

Visual Control through the Trifocal Tensor for Nonholonomic Robots

G. López-Nicolás, J.J. Guerrero and C. Sagüés

DIIS - I3A, Universidad de Zaragoza, C/ María de Luna 1, E-50018 Zaragoza, Spain
{gonlopez, jguerrer, csagues}@unizar.es

Abstract

We present a new vision-based control approach which drives autonomously a nonholonomic vehicle to a target location. The vision system is a camera fixed on the vehicle and the target location is defined by an image taken previously in that location. The control scheme is based on the trifocal tensor model, which is computed from feature correspondences in calibrated retina across three views: the initial, current and target images. The contribution is a trifocal-based control law defined by an exact input-output linearization of the trifocal tensor model. The desired evolution of the system towards the target is directly defined in terms of the trifocal tensor elements by means of sinusoidal functions without needing metric or additional information from the environment. The trifocal tensor presents important advantages for visual control purposes because it is more robust than two view geometry as it includes the information of a third view and, contrary to the epipolar geometry, short baseline is not a problem. Simulations show the performance of the approach, which has been tested with image noise and calibration errors.

1 Introduction

Many contributions have been presented to solve the problem of visual control [1], [2], but still it is a growing field of research. In the framework considered here, the only sensor is a fixed monocular system mounted on a vehicle. We also consider the vehicle with nonholonomic motion constraints. The desired location is defined by the target image taken previously at that location. We present a visual control approach which relies on image information. The information extracted from the initial, current and target images (where the current image is taken by the vehicle's camera at the actual location) is used by the control law to drive the vehicle to the target.

A direct way to face the problem of extracting information from the images is to rely on landmarks or particular features, which are extracted and tracked, directly

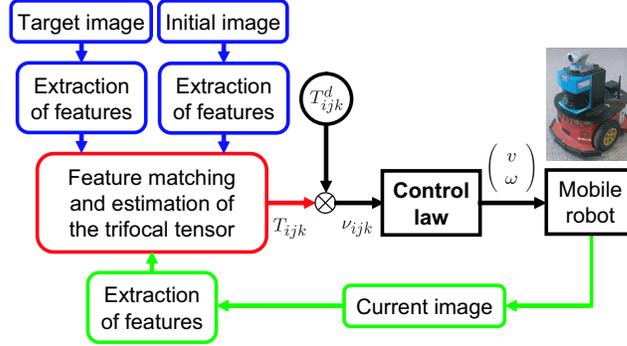


Fig. 1. Diagram of the trifocal tensor-based control loop.

included in the control scheme [3], [4]. In favour of robustness, a good choice is to process the image information through a geometric model relating the acquired images. In this case, there is less chance that spurious correspondences from the set of matches could reduce the control performance. An early work [5] based on the epipolar geometry, where the image information relies on the epipoles, has been followed by others [6] [7] [8]. Nevertheless, the epipolar geometry has main drawbacks, one is that the fundamental matrix is ill-conditioned with short baseline and therefore, an epipolar-based control eventually becomes unstable. Another drawback is that the epipolar geometry becomes degenerated with planar scenes, which are common in man-made environments. A natural way to overcome these drawbacks is using the homography defined by a plane of the scene. This geometric model is robust and well defined with short baseline. A well known hybrid method is [9], [10], more examples of visual control based on homographies are [11], [12], [13], [14], [15], [16]. However, the performance of a homography-based control can be affected if there is no dominant plane in the scene. This problem can be solved through virtual planes [10]. Nevertheless, estimations based on virtual planes with wide baseline are not robust to mismatches, noise or occlusions.

Here we propose a new visual control based on the trifocal tensor. This tensor encapsulates the intrinsic geometry between three views and it is independent of the observed scene [17]. This geometric model has several advantages: It is more robust than the two view geometry models as it involves the information given by a third view, and the set of correspondences obtained is more robust to outliers. Additionally, the trifocal tensor is still useful with short baseline, whereas the epipolar geometry fails. The problem of localization has been discussed in [18], [19], [20] through the 1D trifocal tensor and with the 2D trifocal tensor [21], [22] or using stereo vision [23]. The 1D trifocal tensor has been used for visual control [24], but we have chosen the 2D trifocal tensor rather than 1D to take advantage of all the information available in the 2D images. In our approach, rather than decomposing the trifocal tensor to obtain pose information we design a new method which performs the control directly on the trifocal tensor elements. This methodology is inspired by the same background idea as the homography-based control design of [16]. The control law is obtained by an exact input-output linearization of the system, where

the desired trajectories of the tensor entries are defined in order to reach the desired location. In our approach, the trifocal tensor is computed from three views: the initial, current and target images. So, at the start, the initial and current images are the same and, as the vehicle moves towards the target, the current and target images get similar. A diagram of the visual control law presented is shown in Fig. 1. The input of the control are the initial, current and target images and the output are the velocities that lead the vehicle to the target location.

The paper is organized as follows. The trifocal tensor within our framework is described in Section 2. The vision-based control law is developed by an input-output linearization through the trifocal tensor in Section 3. Stability analysis and experimental validation are given in Sections 4 and 5, respectively. Conclusions are discussed in Section 6.

Notation: In some equations we use for readability the notation $s_\phi = \sin \phi$ and $c_\phi = \cos \phi$. We denote the trifocal tensor deduced theoretically with \mathbf{T}' . Later, we normalize the trifocal tensor to a fixed scale and it is denoted with \mathbf{T} .

2 The Trifocal Tensor

Three perspective images can be geometrically linked by the trifocal tensor. This tensor only depends on the relative locations between the three views and the internal calibration parameters of the cameras. Therefore, the trifocal tensor is independent of the observed scene. The geometric relation between three views given by the trifocal tensor is similar to the geometric relation given by the fundamental matrix between two views. Some publications about the trifocal tensor are [25], [21], [26], [17], [27].

The trifocal tensor can be deduced in several ways, here we obtain the geometric expression starting from the camera locations [17]. A point in the 3D space can be represented as $\mathbf{X} = (X, Y, Z, 1)^T$. On the other hand, a point can be represented in the three images with homogeneous coordinates in a calibrated retina as

$$\mathbf{u} = (u^1, u^2, 1)^T, \quad \mathbf{v} = (v^1, v^2, 1)^T, \quad \mathbf{w} = (w^1, w^2, 1)^T. \quad (1)$$

Let us suppose that the three images are taken with a calibrated camera, represented by the pinhole model. The projection of a 3D point \mathbf{X} in the three images gives:

$$\lambda_1 \mathbf{u} = \mathbf{P}_1 \mathbf{X}, \quad \lambda_2 \mathbf{v} = \mathbf{P}_2 \mathbf{X}, \quad \lambda_3 \mathbf{w} = \mathbf{P}_3 \mathbf{X}, \quad (2)$$

where λ_1 , λ_2 and λ_3 are scale factors. The projection matrices of the calibrated

cameras in a common reference system are

$$\mathbf{P}_1 = [\mathbf{R}_1 | \mathbf{t}_1] , \quad \mathbf{P}_2 = [\mathbf{R}_2 | \mathbf{t}_2] , \quad \mathbf{P}_3 = [\mathbf{R}_3 | \mathbf{t}_3] , \quad (3)$$

where the camera rotations and translations are defined as follows. The global reference system is depicted in Fig. 2 with the origin attached to the third camera. The locations of the cameras in the global reference are $\mathbf{C}_1 = (x_1, y_1, z_1)$, $\mathbf{C}_2 = (x_2, y_2, z_2)$ and $\mathbf{C}_3 = (0, 0, 0)$ with their respective orientations ϕ_1, ϕ_2 and ϕ_3 (with $\phi_3 = 0$). Given that we consider planar motion we have $y_1 = 0, y_2 = 0$. Then, the camera rotation matrices are

$$\mathbf{R}_1 = \begin{bmatrix} c_{\phi_1} & 0 & s_{\phi_1} \\ 0 & 1 & 0 \\ -s_{\phi_1} & 0 & c_{\phi_1} \end{bmatrix} , \quad \mathbf{R}_2 = \begin{bmatrix} c_{\phi_2} & 0 & s_{\phi_2} \\ 0 & 1 & 0 \\ -s_{\phi_2} & 0 & c_{\phi_2} \end{bmatrix} , \quad (4)$$

and $\mathbf{R}_3 = \mathbf{I}_3$. The camera translations in local coordinate systems (Fig. 2) are expressed as

$$\mathbf{t}_1 = \begin{pmatrix} t_{x1} \\ t_{y1} \\ t_{z1} \end{pmatrix} = -\mathbf{R}_1 \mathbf{C}_1 = \begin{pmatrix} -x_1 \cos \phi_1 - z_1 \sin \phi_1 \\ 0 \\ x_1 \sin \phi_1 - z_1 \cos \phi_1 \end{pmatrix} . \quad (5)$$

$$\mathbf{t}_2 = \begin{pmatrix} t_{x2} \\ t_{y2} \\ t_{z2} \end{pmatrix} = -\mathbf{R}_2 \mathbf{C}_2 = \begin{pmatrix} -x_2 \cos \phi_2 - z_2 \sin \phi_2 \\ 0 \\ x_2 \sin \phi_2 - z_2 \cos \phi_2 \end{pmatrix} . \quad (6)$$

and $\mathbf{t}_3 = -\mathbf{R}_3 \mathbf{C}_3 = (0, 0, 0)^T$. The expressions given in (2) can be gathered in matrix form $\mathbf{A} \in \mathbb{R}^{9 \times 7}$ as

$$\mathbf{A} \begin{pmatrix} \mathbf{X} \\ -\lambda_1 \\ -\lambda_2 \\ -\lambda_3 \end{pmatrix} = 0 , \quad \text{with } \mathbf{A} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{u} & 0 & 0 \\ \mathbf{P}_2 & 0 & \mathbf{v} & 0 \\ \mathbf{P}_3 & 0 & 0 & \mathbf{w} \end{bmatrix} . \quad (7)$$

The previous equation must hold for any point of the scene and therefore the maximum rank of matrix \mathbf{A} is six. Thus, any 7×7 minor of \mathbf{A} has determinant zero. This gives the trilinear relations that define the trifocal tensor [17]. Two rows have

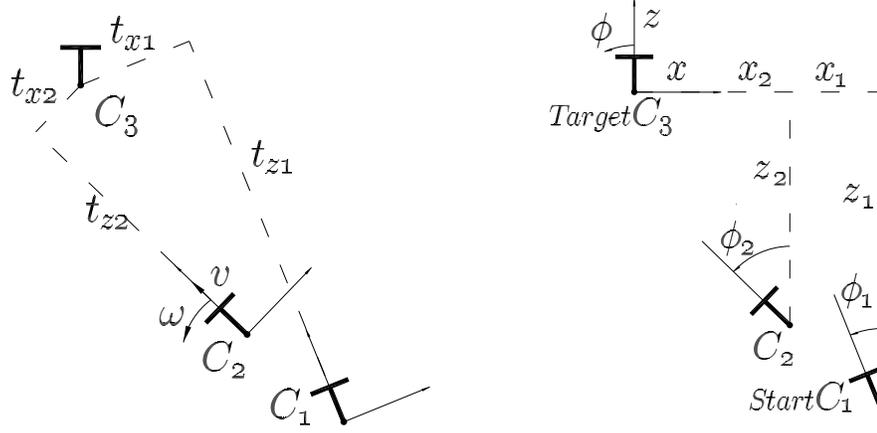


Fig. 2. Local coordinate systems (left) and global reference system (right) defined on the optical center of the cameras C_1 , C_2 and C_3 .

to be removed to develop each minor. We take the three rows given by the third camera P_3 (which is in canonical form $[I|0]$), and two rows from each of P_1 and P_2 . Thus, several determinants are obtained depending on the rows selected. From these determinants the trifocal tensor elements are given up to scale as follows

$$\begin{aligned}
T'_{111} &= -t_{x1} \cos \phi_2 + t_{x2} \cos \phi_1 \\
T'_{113} &= t_{x1} \sin \phi_2 + t_{z2} \cos \phi_1 \\
T'_{131} &= -t_{z1} \cos \phi_2 - t_{x2} \sin \phi_1 \\
T'_{133} &= t_{z1} \sin \phi_2 - t_{z2} \sin \phi_1 \\
T'_{212} &= -t_{x1} \\
T'_{221} &= t_{x2} \\
T'_{223} &= t_{z2} \\
T'_{232} &= -t_{z1} \\
T'_{311} &= -t_{x1} \sin \phi_2 + t_{x2} \sin \phi_1 \\
T'_{313} &= -t_{x1} \cos \phi_2 + t_{z2} \sin \phi_1 \\
T'_{331} &= -t_{z1} \sin \phi_2 + t_{x2} \cos \phi_1 \\
T'_{333} &= -t_{z1} \cos \phi_2 + t_{z2} \cos \phi_1
\end{aligned} \tag{8}$$

The other elements of the trifocal tensor are zero as a result of the planar motion constraint.

The image points are used in calibrated coordinates. In this case and considering

planar motion, the next expressions relating elements of the trifocal tensor hold

$$\begin{aligned} -T'_{111} + T'_{133} + T'_{313} + T'_{331} &= 0, \\ T'_{113} + T'_{131} + T'_{311} - T'_{333} &= 0. \end{aligned} \tag{9}$$

In our visual control framework, (x_1, z_1, ϕ_1) is the initial location of the vehicle, $(x_3, z_3, \phi_3) = (0, 0, 0)$ is the target location and $(x_2(t), z_2(t), \phi_2(t))$ is the current location that varies as the vehicle moves. The goal is to drive the vehicle to the target location. Therefore the objective of the control law is to drive the vehicle to $(x_2, z_2, \phi_2) = (0, 0, 0)$. Therefore, when the vehicle is in the desired target location we have the following values for the trifocal tensor elements

$$T'_{111}{}^d = T'_{212}{}^d = T'_{313}{}^d = -t_{x1}, \tag{10}$$

$$T'_{131}{}^d = T'_{232}{}^d = T'_{333}{}^d = -t_{z1}, \tag{11}$$

$$T'_{113}{}^d = T'_{133}{}^d = T'_{221}{}^d = T'_{223}{}^d = T'_{311}{}^d = T'_{331}{}^d = 0. \tag{12}$$

The trifocal tensor can be computed from feature correspondences across the three views. The image features can be points, lines or a combination of both. The trifocal tensor is defined by 27 parameters, and 26 up to scale. Each triplet of corresponding image points gives 4 equations linearly independent (this can be derived from the trilinear relations expanded from (7)). Therefore, with a minimum set of 7 correspondences of points the trifocal tensor can be computed. We can also include the calibration constraints (9) in the algorithm and then, a minimum of 6 correspondences of points are required. Moreover, considering the planar motion constraint, the trifocal tensor is defined by 12 parameters, 11 up to scale, and it can be computed from a minimum set of 3 correspondences of points. In practice, we consider the presence of image noise and outliers, and therefore the RANSAC method is used [28]. This robust method proceeds by repeatedly generating hypothesis from a minimal set of points. The probability of outliers in the samples is reduced with their smaller size, and the method performance is improved. Thus, it is an advantage to reduce the minimal set of points from 7 in the general tensor to 3 in the case of planar motion. Details for the automatic computation of the trifocal tensor can be found in [17], [26], [29].

3 Control Law

In this section the control law of the vision-based approach is presented. An overview of the visual control law has been depicted in the diagram of Fig. 1. Image features are extracted from the initial and target images, and they are matched with

the features extracted from the current image. Next, the trifocal tensor $\mathbf{T}_{ijk}(t)$ is computed from the feature correspondences. The input of the control is defined by $\boldsymbol{\nu}_{ijk}(t)$ which depends on the trifocal tensor and its desired value $\mathbf{T}_{ijk}^d(t)$. The control law gives the velocities required to drive the vehicle to the target location. This section is organized as follows. First, we present the model of the system relying on the trifocal tensor as output (3.1). Next, the control of the nonlinear system is transformed to a tracking problem where the control is directly performed on the trifocal tensor elements (3.2). Finally, the desired trajectories of the trifocal tensor elements are defined by means of time-varying functions (3.3).

3.1 Model of the System

The nonlinear system of the mobile platform with the vision system to be controlled is described by

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = \mathbf{T}_{ijk}(\mathbf{x}) \end{cases} \quad (13)$$

where $\mathbf{u} = (v, \omega)^T$ denotes the input vector, which includes the translational ($v(t)$) and rotational ($\omega(t)$) velocities of the vehicle, and \mathbf{y} denotes the output vector with $\mathbf{T}_{ijk}(\mathbf{x}, t)$ the trifocal tensor elements. The configuration of the robot system is given by $\mathbf{x} = (x, z, \phi)^T$. The particular nonholonomic differential kinematics of the vehicle $\mathbf{f}(\mathbf{x}, \mathbf{u})$ expressed in state space form as a function of the translation and rotation velocities of the robot (v, ω) is as follows

$$\begin{pmatrix} \dot{x} \\ \dot{z} \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} -\sin \phi \\ \cos \phi \\ 0 \end{pmatrix} v + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \omega . \quad (14)$$

3.2 Input-Output Linearization through the Trifocal Tensor

In the model of the system (13) the output \mathbf{y} (the trifocal tensor) is only indirectly related to the input \mathbf{u} (the robot velocities). Therefore, it is not easy to see how the input \mathbf{u} can be designed to control the desired evolution of \mathbf{y} . We propose to carry out an exact input-output linearization to find a direct relation between \mathbf{y} and \mathbf{u} [31, 32]. So, we transform the problem of nonlinear control into a tracking problem where the desired evolutions of the trifocal tensor elements are defined.

As previously said, the trifocal tensor is computed from three images, the image

at the target location $(x_3, z_3, \phi_3) = (0, 0, 0)$, the image taken at the initial location $(x_1, z_1, \phi_1) = \text{constant}$ and the image at the current location, which varies as the vehicle moves $(x_2(t), z_2(t), \phi_2(t))$. The derivatives of all the trifocal tensor elements with respect to time are now obtained. The derivatives of two elements are shown in detail as example. From (8) we have for the first element:

$$T'_{111} = -t_{x1}c_{\phi_2} + t_{x2}c_{\phi_1} = -t_{x1}c_{\phi_2} - x_2c_{\phi_1}c_{\phi_2} - z_2c_{\phi_1}s_{\phi_2} \quad (15)$$

Its derivative with respect to time gives

$$\dot{T}'_{111} = t_{x1}\dot{\phi}_2s_{\phi_2} + c_{\phi_1}(-\dot{x}_2c_{\phi_2} + x_2\dot{\phi}_2s_{\phi_2} - \dot{z}_2s_{\phi_2} - z_2\dot{\phi}_2c_{\phi_2}) \quad (16)$$

Introducing the kinematic model (14) and (5), (6) we obtain

$$\begin{aligned} \dot{T}'_{111} &= t_{x1}\omega s_{\phi_2} + c_{\phi_1}(v(s_{\phi_2}c_{\phi_2} - s_{\phi_2}c_{\phi_2}) + \omega(x_2s_{\phi_2} - z_2c_{\phi_2})) \\ \dot{T}'_{111} &= (t_{x1}s_{\phi_2} + t_{z2}c_{\phi_1})\omega. \end{aligned} \quad (17)$$

The derivative of another trifocal tensor element is given as example following the same steps,

$$\begin{aligned} T'_{113} &= t_{x1}s_{\phi_2} + t_{z2}c_{\phi_1} = t_{x1}s_{\phi_2} + x_2c_{\phi_1}s_{\phi_2} - z_2c_{\phi_1}c_{\phi_2} \\ \dot{T}'_{113} &= t_{x1}\dot{\phi}_2c_{\phi_2} + c_{\phi_1}(\dot{x}_2s_{\phi_2} + x_2\dot{\phi}_2c_{\phi_2} - \dot{z}_2c_{\phi_2} + z_2\dot{\phi}_2s_{\phi_2}) \\ \dot{T}'_{113} &= t_{x1}\omega c_{\phi_2} + c_{\phi_1}(-v(s_{\phi_2}^2 + c_{\phi_2}^2) + \omega(x_2c_{\phi_2} + z_2s_{\phi_2})) \\ \dot{T}'_{113} &= -c_{\phi_1}v + (t_{x1}c_{\phi_2} - t_{z2}c_{\phi_1})\omega. \end{aligned} \quad (18)$$

Note that no metric information, depth estimation or trifocal tensor decomposition is used in our approach. Therefore, the unknown scale of the trifocal tensor is a key point and we need to define a common scale for the control law during the navigation. For this reason we normalize the trifocal tensor in order to get a fixed scale with T'_N . We need to define the value of T'_N guaranteeing no singularity, and then we have selected constant elements of (8) not equal to zero as follows

$$T_{ijk} = \frac{T'_{ijk}}{T'_N}, \quad (19)$$

$$\text{with } T'_N = \text{sign}(T'_{232})\sqrt{(T'_{212})^2 + (T'_{232})^2}, \quad (20)$$

where T'_{ijk} are the trifocal tensor elements deduced theoretically from the camera locations and T_{ijk} are the normalized trifocal tensor elements. The value of T'_N only

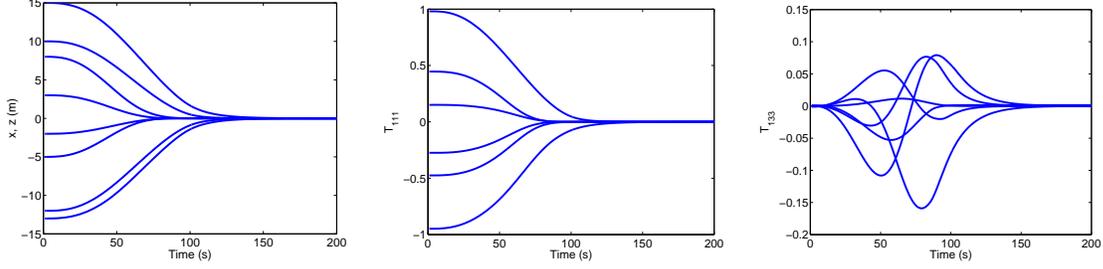


Fig. 3. Examples of several robot motions starting from different locations showing the evolution of x and z coordinates (left). For these different motions, we show the evolution of T_{111} , selected for the control (middle) and T_{133} , not used in the control (right).

is zero if the vehicle is initially in the target position ($x_1 = z_1 = 0$). Therefore we always have $T'_N \neq 0$ when the control is required to reach the target. Note also that this scale factor does not affect the time derivatives since it is a constant value. So, we finally simplify and normalize the derivatives of the elements by using (8) and (19). The same procedure is followed for all the trifocal tensor elements and the following expressions are obtained:

$$\begin{aligned}
 \dot{T}_{111} &= T_{113} \omega \\
 \dot{T}_{113} &= \frac{-\cos \phi_1}{T'_N} v - T_{111} \omega \\
 \dot{T}_{131} &= T_{133} \omega \\
 \dot{T}_{133} &= \frac{\sin \phi_1}{T'_N} v - T_{131} \omega \\
 \dot{T}_{221} &= T_{223} \omega \\
 \dot{T}_{223} &= -v - T_{221} \omega \\
 \dot{T}_{311} &= T_{313} \omega \\
 \dot{T}_{313} &= \frac{-\sin \phi_1}{T'_N} v - T_{311} \omega \\
 \dot{T}_{331} &= T_{333} \omega \\
 \dot{T}_{333} &= \frac{-\cos \phi_1}{T'_N} v - T_{331} \omega
 \end{aligned} \tag{21}$$

After the first derivative we have already obtained a linear relation between the system input and output. From the derivative of the available trifocal tensor elements (21), six of them have been selected for the control of the system. Two velocities of the system are controlled and, in principle, two elements of the trifocal tensor would be enough. However a two-element based control would fail in solving the control task from some location of the workspace (i.e. there are locations of the workspace in which the robot is not controllable with these particular elements). Additionally, with the selection of more elements we can guarantee no singularity of the control as explained later. The selection of the elements has been studied experimentally and the best elements to work with have been found to be T_{111} , T_{113} ,

T_{131} , T_{313} , T_{331} and T_{333} . Plots of one of the elements selected for the control and other discarded are given as example (Fig. 3). Several simulations are carried out in which the robot starts from different initial locations and navigates to the target location. The evolution of the robot coordinates (x, z) is shown in Fig. 3(left). The evolution of one of the elements selected for the control (T_{111}) and one not used in the control (T_{133}) are shown in Fig. 3(middle) and (right), respectively. It can be seen that the different plots of the selected element are simpler than the one not selected and, additionally, they have a straightforward relation with respect to the evolution of the robot coordinates. Given that no metric information is used, this criterion for the selection simplifies subsequently the design of the desired trajectories of the control inputs. With a similar analysis for all the elements of the trifocal tensor, six have been selected for the control, and we have

$$\dot{\mathbf{T}}_{ijk} = \begin{pmatrix} \dot{T}_{111} \\ \dot{T}_{113} \\ \dot{T}_{131} \\ \dot{T}_{313} \\ \dot{T}_{331} \\ \dot{T}_{333} \end{pmatrix} = \mathbf{L} \begin{pmatrix} v \\ \omega \end{pmatrix}, \quad (22)$$

where $\dot{\mathbf{T}}_{ijk} \in \mathbb{R}^6$ and the interaction matrix $\mathbf{L} \in \mathbb{R}^{6 \times 2}$ is

$$\mathbf{L} = \begin{bmatrix} 0 & T_{113} \\ \frac{-\cos \phi_1}{T'_N} & -T_{111} \\ 0 & T_{133} \\ \frac{-\sin \phi_1}{T'_N} & -T_{311} \\ 0 & T_{333} \\ \frac{-\cos \phi_1}{T'_N} & -T_{331} \end{bmatrix}. \quad (23)$$

The trifocal tensor elements have been already computed and normalized and the value of ϕ_1 is solved in Section 3.3. It turns out that the value of T'_N is also required in the \mathbf{L} matrix. However, T'_N is different with each trifocal tensor and different values of T'_N in the \mathbf{L} matrix would introduce different scales for the velocity. Thus, given that T'_N appears as a parameter of the input (v) , we have transferred its value as a constant gain of the control. This selection of T'_N in the \mathbf{L} matrix is tested in

Section 5. Solving (22) for the control outputs we have

$$\begin{pmatrix} v \\ \omega \end{pmatrix} = \mathbf{L}^+ \boldsymbol{\nu}_{ijk} = \mathbf{L}^+ \begin{pmatrix} \nu_{111} \\ \nu_{113} \\ \nu_{131} \\ \nu_{313} \\ \nu_{331} \\ \nu_{333} \end{pmatrix}, \quad (24)$$

where $\mathbf{L}^+ \in \mathbb{R}^{2 \times 6}$ is the pseudo-inverse of \mathbf{L} and $\boldsymbol{\nu}_{ijk} \in \mathbb{R}^6$ are the new inputs defined as

$$\boldsymbol{\nu}_{ijk} = \dot{T}_{ijk}^d - k(T_{ijk} - T_{ijk}^d), \quad (25)$$

$k > 0$ being constant gain and $(ijk) = \{111, 113, 131, 313, 331, 333\}$. The tracking error is defined as $\mathbf{e} = \mathbf{T}_{ijk} - \mathbf{T}_{ijk}^d$ with T_{ijk}^d the desired evolution of the trifocal tensor elements defined in Section 3.3.

The interaction matrix of the control (23) is not squared and we need to compute the left pseudo-inverse matrix \mathbf{L}^+

$$\mathbf{L}^+ = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T. \quad (26)$$

We need to guarantee that there is no singularity and therefore that $\det(\mathbf{L}^T \mathbf{L}) \neq 0$. The expression of this determinant results in

$$\begin{aligned} \det(\mathbf{L}^T \mathbf{L}) &= \frac{(1 + \cos^2 \phi_1)}{T_N'^2} (T_{111}^2 + T_{113}^2 + T_{133}^2 + T_{311}^2 + T_{331}^2 + T_{333}^2) \\ &\quad - \frac{1}{T_N'^2} (T_{111} \cos \phi_1 + T_{331} \cos \phi_1 + T_{311} \sin \phi_1)^2. \end{aligned} \quad (27)$$

This determinant only is zero if the initial position is on the target position, and so $(x_1, z_1) = (0, 0)$ and $(x_2, z_2) = (0, 0)$, which is not our case. Therefore, the control matrix \mathbf{L} is never singular. Additionally, the control law will generate velocities of zero for the system if $\boldsymbol{\nu}_{ijk}$ belongs to the nullspace of $\mathbf{L}\mathbf{L}^+$, this issue is analyzed in Section 4.

3.3 Desired Trajectories of the Control Input

Once the input-output linearization is carried out and the control law is obtained, the desired evolution of the input control in order to reach the target has to be defined. The input control consists of the tensor elements T_{111} , T_{113} , T_{131} , T_{313} , T_{331} and T_{333} . Next, the objective is to define smooth functions which lead from the initial to the final desired values of the trifocal tensor elements.

If we analyze the expressions that define the tensor elements (8), we can see that, for those used as input of the control, we can make a classification. On the one hand, there are trifocal tensor elements depending directly on t_{z2} (T_{113} , T_{313} and T_{333}), on the other hand there are elements that depend on t_{x2} (T_{111} , T_{131} and T_{331}). For those elements depending on t_{z2} we define sinusoids that guarantee smooth motion of the vehicle towards the target. Although we design the desired trajectories of the control input in terms of sinusoids, they could also be defined with other different criteria, for example using polynomials or parabolic functions [7]. The use of sinusoidal inputs for steering systems was discussed in [30]. That work tackles the problem of motion planning using sinusoids and considering nonholonomic constraints. There, the use of sinusoidal inputs for steering systems was discussed showing their advantages. The other elements, depending on t_{x2} , are highly related on how the lateral and orientation errors are corrected. Due to the nonholonomic constraints of the mobile platform, both functions have to be defined properly in such a way they evolve accordingly to reach successfully the target location. This issue, that was handled in [30] by constraining the initial configuration of the robot, is addressed here by the closed-loop tracking design. Therefore, the desired trajectories given as a function of $\psi(t)$ are autonomously modified by the current value of this parameter in such a way that motion constraints are overcome. Thus, we propose the continuous and time differentiable functions for the elements of the input control to be tracked as:

If ($0 \leq t \leq t_b$)

$$\begin{cases} T_{111}^d(t) = (T_{111}(0) - T_{212}) \frac{\psi(t)}{\psi(0)} + T_{212} \\ T_{113}^d(t) = \frac{T_{113}(0)}{2} + \frac{T_{113}(0)}{2} \cos\left(\frac{\pi t}{t_b}\right) \\ T_{131}^d(t) = (T_{131}(0) - T_{232}) \frac{\psi(t)}{\psi(0)} + T_{232} \\ T_{313}^d(t) = \frac{T_{313}(0) + T_{212}}{2} + \frac{T_{313}(0) - T_{212}}{2} \cos\left(\frac{\pi t}{t_b}\right) \\ T_{331}^d(t) = T_{331}(0) \frac{\psi(t)}{\psi(0)} \\ T_{333}^d(t) = \frac{T_{333}(0) + T_{232}}{2} + \frac{T_{333}(0) - T_{232}}{2} \cos\left(\frac{\pi t}{t_b}\right) \end{cases} \quad (28)$$

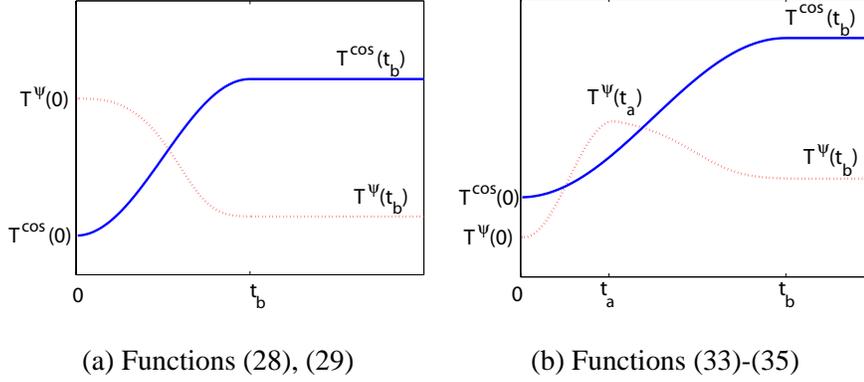


Fig. 4. Examples of the functions used to define the desired evolution of the trifocal tensor elements. These functions are defined along the time intervals $(0, t_a)$, (t_a, t_b) and (t_a, ∞) .

If $(t_b < t < \infty)$

$$\left\{ \begin{array}{l} T_{111}^d(t) = T_{212} \\ T_{113}^d(t) = 0 \\ T_{131}^d(t) = T_{232} \\ T_{313}^d(t) = T_{212} \\ T_{331}^d(t) = 0 \\ T_{333}^d(t) = T_{232} \end{array} \right. \quad (29)$$

where t_b is the time defined to reach the target location. The final desired values of the trifocal tensor elements are taken from (12) and (8). These previous functions are summarized and depicted qualitatively in Fig. 4(a). Functions for elements T_{113} , T_{313} and T_{333} are defined following a sinusoid depicted in the figure as $T^{cos}(t)$ where the initial value $T^{cos}(0)$ of the corresponding element is given by the trifocal tensor estimated at $t = 0$ and the final value $T^{cos}(t_b)$ is given by (12). The other type of function for elements T_{111} , T_{131} and T_{331} are denoted in the figure as $T^\psi(t)$, where initial and final values follow the same reasoning, but in this case the evolution of the function is related with ψ . We define $\psi = -\arctan(x_2/z_2)$ as the angular coordinate respect the world z -axis. Thus, this function ensure that the ϕ -error is corrected accordingly to x_2 and z_2 .

We propose the expressions to compute the orientation in the initial and current locations, ϕ_1 and ϕ_2 respectively, as a function of the trifocal tensor elements. We use the expressions of the trifocal tensor (8) and the relative translation vector \mathbf{t}_2 (6) to obtain

$$\phi_1 = \arcsin \left(\frac{T_{232} T_{313} - T_{212} T_{131}}{T_{232} T_{223} + T_{212} T_{221}} \right). \quad (30)$$

$$\phi_2(t) = \arccos\left(\frac{T_{223} T_{131} + T_{221} T_{313}}{T_{223} T_{232} + T_{221} T_{212}}\right). \quad (31)$$

Finally, ψ can be computed from the trifocal tensor as

$$\psi(t) = \arctan\left(\frac{T_{223} \sin \phi_2 - T_{221} \cos \phi_2}{T_{223} \cos \phi_2 + T_{221} \sin \phi_2}\right). \quad (32)$$

Note that these expressions are independent of the scale of the estimated trifocal tensor.

The initial location of the mobile platform is not restricted except that part of the scene has to be shared between the initial and target image in order to compute properly the trifocal tensor from corresponding image points. In the desired evolution of the trifocal tensor elements defined in (28) and (29) we suppose that the initial orientation is $(\psi - \phi_2 < \pi)$, allowing a smooth motion towards the target while correcting simultaneously lateral and orientation error. At the beginning of the navigation we check if $(\psi - \phi_2 < \pi)$ holds, otherwise an initial phase is needed before following the desired evolution of the trajectories expressed by (28) and (29). Thus, if two phases are required, we define T_{111}^d , T_{131}^d and T_{331}^d as follows (while T_{113}^d , T_{313}^d and T_{333}^d remain the same),

If $(0 \leq t \leq t_a)$

$$\begin{cases} T_{111}^d(t) = \frac{T_{111}(0) + T_{111}^d(t_a)}{2} + \frac{T_{111}(0) - T_{111}^d(t_a)}{2} \cos\left(\frac{\pi t}{t_a}\right) \\ T_{131}^d(t) = \frac{T_{131}(0) + T_{131}^d(t_a)}{2} + \frac{T_{131}(0) - T_{131}^d(t_a)}{2} \cos\left(\frac{\pi t}{t_a}\right) \\ T_{331}^d(t) = \frac{T_{331}(0) + T_{331}^d(t_a)}{2} + \frac{T_{331}(0) - T_{331}^d(t_a)}{2} \cos\left(\frac{\pi t}{t_a}\right) \end{cases} \quad (33)$$

If $(t_a < t \leq t_b)$

$$\begin{cases} T_{111}^d(t) = (T_{111}(t_a) - T_{212}) \frac{\psi(t)}{\psi(t_a)} + T_{212} \\ T_{131}^d(t) = (T_{131}(t_a) - T_{232}) \frac{\psi(t)}{\psi(t_a)} + T_{232} \\ T_{331}^d(t) = T_{331}(t_a) \frac{\psi(t)}{\psi(t_a)} \end{cases} \quad (34)$$

If $(t_b < t < \infty)$

$$\begin{cases} T_{111}^d(t) = T_{212} \\ T_{131}^d(t) = T_{232} \\ T_{331}^d(t) = 0 \end{cases} \quad (35)$$

where t_a is the time defined to perform the first phase and t_b has been defined previously. The only difference of these functions with the ones defined in (28) lies in the phase defined in (33) with $(0 \leq t \leq t_a)$. This new function denoted as $T^\psi(t)$ has been depicted qualitatively in Fig. 4(b). There, the desired orientation at t_a is defined qualitatively in terms of trifocal tensor elements. The values of $T^\psi(t_a)$, i.e. $T_{111}^d(t_a)$, $T_{131}^d(t_a)$ and $T_{331}^d(t_a)$, are defined experimentally as detailed in section 5. These goal values are not critical because the control allows a high margin on the orientation for the next phase.

4 Stability Analysis

In this section we analyze the stability of the proposed control law in the sense of *Lyapunov* [33]. As we measure the state of the system through the trifocal tensor elements we need first to ensure that when their desired values are reached the vehicle is actually in the target location. Thus, we need to prove that there is only one equilibrium state and it is the desired target location.

Proposition 4.1 *The vehicle is in the target location if and only if the desired values of the selected trifocal control elements have been reached. Therefore, the desired target location $(x_2, z_2, \phi_2) = (0, 0, 0)$ is the only equilibrium state of the system.*

Proof. It is straightforward to see that, if the vehicle is in the target location, the desired values of the trifocal tensor have been reached (12). Next we show that if the desired values of the trifocal tensor have been reached the vehicle is in the target location. From (8) we make equal some of the equations to the desired trifocal tensor values obtaining the following expressions

$$T'_{111} - T_{111}^d = t_{x1}(1 - \cos \phi_2) + t_{x2} \cos \phi_1 = 0 \quad (36)$$

$$T'_{131} - T_{131}^d = t_{z1}(1 - \cos \phi_2) - t_{x2} \sin \phi_1 = 0 \quad (37)$$

$$T'_{313} - T_{313}^d = t_{x1}(1 - \cos \phi_2) + t_{z2} \sin \phi_1 = 0 \quad (38)$$

$$T'_{333} - T_{333}^d = t_{z1}(1 - \cos \phi_2) + t_{z2} \cos \phi_1 = 0 \quad (39)$$

From (36) and (38) we obtain

$$t_{x2} \cos \phi_1 - t_{z2} \sin \phi_1 = 0 . \quad (40)$$

Similarly, from (37) and (39) we obtain

$$t_{x2} \sin \phi_1 + t_{z2} \cos \phi_1 = 0 . \quad (41)$$

Substituting t_{x2} from (40) into (41) we have:

$$t_{z2} \frac{\sin^2 \phi_1}{\cos \phi_1} + t_{z2} \cos \phi_1 = 0 . \quad (42)$$

Which gives $t_{z2} = 0$ since $\frac{\sin^2 \phi_1}{\cos \phi_1} + \cos \phi_1 = \frac{1}{\cos \phi_1} \neq 0$. Similarly, substituting t_{z2} from (41) into (40) we have:

$$t_{x2} \frac{\cos^2 \phi_1}{\sin \phi_1} + t_{x2} \sin \phi_1 = 0 . \quad (43)$$

And therefore $t_{x2} = 0$ since $\frac{\cos^2 \phi_1}{\sin \phi_1} + \sin \phi_1 = \frac{1}{\sin \phi_1} \neq 0$. Substituting $t_{x2} = 0$ or $t_{z2} = 0$ in (8), we have $\phi_2 = 0$ and finally $x_2 = 0$ and $z_2 = 0$. Therefore $(x_2, z_2, \phi_2) = (0, 0, 0)$ is the only equilibrium state of the system reached when the desired values of the trifocal tensor elements used in the control are achieved. \square

We consider the error to minimize as $\mathbf{e} = \mathbf{T}_{ijk} - \mathbf{T}_{ijk}^d$ and define the candidate Lyapunov function by the squared error norm

$$V(\mathbf{x}, t) = \frac{1}{2} \|\mathbf{e}\|^2 . \quad (44)$$

We need to prove that V is positive definite, \dot{V} is negative definite and V is radially unbounded. The function V is positive definite given that $V > 0$ for all $\mathbf{x} \neq \mathbf{0}$ and $V(\mathbf{0}, t) = 0$ (see proposition 4.1). We also have that V is radially unbounded given that $V(\mathbf{x}, t) \rightarrow \infty$ as $\|\mathbf{x}\| \rightarrow \infty$. Next, we study if the derivative $\dot{V}(\mathbf{x}, t)$ is negative definite. The Lyapunov candidate function derivative is

$$\begin{aligned} \dot{V} &= \mathbf{e}^T \dot{\mathbf{e}} \\ &= \mathbf{e}^T \left(\mathbf{L}(v, \omega)^T - \dot{\mathbf{T}}_{ijk}^d \right) \\ &= \mathbf{e}^T \left(\mathbf{L}\mathbf{L}^+ \boldsymbol{\nu}_{ijk} - \dot{\mathbf{T}}_{ijk}^d \right) \\ &= \mathbf{e}^T \left(\mathbf{L}\mathbf{L}^+ (\dot{\mathbf{T}}_{ijk}^d - k\mathbf{e}) - \dot{\mathbf{T}}_{ijk}^d \right) \\ &= -k\mathbf{e}^T \mathbf{L}\mathbf{L}^+ \mathbf{e} + \mathbf{e}^T \left(\mathbf{L}\mathbf{L}^+ - \mathbf{I} \right) \dot{\mathbf{T}}_{ijk}^d \end{aligned} \quad (45)$$

The global asymptotic stability of the system would need first that $\mathbf{L}\mathbf{L}^+ > 0$. However this matrix is not definite positive since $\mathbf{L}\mathbf{L}^+ \in \mathbb{R}^{6 \times 6}$ is at most of rank 2. Therefore there exist a null space in such a way that $\boldsymbol{\nu}_{ijk} \in \text{Ker}(\mathbf{L}^+)$ corresponds to local minima (i.e. the control law gives zero velocities despite the target has not

been reached [34]). In particular, we have shown with (27) that the $\text{rank}(\mathbf{L}\mathbf{L}^+) = 2$ and then, $\dim(\text{Ker}(\mathbf{L}^+)) = 4$. A basis of the null space of \mathbf{L}^+ is defined as follows

$$\left\{ \begin{pmatrix} -T_{133} \\ 0 \\ T_{113} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} T_{311} - \tan(\phi_1)T_{111} \\ -\tan(\phi_1)T_{113} \\ 0 \\ T_{113} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -T_{333} \\ 0 \\ 0 \\ 0 \\ T_{113} \\ 0 \end{pmatrix}, \begin{pmatrix} T_{331} - T_{111} \\ -T_{113} \\ 0 \\ 0 \\ 0 \\ T_{113} \end{pmatrix} \right\} \quad (46)$$

This does not imply that local minima always exist, because any configuration belonging to the nullspace must hold the trifocal tensor constraints as well. Local minima occur only with particular robot configurations, and we have not obtained general expressions because of the complexity of the computations. However, extensive simulations show that the area of convergence is very large in practice within the workspace.

Since $\dot{V}(\mathbf{0}, t) = 0$ and $\dot{V}(\mathbf{x}, t) < 0$ for all $\mathbf{x} \neq \mathbf{0}$ cannot be guaranteed, only local asymptotic stability can be obtained. Following [1], we define the new error $\mathbf{e}' = \mathbf{L}^+\mathbf{e}$ to study local asymptotic stability. The time derivative of this error is given by

$$\begin{aligned} \dot{\mathbf{e}}' &= \mathbf{L}^+\dot{\mathbf{e}} + \dot{\mathbf{L}}^+\mathbf{e} \\ &= \mathbf{L}^+ \left(\mathbf{L}(v, \omega)^T - \dot{\mathbf{T}}_{ijk}^d \right) + \dot{\mathbf{L}}^+\mathbf{e}. \end{aligned} \quad (47)$$

From [35] and [1], $\dot{\mathbf{L}}^+\mathbf{e}$ can be written as $\mathbf{O}(v, \omega)^T$ with $\mathbf{O} \in \mathbb{R}^{2 \times 2}$, and in the previous expression $\mathbf{O} \rightarrow \mathbf{0}$ and $\dot{\mathbf{T}}_{ijk}^d \rightarrow \mathbf{0}$ if $\mathbf{e} \rightarrow \mathbf{0}$. Using the control outputs (24) we obtain

$$\begin{aligned} \dot{\mathbf{e}}' &= (\mathbf{L}^+\mathbf{L} + \mathbf{O}) (v, \omega)^T - \mathbf{L}^+\dot{\mathbf{T}}_{ijk}^d \\ &= (\mathbf{L}^+\mathbf{L} + \mathbf{O}) \mathbf{L}^+\dot{\mathbf{T}}_{ijk}^d - k (\mathbf{L}^+\mathbf{L} + \mathbf{O}) \mathbf{L}^+\mathbf{e} - \mathbf{L}^+\dot{\mathbf{T}}_{ijk}^d \\ &= (\mathbf{L}^+\mathbf{L} + \mathbf{O} - \mathbf{I}) \mathbf{L}^+\dot{\mathbf{T}}_{ijk}^d - k (\mathbf{L}^+\mathbf{L} + \mathbf{O}) \mathbf{e}' \end{aligned} \quad (48)$$

To obtain the local asymptotic stability we consider the linearized system around $\mathbf{e} = \mathbf{0}$,

$$\dot{\mathbf{e}}' = -k\mathbf{L}^+\mathbf{L}\mathbf{e}', \quad (49)$$

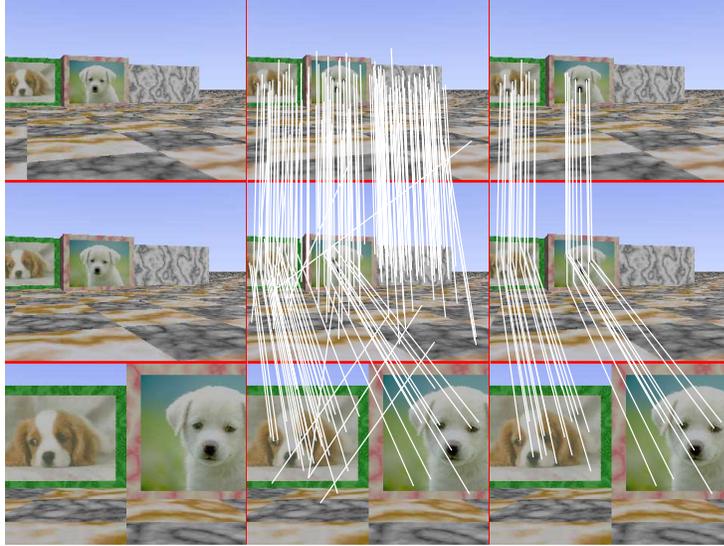


Fig. 5. Examples of images taken from the virtual scene. First row is the initial image, second row is a current image taken during the navigation and third row is the target image. Second column shows the SIFT matches between initial-current and current-target images. Third column shows the matches remaining across the three views after the robust computation of the trifocal tensor. Video attachments show the complete sequence of two simulations (videos 1 and 2).

which is locally asymptotically stable in a neighborhood of $\mathbf{e} = 0$ if $\mathbf{L}^+\mathbf{L} > 0$, where $\mathbf{L}^+\mathbf{L} \in \mathbb{R}^{2 \times 2}$. If the estimation of \mathbf{L}^+ is not too coarse, $\mathbf{L}^+\mathbf{L} > 0$ is ensured.

5 Experimental Validation

Several simulations are presented to show the validity of the approach and its performance with image noise and calibration errors. The virtual scene has been created using the Persistence of Vision Ray-Tracer¹ (POV-RayTM). The scene consists of a checkered floor and walls with real images attached as posters. The virtual scene is rendered and projected into the image plane through a pin-hole camera model in each control loop. The coordinates of the image points are transformed to calibrated coordinates before the trifocal tensor computation. Examples of an initial, current and target images obtained during a simulation is given in Fig. 5. The image size is 640×480 pixels. The putative set of matches is obtained by means of SIFT features [36], which are highly invariant to scale and rotation. Mismatches are eliminated through the robust estimation of the trifocal tensor (Fig. 5), we refer for details to [26], [17]. In the following experiments the control gain has been tuned to $k = 1$. The values of t_a and t_b used in the definition of the desired evolution of the trifocal tensor elements are $t_a = 0$ or 50 s and $(t_b - t_a) = 100$ s. These val-

¹ <http://www.povray.org/>

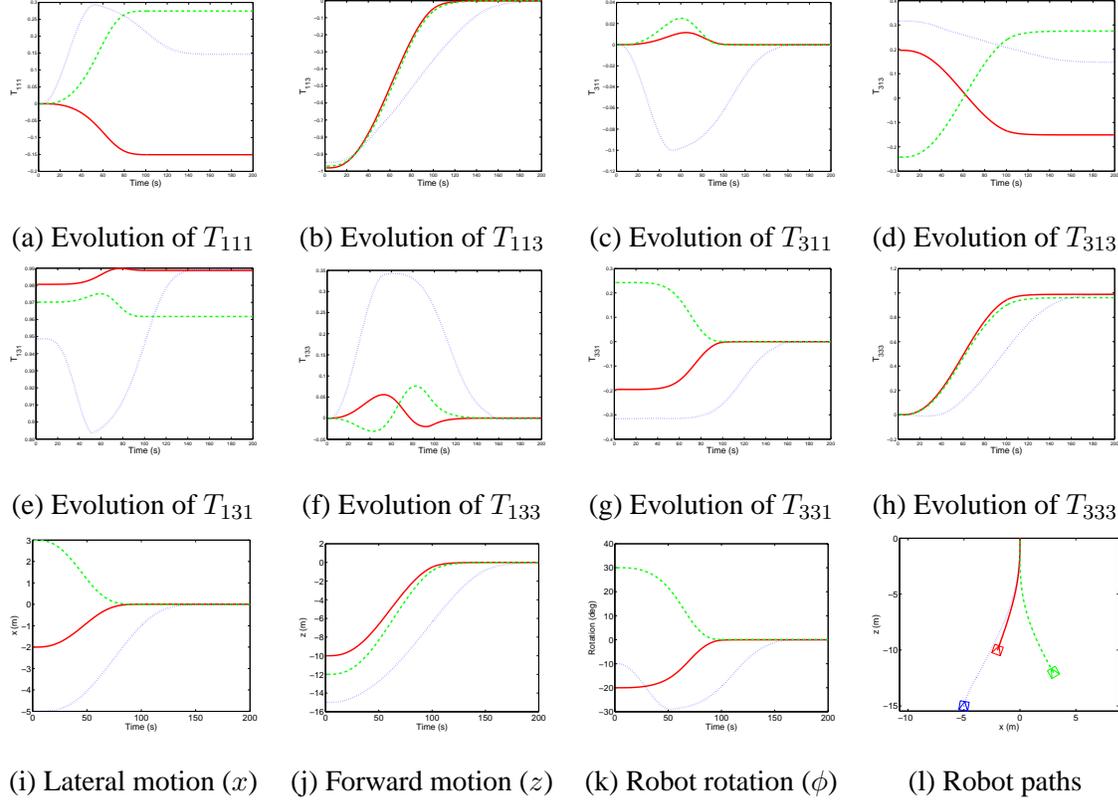


Fig. 6. Three simulations with initial locations at $(-2, -10, -20^\circ)$ with solid line, $(3, -12, 30^\circ)$ with dashed line and $(-5, -15, -10^\circ)$ with dotted line. The target location is $(0, 0, 0^\circ)$. Video attachment 3 shows additional examples.

ues have been tuned experimentally, higher values give lower robot velocities and lower values give faster motion.

Three simulations from different initial locations are presented in Fig. 6. The three examples are superposed with different line style. The evolution of the trifocal elements along time are shown in Fig. 6(a)-(h). The trifocal tensor elements converge to their desired values as defined in (29). The evolution of the x -coordinate, the z -coordinate and the orientation is shown in Fig. 6(i)-(k). The motion of the vehicle is also shown in Fig. 6(l). As it can be seen, the resultant motion is smooth and converges properly to the target location. Depending on the initial location of the vehicle, the desired trifocal tensor element trajectories have to be defined in two phases (33). An example which needs two phases is $(-5, -15, -10^\circ)$ in Fig. 6. In this case the values of $T_{111}^d(t_a)$, $T_{131}^d(t_a)$ and $T_{331}^d(t_a)$ are required qualitatively. These values have been chosen experimentally where $T_{111}^d(t_a)$ is defined as $2(T_{111}(0) + T_{212})$, $T_{131}^d(t_a)$ is defined as 80% of $T_{131}(0)$ and $T_{331}^d(t_a)$ is defined as 20% over $T_{331}(0)$. These are valid values in all the work space, taking into account that the motion of the vehicle is restricted by the limited camera field of view and only direct motion towards the target is allowed. In case that another particular evolution of the vehicle path is desired, for example parking manoeuvres, a different strategy could be defined for the desired trifocal tensor element trajectories.

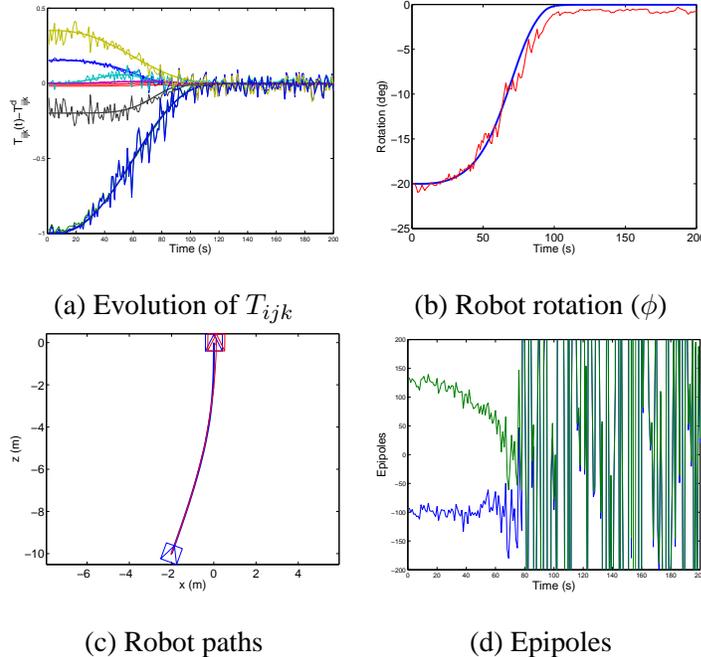


Fig. 7. Simulations with (thin line) and without (thick line) image noise of $\sigma = 1$ pixel. The initial location is $(-2, -10, -20^\circ)$. (d) Estimation of the epipoles to show the short baseline problem.

The robustness of the trifocal-based control has been tested in presence of image noise. Simulations with and without image noise are superposed in Fig. 7. The thick line is the simulation without image noise and the thin line is with image noise. The evolution of the trifocal tensor elements are shown together in Fig. 7(a). For readability, the final desired value of each element has been subtracted to draw the convergence to zero of all the elements ($T_{ijk}(t) - T_{ijk}^d$). The evolution of the orientation is shown in Fig. 7(b) and the evolution of the location is shown in Fig. 7(c). The image noise of the point correspondences consist of Gaussian noise with a standard deviation of $\sigma = 1$ pixel in point coordinates. The results show that the control law can cope with image noise converging to the target location successfully. It can be seen that a small continuous final error is obtained in x and ϕ , revealing the non-holonomic nature of the system. The motion constraint affects to x but because the coupling of the system coordinates the ϕ coordinate is also affected. The epipolar geometry is also estimated from the feature correspondences across the current and target images, and the resultant horizontal-coordinate of the epipoles are depicted in Fig. 7(d). It can be seen how the epipolar geometry becomes unstable with short baseline and therefore, epipolar-based approaches like [7], [8] requires additional procedures to reach the target.

Before computing the trifocal tensor from the correspondences they have to be transformed to calibrated coordinates by means of the internal camera calibration parameters. Next simulations show the performance of the visual control approach with errors in the camera calibration parameters. In the simulations depicted in

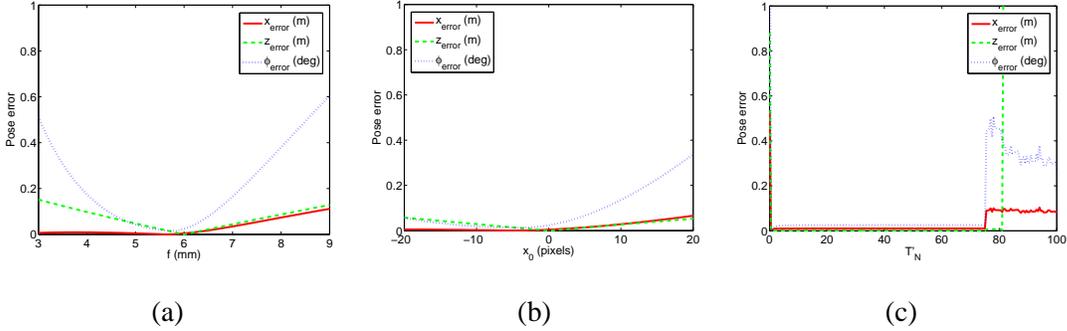


Fig. 8. Final location error in (x, z, ϕ) varying the camera calibration parameters: (a) focal length and (b) principal point coordinates. (c) Final location error as a function of T'_N .

Fig. 8(a) the known focal length is fixed to 6 mm while its real value is changed from 3 to 9 mm. In Fig. 8(b) the principal point coordinates are zero in the control law while their real value is changed. In both cases the final error obtained in the simulation when varying the camera calibration parameters is depicted. As expected, the results show that the final location error increases with the camera calibration parameter errors. This is because the control law is defined in terms of the trifocal tensor elements, and inaccuracy in the calibration parameters is transferred to the tensor. Therefore, the trajectory is altered and because of the motion constraints a final error is obtained. However, it is stable and the performance is still acceptable with small calibration errors. The parameter T'_N appears in the interaction matrix of the system (23) and it has been transferred as a constant gain of the velocity v . In the previous tests we have chosen $T'_N = 1$ for the computation of (23). The final location error as a function of T'_N is shown in Fig. 8(c). It can be seen that the final error is not affected by T'_N except for extreme values. When the selected T'_N is close to zero the velocities are too low and the robot does not move enough to reach the target. With high values of T'_N the velocities are high and the robot oscillates around the target location in z -axis (forward and backward). This is the expected behavior of the system and it includes the usual tuning of the control gains. Fig. 8(c) shows that quite flexibility is allowed for choosing T'_N in practice.

6 Conclusions

This paper considers the problem of autonomous visual control of a nonholonomic vehicle. We have presented a new vision-based control approach which is based on the trifocal tensor. The control law is defined by the exact input-output linearization of the system through the trifocal tensor. With this control law, the vehicle is autonomously driven towards the target as the desired values of the trifocal tensor elements are achieved. This approach avoids the need of metric information or additional data from the environment by relying directly on terms of the trifocal tensor elements. Some advantages of the trifocal tensor based approach are that it is more

robust than two view geometry thanks to additional information of a third view and that the problem of short baseline with epipolar geometry is overcome. The stability analysis of the system in the Lyapunov sense is also presented. Simulations have been carried out to test the approach showing good performance. Simulations in the presence of image noise have been carried out to show that the method performs correctly. Results with calibration errors in the internal camera parameters have been also presented. We have considered planar motion with nonholonomic constraints, an interesting issue to address is to extend this approach to robot manipulators, in that case planar motion can not be assumed but on the other hand there are no motion constraints.

References

- [1] F. Chaumette and S. Hutchinson, "Visual servo control, part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, Dec. 2006.
- [2] —, "Visual servo control, part II: Advanced approaches," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 109–118, Mar. 2007.
- [3] N. R. Gans and S. A. Hutchinson, "A stable vision-based control scheme for nonholonomic vehicles to keep a landmark in the field of view," in *IEEE International Conference on Robotics and Automation*, Apr. 2007, pp. 2196–2200.
- [4] E. Marchand and F. Chaumette, "Feature tracking for visual servoing purposes," *Robotics and Autonomous Systems*, vol. 52, pp. 53–70, 2005.
- [5] R. Basri, E. Rivlin, and I. Shimshoni, "Visual homing: Surfing on the epipoles," *International Journal of Computer Vision*, vol. 33, no. 2, pp. 117–137, 1999.
- [6] P. Rives, "Visual servoing based on epipolar geometry," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, 2000, pp. 602–607.
- [7] G. L. Mariottini, D. Prattichizzo, and G. Oriolo, "Epipole-based visual servoing for nonholonomic mobile robots," *IEEE International Conference on Robotics and Automation*, pp. 497–503, 2004.
- [8] G. López-Nicolás, C. Sagüés, J. J. Guerrero, D. Kragic, and P. Jensfelt, "Nonholonomic epipolar visual servoing," in *IEEE International Conference on Robotics and Automation*, May 2006, pp. 2378–2384.
- [9] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 D visual servoing," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 2, pp. 234–246, Apr. 1999.
- [10] E. Malis and F. Chaumette, "2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *International Journal of Computer Vision*, vol. 37, no. 1, pp. 79–97, 2000.
- [11] E. Malis and S. Benhimane, "A unified approach to visual tracking and servoing," *Robotics and Autonomous Systems*, vol. 52, pp. 39–52, 2005.

- [12] Y. Fang, W. E. Dixon, D. M. Dawson, and P. Chawda, "Homography-based visual servo regulation of mobile robots," *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, vol. 35, no. 5, pp. 1041–1050, 2005.
- [13] S. Benhimane, E. Malis, P. Rives, and J. R. Azinheira, "Vision-based control for car platooning using homography decomposition," in *IEEE International Conference on Robotics and Automation, Barcelona, Spain*, April 2005, pp. 2173–2178.
- [14] C. Sagüés and J. J. Guerrero, "Visual correction for mobile robot homing," *Robotics and Autonomous Systems*, vol. 50, no. 1, pp. 41–49, 2005.
- [15] J. Chen, W. Dixon, M. Dawson, and M. McIntyre, "Homography-based visual servo tracking control of a wheeled mobile robot," *Robotics, IEEE Transactions on [see also Robotics and Automation, IEEE Transactions on]*, vol. 22, no. 2, pp. 406–415, April 2006.
- [16] G. López-Nicolás, C. Sagüés, and J. J. Guerrero, "Homography-based visual control of nonholonomic vehicles," in *IEEE International Conference on Robotics and Automation*, Apr. 2007, pp. 1703–1708.
- [17] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.
- [18] K. Åström and M. Oskarsson, "Solutions and ambiguities of the structure and motion problem for 1D retinal views," *Journal of Mathematical Imaging and Vision*, vol. 12, pp. 121–135, 2000.
- [19] A. Murillo, C. Sagüés, J. Guerrero, T. Goedemé, T. Tuytelaars, and L. Van-Gool, "From omnidirectional images to hierarchical localization," *Robotics and Autonomous Systems*, vol. 55, no. 5, pp. 372–382, 2007.
- [20] F. Dellaert and A. W. Stroupe, "Linear 2D localization and mapping for single and multiple robot scenarios," in *IEEE International Conference on Robotics and Automation*, May 2002, pp. 688–694.
- [21] L. Quan and M. Lhuillier, "Structure from motion from three affine views," in *Proceedings of the 16th International Conference on Pattern Recognition*, vol. IV, Aug. 2002, pp. 1–6.
- [22] Y. K. Yu, K. H. Wong, M. M. Y. Chang, and S. H. Or, "Recursive camera-motion estimation with the trifocal tensor," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 36, no. 5, pp. 1081–1090, Oct. 2006.
- [23] Y. K. Yu, K. H. Wong, S. H. Or, and M. M. Y. Chang, "Robust 3-D motion tracking from stereo images: A model-less method," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 3, pp. 622–630, March 2008.
- [24] B. H. Becerra and C. Sagüés, "A Novel 1D Trifocal Tensor-Based Control for Differential-Drive Robots," in *IEEE International Conference on Robotics and Automation*, May 2009, pp. 1104–1109.
- [25] A. Shashua and M. Werman, "Trilinearity of three perspective views and its associated tensor," in *Proceedings of the Fifth International Conference on Computer Vision*, 1995, pp. 920–925.

- [26] P. H. S. Torr and A. Zisserman, “Robust parameterization and computation of the trifocal tensor,” *Image and Vision Computing*, vol. 15, pp. 591–607, 1997.
- [27] R. Hartley and R. Vidal, “The multibody trifocal tensor: Motion segmentation from 3 perspective views,” in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 27 June–2 July 2004, pp. 769–775.
- [28] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [29] M. I. A. Lourakis and A. A. Argyros, “Fast trifocal tensor estimation using virtual parallax,” in *IEEE International Conference on Image Processing*, vol. 2, Sept. 2005, pp. 93–96.
- [30] R. Murray and S. Sastry, “Nonholonomic Motion Planning: Steering Using Sinusoids,” in *IEEE Transactions on Automatic Control*, vol. 38, no. 5, pp. 700–716, 1993.
- [31] J.-J. E. Slotine and W. Li, *Applied nonlinear control*. Prentice Hall, Englewood Cliffs NJ, 1991.
- [32] A. Isidori, *Nonlinear Control Systems*. Springer, 1995.
- [33] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2001.
- [34] F. Chaumette, “Potential problems of stability and convergence in image-based and position-based visual servoing,” *The Confluence of Vision and Control*, D. Kriegman, G. Hager, and A. Morse, Eds. LNCIS, no 237, Springer-Verlag, 1998, pp. 66–78.
- [35] E. Malis, “Visual servoing invariant to changes in camera-intrinsic parameters,” *IEEE Transactions on Robotics and Automation*, vol. 20, no. 1, pp. 72–81, 2004.
- [36] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.