2 - t_{z_2}s\phi_1 \\ -t_{x_1}c\phi_2 + t_{x_2}c\phi_1 & -t_{x_1}s\phi_2 + t_{x_2}s\phi_1 \end{bmatrix} \end{aligned} \tag{3}$$

where $t_{x_i} = -x_ic\phi_i - z_is\phi_i$, $t_{z_i} = x_is\phi_i - z_ic\phi_i$ for $i = 1, 2$. Equations in (3) can be verified as described in [7].

We can note that in the elements of the TT we have eight nonlinear relationships relating information of camera locations. From now on, let us define the initial location of a camera mounted on a robot as described in Fig. 1 to be (x_1, z_1, ϕ_1), the target location $(0, 0, 0)$ and $(x_2(t), z_2(t), \phi_2(t))$ the current location, which varies as the robot moves. The rates of change of the current location

are given by (1). We can obtain the Jacobian of each element of the tensor as follows

$$\frac{\partial \mathbf{T}_{ijk}}{\partial \mathbf{x}_2} = \begin{bmatrix} -s\phi_1 s\phi_2 & s\phi_1 c\phi_2 & t_{z1}c\phi_2 + t_{x2}s\phi_1 \\ c\phi_1 s\phi_2 & -c\phi_1 c\phi_2 & t_{z1}s\phi_2 - t_{x2}c\phi_1 \\ -s\phi_1 c\phi_2 & -s\phi_1 s\phi_2 & -t_{z1}s\phi_2 + t_{z2}s\phi_1 \\ c\phi_1 c\phi_2 & c\phi_1 s\phi_2 & t_{z1}c\phi_2 - t_{z2}c\phi_1 \\ -c\phi_1 s\phi_2 & c\phi_1 c\phi_2 & -t_{x1}c\phi_2 + t_{x2}c\phi_1 \\ -s\phi_1 s\phi_2 & s\phi_1 c\phi_2 & -t_{x1}s\phi_2 + t_{x2}s\phi_1 \\ -c\phi_1 c\phi_2 & -c\phi_1 s\phi_2 & t_{x1}s\phi_2 + t_{z2}c\phi_1 \\ -s\phi_1 c\phi_2 & -s\phi_1 s\phi_2 & -t_{x1}c\phi_2 + t_{z2}s\phi_1 \end{bmatrix}. \quad (4)$$

In the sequel, we avoid to use the subscript for the current location, second view.

III. EKF FOR THE 1D TT

The elements of the TT are very useful providing information of position and orientation of a camera [7]. In our previous work [13], we have introduced the use of the 1D TT in visual servoing by taking the elements of the tensor directly in the control law. In this paper, we propose to make use of the information provided by the TT to estimate the camera motion according to the nonholonomic motion model (2). Once the state estimation is available, we can use such model to design a feedback control law.

Consider as the framework for the EKF the discrete version of the system that describes the kinematic motion of the fixed camera on the robot (2) together with a measurement model given by some elements of the TT. Due to the robot system (1) is a driftless one, we consider that the control input is the source of noise in the state, which is modeled as a Gaussian noise process $\mathbf{v}_k$. Besides, the noisy measurements can be modeled adding a Gaussian noise $\mathbf{w}_k$ to a nonlinear function h as follows

$$\begin{aligned} \mathbf{x}_{k+1} &= f\left(\mathbf{x}_k, \mathbf{u}_k + \mathbf{v}_k\right), \\ \mathbf{y}_k &= h\left(\mathbf{x}_k\right) + \mathbf{w}_k \end{aligned} \quad (5)$$

where $\mathbf{v} \sim N\left(0, \mathbf{M}\right), \mathbf{w} \sim N\left(0, \mathbf{N}\right)$, and $E\left[\mathrm{v}_i \mathrm{w}_j^T\right] = 0$. These are the continuous noise processes, with $\mathbf{M} \in \mathbb{R}^{2\times 2}$ the process noise covariance and $\mathbf{N} \in \mathbb{R}^{3\times 3}$ the measurement noise covariance. Under this framework an extended direct Kalman filter can be designed. The prediction equations to compute the estimates are

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= f\left(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k, \mathbf{0}, k\right), \\ \mathbf{P}_{k+1|k} &= \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{G}_k \mathbf{M}_k \mathbf{G}_k^T \end{aligned} \quad (6)$$

where the linear approximation $\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k \hat{\mathbf{x}}_{k|k} + \mathbf{G}_k \mathbf{u}_k$, $\hat{\mathbf{y}}_{k|k} = \mathbf{H}_k \hat{\mathbf{x}}_{k|k}$ of the nonlinear system is used.

$$\begin{aligned} \mathbf{F}_k &= \left.\frac{\partial f}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}, \mathbf{v}_k = 0} = \\ &= \begin{bmatrix} 1 & 0 & \ell\delta\omega_k s\phi_k - \delta\upsilon_k c\phi_k \\ 0 & 1 & -\ell\delta\omega_k c\phi_k - \delta\upsilon_k s\phi_k \\ 0 & 0 & 1 \end{bmatrix}_{\phi_k = \hat{\phi}_{k|k}}, \\ \mathbf{G}_k &= \left.\frac{\partial f}{\partial \mathbf{u}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}} = \begin{bmatrix} -s\phi_k & -\ell c\phi_k \\ c\phi_k & -\ell s\phi_k \\ 0 & 1 \end{bmatrix}_{\phi_k = \hat{\phi}_{k|k}}. \end{aligned}$$

The update equations to correct the estimates are

$$\begin{aligned} \mathbf{Q}_{k+1} &= \mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T + \mathbf{N}_{k+1}, \\ \mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T \mathbf{Q}_{k+1}^{-1}, \\ \hat{\mathbf{x}}_{k+1|k+1} &= \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\left[\mathbf{y}_{k+1} - h\left(\hat{\mathbf{x}}_{k+1|k}\right)\right], \\ \mathbf{P}_{k+1|k+1} &= \left[\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]\mathbf{P}_{k+1|k}. \end{aligned} \quad (7)$$

In these equations, $\hat{\mathbf{x}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$ represent an *a priori* estimate of the state and its covariance, and $\hat{\mathbf{x}}_{k+1|k+1}$, $\mathbf{P}_{k+1|k+1}$ provide an *a posteriori* estimated state for step k. It means that the a posteriori information utilizes feedback error in order to improve the state estimation. The required output matrix can be computed as follows

$$\mathbf{H}_{k+1} = \left.\frac{\partial h}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k+1|k}, \mathbf{w}_k = 0},$$

and the resulting matrix from the 1D TT will be specified in the next section.

A. Observability of the EKF with the TT as output

There are few works concerned about observability when an estimation based on Kalman filtering is applied. Some of them are [14] and [15]. To analyze our case, let us consider the linear approximation ($\mathbf{F}_k$, $\mathbf{G}_k$, $\mathbf{H}_k$) of the system (2) in the time k. Due to the matrices $\mathbf{F}_k$ and $\mathbf{H}_k$ are changing at each instant time, observability may be lost, which affects the convergence properties of the estimation algorithm. As is mention in [14], a system that is locally observable over every time segment $[t_k, t_{k+1}]$ in the interval $[t_0, t_{k+1}]$ will also be completely observable over the interval $[t_0, t_{k+1}]$. Then, the condition to accomplish for every k to ensure the system to be completely observable is

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

Because of the left superior identity matrix in $\mathbf{F}_k$, the rows of the observability matrix become linearly dependent. The only possibility of reaching the full rank condition is by building $\mathbf{H}_k$ of full space. It can be done by taking three elements of the TT as outputs. From (4), we can see that a suitable selection of measurements is T_{122}, T_{211}, T_{222}, in such a way that

$$\mathbf{H}_{k+1} = \begin{bmatrix} c\phi_1 c\hat{\phi} & c\phi_1 s\hat{\phi} & t_{z1}c\hat{\phi} - \hat{t}_z c\phi_1 \\ -c\phi_1 s\hat{\phi} & c\phi_1 c\hat{\phi} & -t_{x1}c\hat{\phi} + \hat{t}_x c\phi_1 \\ -s\phi_1 s\hat{\phi} & s\phi_1 c\hat{\phi} & t_{z1}c\hat{\phi} + \hat{t}_x s\phi_1 \end{bmatrix} \quad (8)$$

where $\hat{\phi} = \hat{\phi}_{k+1|k}$, $\hat{t}_x = -\hat{x}_{k+1|k}c\hat{\phi}_{k+1|k} - \hat{z}_{k+1|k}s\hat{\phi}_{k+1|k}$, $\hat{t}_z = \hat{x}_{k+1|k}s\hat{\phi}_{k+1|k} - \hat{z}_{k+1|k}c\hat{\phi}_{k+1|k}$, and t_{x1}, t_{z1} and ϕ_1 are constant values. The output matrix in (8) ensures local observability for every k even for some particular initial conditions, for instance $\phi_1 = 0$, in which case this matrix remains full rank due to the cosines in the main diagonal. It is worth noting that the measurement Jacobian requires to know the initial location $\mathbf{C}_1$, which can be computed using the localization scheme presented in [7]. It can also be used for the EKF initialization. As the scale factor is introduced in the initial location, the normalization of the values of the TT is not required.

IV. NONLINEAR CONTROLLER DESIGN

We present the design of a static state feedback linearizing control law that is able to drive the camera on the robot to a desired location without switching to a different control law. We solve a tracking problem for the nonlinear system (1) in order to correct x and z positions. At the same time, by tracking suitable reference signals, orientation correction is also ensured. This controller can be applied considering that the camera location is known, in our case, estimated by the EKF as described previously.

A. Input-Output Linearization

The goal is to drive the robot to the target location, which means to reach $(x_2, z_2, \phi_2) = (0, 0, 0)$. We will control the location of the point $\mathbf{c}$ instead of the robot reference frame. In this section we assume that this location is known and is given by the EKF estimation by using three elements of the TT as measurements. The advantage of our proposal with respect to previous works is that the state estimation makes possible to tackle the visual servoing problem as a trajectory tracking problem. The outputs to be controlled are the camera position coordinates $y_1 = x$, $y_2 = z$. Consequently, the orientation (ϕ) is left as a DOF which is automatically corrected by tracking suitable desired trajectories as will be proven in Section V.

To take the value of both outputs to zero in a smooth way, we design a tracking controller. Let us define the tracking errors as $e_1 = x - x^d$, $\quad e_2 = z - z^d$. Thus, the error system is given as

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} = \begin{bmatrix} -\sin\phi & -\ell\cos\phi \\ \cos\phi & -\ell\sin\phi \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} - \begin{bmatrix} \dot{x}^d \\ \dot{z}^d \end{bmatrix}. \quad (9)$$

This system has the form $\dot{\mathbf{e}} = \mathbf{D}(\phi, \ell)\,\mathbf{u} - \dot{\mathbf{y}}^d$, where $\mathbf{D}(\phi, \ell)$ corresponds to the decoupling matrix, whose inverse matrix is given in (10), and $\dot{\mathbf{y}}^d$ represents a known disturbance for the error dynamics. We can see that the control inputs appear in the first derivative of each output. Then the system (1) with outputs (x, z) has a vector relative degree $\{1,1\}$. Due to the sum of the indices of the system $(1 + 1)$ is less than the order of the system ($n = 3$) we have a first order zero dynamics, which will be analyzed in Section V.

$$\mathbf{D}^{-1}(\phi, \ell) = \frac{1}{\ell}\begin{bmatrix} -\ell\sin\phi & \ell\cos\phi \\ -\cos\phi & -\sin\phi \end{bmatrix}. \quad (10)$$

A static state feedback control law that achieves global stabilization of the system (9) has the form $\mathbf{u} = \mathbf{D}^{-1}(-\mathbf{k}\mathbf{e} + \dot{\mathbf{y}}_d)$, which is

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \frac{1}{\ell}\begin{bmatrix} -\ell\sin\phi & \ell\cos\phi \\ -\cos\phi & -\sin\phi \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (11)$$

where $u_1 = -k_1 e_1 + \dot{x}^d$, and $u_2 = -k_2 e_2 + \dot{z}^d$. The error behavior will be exponentially stable iff $k_1 > 0$, $k_2 > 0$.

Note that this input-output linearization via static feedback is only possible for the system (1) with $\ell \neq 0$. Otherwise, a singular decoupling matrix is obtained and a static feedback fails to solve the input-output linearization problem. However, the case of having the camera shifted from the robot rotation axis over the longitudinal axis is a common situation. Besides, the value of ℓ can be easily measured.

B. Desired Trajectories

The objective of reference tracking is to take the outputs to zero in a smooth way and consequently, the robot performs a smooth motion in a desired time. We propose the following references

$$\begin{aligned} z^d &= \tfrac{z(0)}{2}\left(1 + \cos\left(\tfrac{\pi}{\tau}t\right)\right),\ 0 \le t \le \tau \quad (12)\\ z^d &= 0, \qquad\qquad t > \tau \\ x^d &= \tfrac{x(0)}{z(0)^2}\left(z^d\right)^2, \qquad 0 \le t \le \tau \\ x^d &= 0, \qquad\qquad t > \tau \end{aligned}$$

where τ is the time to reach the target. Note that the robot always begins over the desired path and the controller has to maintain it tracking that path.

V. STABILITY ANALYSIS

The controller behavior is based on zeroing the defined outputs, therefore, when these outputs reach to zero the so-called *zero dynamics* in the robot system is achieved. Zero dynamics is described by a subset of the state space which makes the output to be identically zero [16]. In the particular case of the robot system (1) with outputs $y_1 = x$, $y_2 = z$, this set is given as follows

$$\begin{aligned} Z^* &= \left\{ \begin{bmatrix} x & z & \phi \end{bmatrix}^T \mid y_1 \equiv 0,\ y_2 \equiv 0 \right\} \\ &= \left\{ \begin{bmatrix} 0 & 0 & \phi \end{bmatrix}^T,\ \phi = \text{constant} \in \mathbb{R} \right\}. \end{aligned}$$

The constant value of ϕ is the result of the differential equation that characterize the zero dynamics, that in this case is

$$\dot{\phi} = -\frac{1}{\ell}\left(u_1\cos\phi + u_2\sin\phi\right) = 0$$

because $u_1 = 0$ and $u_2 = 0$ when $y_1 \equiv 0$ and $y_2 \equiv 0$. Thus, zero dynamics in this control system means that when x and z-coordinates of the robot are corrected, the orientation may be different to zero. Now we focus on proving how orientation correction is achieved.

Proposition 1. The control inputs in (11) with complete knowledge of the state $(x, z, \phi)^T$ provided by the 1D TT-based EKF estimation and using the reference signals in (12) drive the robot-system (1) to reach the location $(x = 0, z = 0, \phi = 0)^T$, i. e. orientation is also corrected.

Proof: It is clear the global exponential stability of the error system (9) with the inputs in (11), which drives the depth and lateral error to zero after τ seconds. It only remains to prove that the orientation is also zero after τ seconds. From the decomposition of the translational velocity vector ($\dot{x} = -v\sin\phi$, $\dot{z} = v\cos\phi$), we have that

$$\phi = \arctan\left(-\frac{\dot{x}}{\dot{z}}\right).$$

Let us be the parabolic relationship between Cartesian coordinates $x = \frac{x(0)}{z(0)^2} z^2$ according to the desired trajectories. Its corresponding time-derivative is $\dot{x} = 2\frac{x(0)}{z(0)^2} z\dot{z}$. Thus, the orientation angle can be computed as follows when the x and z-coordinates track the desired trajectories

$$\phi = \arctan\left(-2\frac{x(0)}{z(0)^2} z\right).$$

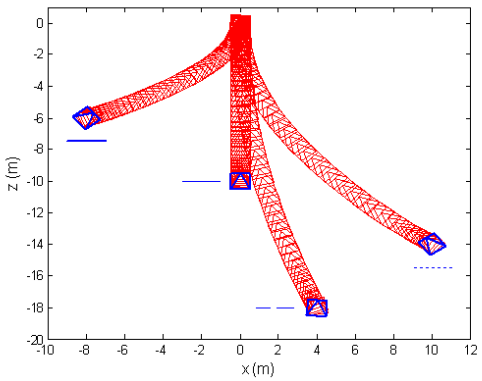
As have been mentioned, when the robot has followed the reference path and $t = \tau$ the position reaches zero ($x = 0$, $z = 0$), and consequently $\phi = \arctan(0) = 0$. This proves that the location $(x = 0, z = 0, \phi = 0)^T$ is always reached in τ seconds, even though the orientation is a DOF for the control system. ■

VI. RESULTS

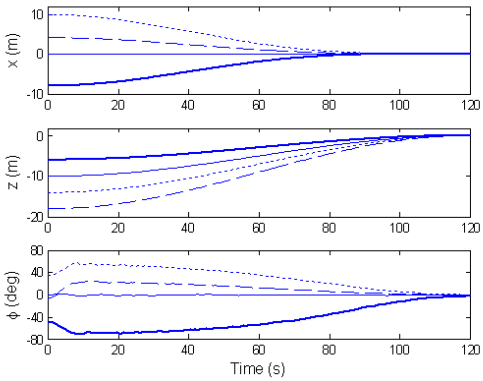
This section presents some simulation results that have been performed in Matlab by applying the control system as established in the Proposition 1. These results show that the main objective of driving the robot to a desired pose is always attained with good precision. For these simulations, we compute the 1D TT from point correspondences on virtual images using the approach described in [7]. The distance from the rotation axis to the camera position on the robot is set to $\ell = 10$ cm. For the controllers, the time to reach the target position (τ) is fixed to 120 s and the control gains are set to $k_1 = 1$, $k_2 = 1$. Related to the Kalman filtering, the sampling time T_s is set to 0.5 s. We have tuned the matrices $\mathbf{M}$ and $\mathbf{N}$ in such a way that we have more confidence in the model than in the measurements, otherwise, the robot motion may be uncertain. This is done by using small standard deviations in $\mathbf{M}$ and relatively large standard deviations in $\mathbf{N}$. The standard deviation of the noise affecting the control inputs (related to $\mathbf{M}$) is set to $\sigma_{\mathbf{v}} = [0.01$ m/s, 0.001 rad/s$]$. The standard deviation of the measurement noise process (related to $\mathbf{N}$) is set to $\sigma_{\mathbf{w}} = [20$ cm, 30 cm, 5 cm$]$. Finally, we suggest initial standard deviation for the state estimation errors $\mathbf{P} = diag(5^2$ cm, 10^2 cm, 1^2 deg$)$.

Fig. 3 shows the paths traced by the robot and the state variables evolution from four different initial locations. The thick solid line begins from (-8,-6,-50^o), the solid line from (0,-10,0^o), the long dashed line from (4,-18,-5^o) and the short dashed line from (10,-14,35^o). We can see that an initial rotation is performed automatically in order to align the robot with the parabolic path to be tracked. This kind of path is traced except for the case when there is no lateral error to correct. In Fig. 3(b) we can see that both outputs (x and z positions) are driven to zero in 120 s and the robot reaches the target with the right orientation.

An example of the good performance of the tracking control law is presented in Fig. 4(a). For the initial location (4,-18,-5^o), both tracking errors are maintained bounded with a bandwidth around 0.5 cm. These errors are used by the feedback control law and they are computed as the difference between estimated positions and their reference values. The smooth control inputs computed from the shown tracking



(a) Paths on the $x - z$ plane



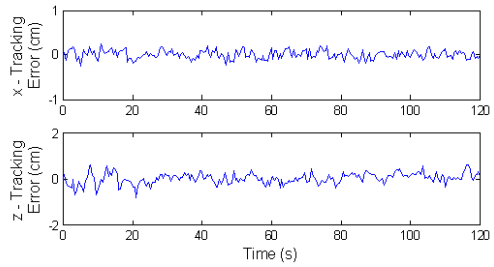
(b) State variables of the robot

Fig. 3. Resultant paths and state evolution.

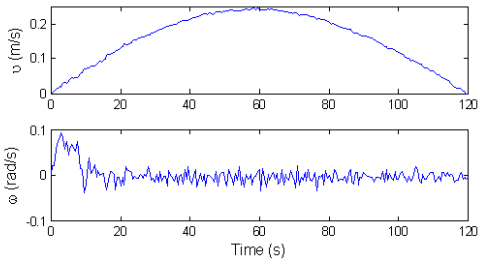
errors are presented in Fig. 4(b). Note that these signals start and end with zero value and they are always well defined.

In Fig. 5(a) the evolution of the elements of the TT that are taken as measurements for the EKF is plotted for the initial location (4,-18,-5^o). From these noise measurements the state estimation errors of Fig. 5(b) are determined. The first plot presents the evolution of the *trace* of the state estimation error covariance matrix ($\mathbf{P}$). It shows that the EKF is converging because the standard deviations of the state estimation error is decreasing at each step. Convergence is expected because the discrete linear approximation of the nonlinear system remains observable along the navigation. However, the plot *trace*($\mathbf{P}$) could increase during short periods of time in some cases due to the nonlinearity of the system. The next three plots in Fig. 5(b) show that the estimation errors are maintained within the 1σ confidence bounds. The errors are computed from the truth state (available in simulation) and the estimated state.

Finally, Table I shows that the target location (0,0,0^o) is reached with good precision, in the order of centimeters for position and less than one degree for rotation in most of the cases. The final rotation larger than one degree (first column) corresponds to a large initial angle (-50^o) and additionally, the initial depth is less than the initial lateral error, which forces a large rotation to reach the target. However, in this case the depth is corrected better than in any other case.



(a) Tracking Errors



(b) Control Inputs

Fig. 4. Tracking errors and control inputs for the initial location $(4, -18, -5^o)$.

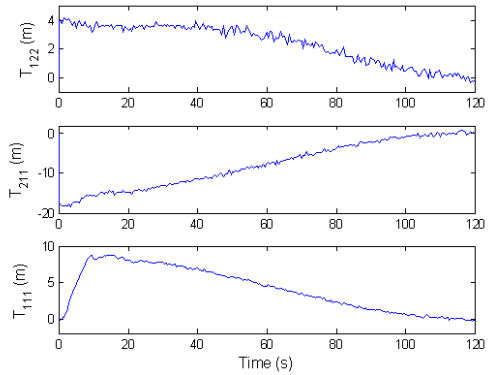
TABLE I

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

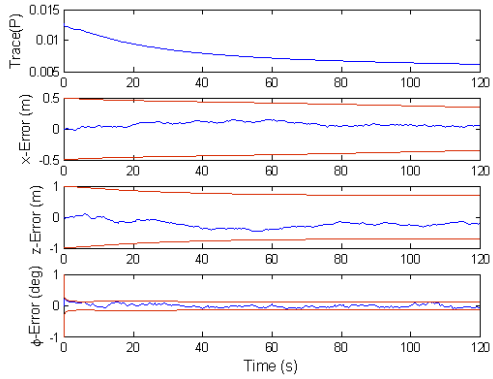
	(-8,-6,-50) (m,m,^{o})	(0,-10,0) (m,m,^{o})	(4,-18,-5) (m,m,^{o})	(10,-14,35) (m,m,^{o})
x_{final} (cm)	-0.69	0.85	0.62	0.18
z_{final} (cm)	-0.5	-1.23	-1.84	-1.33
ϕ_{final} (^{o})	-2.46	-0.84	-0.84	0.84

VII. CONCLUSIONS

In this paper we have presented a new pose-estimation-based visual servoing scheme to perform visual control for differential-drive robots using a teach-by-showing strategy. This scheme allows knowing the real world path performed by the robot without the computational load introduced by position-based approaches. It recovers the camera location (position and orientation) using an Extended Kalman Filter (EKF) algorithm with the 1D trifocal tensor (TT) as measurements. An observability analysis of the discrete linear approximation is discussed. We solve a tracking problem for the lateral and longitudinal robot coordinates with a static state-estimated feedback control law. The desired trajectories to be tracked ensure total correction of lateral error, depth and orientation using a unique control law, even though the orientation is a DOF in the control system. The performance of our approach is proven via simulations.



(a) Noisy measurements, elements of the TT



(b) State estimation errors and 1σ uncertainty bounds

Fig. 5. Performance of the EKF for the initial location $(4, -18, 5^o)$.

REFERENCES

[1] G. L. Mariottini, G. Oriolo, and D. Prattichizzo. Image-based visual servoing for nonholonomic mobile robots using epipolar geometry. *IEEE Transactions on Robotics*, 23(1):87–100, 2007.

[2] H. M. Becerra and C. Sagues. A sliding mode control law for epipolar visual servoing of differential-drive robots. In *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, pages 3058–3063, 2008.

[3] S. Benhimane and E. Malis. Homography-based 2D visual servoing. In *IEEE Int. Conf. on Robotics and Autom.*, pages 2397–2402, 2006.

[4] G. López-Nicolás, C. Sagüés, and J.J. Guerrero. Homography-based visual control of nonholonomic vehicles. In *IEEE Int. Conf. on Robotics and Automation*, pages 1703–1708, 2007.

[5] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

[6] F. Dellaert and A. W. Stroupe. Linear 2D localization and mapping for single and multiple robot scenarios. In *IEEE Int. Conf. on Robotics and Automation*, pages 688–694, 2002.

[7] J.J. Guerrero, A.C. Murillo, and C. Sagüés. Localization and matching using the planar trifocal tensor with bearing-only data. *IEEE Transactions on Robotics*, 24(2):494–501, 2008.

[8] G. Lopez-Nicolas. *Visual Control of Mobile Robots Through Multiple View Geometry*. PhD thesis, DIIS, Universidad de Zaragoza, Spain, June 2008.

[9] K. Hashimoto and H. Kimura. Visual servoing with nonlinear observer. In *IEEE Int. Conf. on Robotics and Automation*, pages 484–489, 1995.

[10] K. Hashimoto and T. Noritsugu. Visual servoing with linearized observer. In *IEEE Int. Conf. on Robotics and Automation*, pages 263–268, 1999.

[11] A. De Luca, G. Oriolo, and P. Robuffo. On-line estimation of feature depth for image-based visual servoing schemes. In *IEEE Int. Conf. on Robotics and Automation*, pages 2823–2828, 2007.

[12] Y.K. Yu, K.H. Wong, M.M.Y Chang, and S.H. Or. Recursive camera-motion estimation with the trifocal tensor. *IEEE Trans. on Syst., Man, and Cyb. - Part B: Cybernetics*, 36(5):1081–1090, 2006.

[13] H. M. Becerra and C. Sagues. A novel 1D trifocal tensor-based control for differential drive robots. In *IEEE Int. Conf. on Robotics and Automation*, pages 1104–1109, 2009.

[14] M. Bryson and S. Sukkarieh. Observability analysis and active control for airbone SLAM. *IEEE Trans. on Aerospace and Electronic Systems*, 44(1):261–280, 2008.

[15] 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.

[16] S. Sastry. *Nonlinear Systems: Analysis, Stability and Control*. Springer, New York, 1999.

IEEE

The 2009 IEEE/RSJ International Conference on Robots and Intelligent Systems

IROS 2009

October 11 - 15, 2009

St. Louis, USA

IEEE Catalog Number: CFP09IRO-DVD
ISBN: 978-1-4244-3804-4
Library of Congress: 2009900280

Foreword

We offer a warm welcome to all of the attendees of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2009). This is a conference that over the years has established itself as a truly global enterprise of major technical importance. Our conference venue is St. Louis, Missouri located on the Mississippi River. St. Louis is the historical gateway to the West during the early years of growth and expansion in the United States. Crossing the Mississippi was the demarcation from east to west for early pioneers embarking on new hopes and horizons in the unsettled regions to the west. Building on this history, the conference theme is "Exploring New Horizons in Intelligent Robots and Systems." Indeed, most think of IROS as a conference where an attractive blend of recent results encompassing basic through applied research can be viewed. We believe this year's technical program offers just that.

Even under the current conditions of global economic stress, IROS 2009 continues the tradition of wide interest and participation from all over the world. 1599 papers were submitted from 53 countries. After a rigorous review process, 936 papers were accepted for presentation at the conference. The technical program consists of three plenary talks, 192 technical sessions in 16 tracks, 18 tutorials/workshops, and 13 video presentations. Session technical topics cover the full gamut from emerging areas of interest to more traditional subject areas within intelligent systems and robots. Some of the largest contributions have occurred in areas such as robot audition, biological inspiration, haptics, human-robot interactions, humanoids, medical robotics, and rehabilitation robotics. Like all recent IROS's, you face a challenge to navigate amongst the 16 tracks to hear the papers you are most interested. We have prepared this digest to assist you as much as possible in this challenge. We sincerely hope that you enjoy the conference, and that it provides you with useful and important information about current research.

We would like to thank all of the members of the IROS 2009 Organizing Committee for their contributions. It takes a dedicated team to be able to accomplish a conference of this size, and we are grateful to each of the committee members who have dedicated so much of their time and effort to IROS 2009. Finally, our thanks to all of the authors and participants who provide the intellectual exchanges that are the essence of IROS.

Ning Xi, General Chair

William R. Hamel, Program Chair