Citation: G. López-Nicolás, C. Sagüés, J. J. Guerrero, D. Kragic, and P. Jensfelt. Switching visual control based on epipoles for mobile robots. Robotics and Autonomous Systems, vol. 56, no. 7, pp. 592-603, 2008

Switching Visual Control Based on Epipoles for Mobile Robots

G. López-Nicolás[1] [a] C. Sagüés [a] J.J. Guerrero [a] D. Kragic [b]
P. Jensfelt [b]

[a] *DIIS - I3A, Universidad de Zaragoza, C/ María de Luna 1, E-50018 Zaragoza, Spain*
[b] *CAS - CVAP, Royal Institute of Technology, SE-100 44 Stockholm, Sweden*

Abstract

In this paper, we present a visual control approach consisting in a switching control scheme based on the epipolar geometry. The method facilitates a classical teach-by-showing approach where a reference image is used to control the robot to the desired pose (position and orientation). As a result of our proposal a mobile robot carries out a smooth trajectory towards the target and the epipolar geometry model is used through the whole motion. The control scheme developed considers the motion constraints of the mobile platform in a framework based on the epipolar geometry that does not rely on artificial markers or specific models of the environment. The proposed method is designed in order to cope with the degenerate estimation case of the epipolar geometry with short baseline. Experimental evaluation has been performed in realistic indoor and outdoor settings.

Key words: Visual control, the epipolar geometry, mobile robot.

1 Introduction

It is generally accepted that machine vision is the most promising sensory modality in the context of navigation. Visual servoing is an important building block of mobile manipulation systems as it provides rich information on the environment [9]. A significant amount of work has been reported in the area of visual servoing during the last decade, [5, 7, 11, 16–18, 21]. However, many of the above references concentrate mainly on problems inherent to image, position and hybrid visual servoing

[1] Corresponding author. Fax: +34 976 761914. E-mail: gonlopez@unizar.es
This work was supported by projects DPI2006-07928, IST-1-045062-URUS-STP.

assuming that there are no constraints on the robot motion itself. Most commercial platforms available today are nonholonomic and most of the above approaches cannot be used for their motion control.

The basic idea of teach-by-showing visual servoing is to control a robot to a desired reference position by regulation an error term to zero. which is estimated by matching image features between current and target images. A pioneering example of this approach is presented in [1]. There, the control is performed by recovering the relative pose of the robot between the current and desired images. The method is based on the computation of the essential matrix, assuming that the internal parameters of the camera are precisely known. In [3], a framework is proposed based in a visual route consisting of a sequence of key images of the environment where the control scheme is based on the homography of the ceiling. In [15], the navigation task is considered as tracking an arbitrarily shaped ground curve and they present a stabilizing control law for tracking analytic curves.

In [19], a visual servoing method dealing with the motion constraints of a mobile robot is presented. The controller is based on the epipolar geometry and consists in an input-output feedback linearizing control law. However, the motion performed drives the robot from and not to the target, i.e. just moving away while the lateral error is corrected. After that, the robot moves backwards in a straight line towards the target with a different controller based on correlation. In the present work the motion is performed directly towards the target, but instead of following a straight path like in [13] the robot is driven in a smooth trajectory in order to reach the target with the desired orientation. Moving backwards may be not practical in visual-based robot navigation where the only sensor is a camera pointing forward. Therefore we constrained the design of the controller with forward motion; so, backward manoeuvres are not allowed, resulting in a direct motion towards the target. This implies that the robot motion is restricted by the camera field of view. Therefore, it is assumed that distance in depth from the initial position to the target is greater than the side distance from the initial position to the target, avoiding the need of high rotations to reach the target.

The estimation of the epipolar geometry becomes ill-conditioned when the robot is close to the target and short baseline image pairs are considered. Working with homographies eliminates this problem. Fang et al. [6] proposed the asymptotic regulation of the position and orientation of a mobile robot by exploiting homography-based visual servo control strategies. The approach presented in [2] consists in a system for car platooning, using visual tracking by estimating the homography between a selected reference template attached to the leading vehicle. Here, we propose an auxiliary procedure to avoid the degeneracies due to short baseline when using the epipolar based control and we define an auxiliary fundamental matrix with another image taken during the motion. Then, the navigation is performed using a control based on the epipoles during the whole motion.

Our method consists in a switching control law based on the epipolar geometry estimation where only one image taken at the target position is given as previous information. The proposed approach is suited for a mobile platform and does not require complete camera calibration or any specific knowledge of the scene geometry. The use of uncalibrated cameras has many advantages as previously discussed [9]. Methods that do not require any special modelling of the environment, and those that do not consider the use of fiducial markers, are of significant importance in practise. The experimental evaluation, performed in realistic settings, shows the effectiveness of our method.

This paper is organized as follows. In Section 2, we present the estimation of the epipolar geometry based on scale invariant image features. Perceptual and motion models are described in Section 3, followed by the design of the control law in Section 4. Experimental evaluation and conclusions are given in Section 5 and 6 respectively.

2 From Images to Epipoles

The visual information used in our approach is obtained from image pairs (current and target image). In this section, we shortly describe the computation of the epipolar geometry and the visual features used in the process.

In a recent study, Mikolajczyk and Schmid [20] analyzed a large number of interest point descriptors and their behaviors under changes, such as scale and illumination variation. The descriptor that turned out to be most robust in this study was the Scale Invariant Feature Transform - SIFT, which was proposed in [14]. It was also concluded that the point detector used was less significant. In the current implementation, we use SIFT features as originally proposed, and refer to [14] for all details. As it can be seen in Fig. 1, the descriptor corresponds to highly distinctive image locations and is robustly invariant to image plane transformations such as translation, rotation and scaling. Matching between images is performed using a simple squared distance measure between descriptors.

Once matches are available, the position of the epipoles can be estimated. We first start by estimating the fundamental matrix using the robust approach proposed by Torr [24] called MAximum a Posteriori SAmple Consensus (MAPSAC) which is based on a well-known 7-point algorithm [8]. Similar to RANSAC, the method proceeds by repeatedly generating hypothesis from a minimal seven point correspondences and minimizing an error term for a predefined set of point combinations. Fig. 1 shows the estimated epipoles and some corresponding epipolar lines for one of the current-target image pairs. Matching scale and rotation invariant features allows us successful control under significant viewpoint changes.

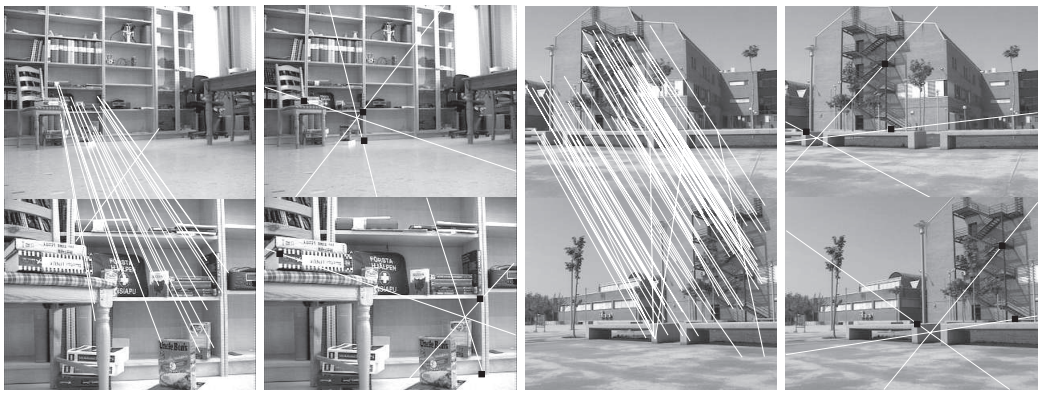


Fig. 1. Indoor and outdoor examples of matching between a current image (top) and a target image (bottom). 25 and 74 matches have been found respectively. After MAPSAC, 21 and 62 matches remain to estimate the epipoles; three epipolar lines are drawn as example in each case.

3 Perceptual and Motion Models

In this section we present the mobile robot system and define the expressions that relate the epipoles with the state of the robot. Let us suppose that the state of the robot is given by its position and orientation coordinates $\mathbf{x} = (x, z, \phi)^T$. The coordinate frame is defined in the target position as depicted in Fig. 2. From the perspective projection of Fig. 2, the x-coordinates of the epipole in the current image (e_{cx}) and in the target image (e_{tx}) can be expressed as a function of the state of the robot [13, 19] (See Appendix A for detailed derivation of these expressions)

$$e_{cx} = \alpha_x \frac{x\cos\phi + z\sin\phi}{z\cos\phi - x\sin\phi}\,, \tag{1}$$

$$e_{tx} = \alpha_x \frac{x}{z}\,. \tag{2}$$

The parameter α_x is the focal length of the camera in pixel dimensions. We have that $\alpha_x = f\,m_x$, where f is the focal length in distance units and m_x is pixel per distance unit. In practise, we assume that the principal point is in the center of the image and there is no skew. The origin of the coordinate system is fixed on the center of the image. In Section 5 we analyze the robustness of the controller to changes on these assumptions.

Let us also suppose that the nonholonomic differential kinematics to be expressed in a general way as

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = h(\mathbf{x}) \end{cases}$$

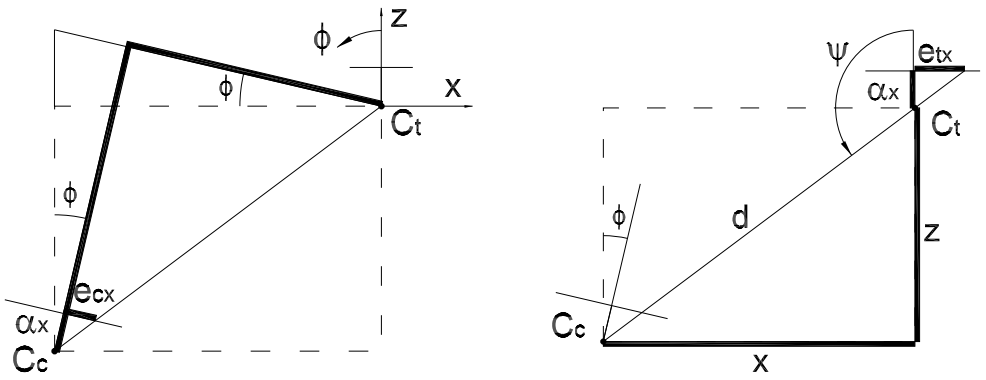


Fig. 2. Geometric relations for e_{cx} (left) and e_{tx} (right). C_c and C_t are the current and target camera positions respectively.

where $\mathbf{u} = (v, \omega)^T$ denotes the input vector, which includes the translational (v) and rotational (ω) velocities, and $\mathbf{y} = (e_{cx}, e_{tx})^T$ the output vector. The particular nonholonomic differential kinematics of the robot expressed in state space form as a function of the translation and rotation velocities of the robot (v, ω) is as follows

$$\begin{pmatrix} \dot{x} \\ \dot{z} \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} v + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \omega \,, \tag{3}$$

$$\mathbf{y} \quad = \quad (e_{cx} \;\; e_{tx})^T \;\; .$$

Also we can express the x and z coordinates as a function of the polar coordinates d and ψ (Fig. 2) as

$$x = -d\sin\psi \;, \;\; z = d\cos\psi \;, \tag{4}$$

with $\psi = -\arctan(e_{tx}/\alpha_x)$ and $\phi = \psi + \arctan(e_{cx}/\alpha_x)$.

4 Control Law

A common way to face the problem of controlling nonlinear systems is through linearization [10,23]. The objective is to perform the navigation by using a feedback control law where the inputs are the coordinates of the epipoles. Then the visual servoing problem is transformed in a tracking problem in a nonlinear system where the reference trajectories of the epipoles are defined.

4.1 Input-output Linearization

As the system is nonlinear, an input-output linearization is made first to design the controller. This input-output linearization [10, 23] is carried out as originally presented in [19] by differentiating the system outputs until the inputs appear explicitly, and then solving for the control inputs. So, for the epipole in the current image according to (1), and using (3) and (4) after the differentiation, we have

$$\dot{e}_{cx} = -\frac{\alpha_x \sin(\phi - \psi)}{d\cos^2(\phi - \psi)}\ v + \frac{\alpha_x}{\cos^2(\phi - \psi)}\ \omega\ . \tag{5}$$

Similarly for the epipole in the target image (2), and using again (3) and (4) it follows that

$$\dot{e}_{tx} = -\frac{\alpha_x \sin(\phi - \psi)}{d\cos^2\psi}\ v\ . \tag{6}$$

The derivation of (5) and (6) is detailed in Appendix B. Now, we define the new inputs in closed loop ν_c and ν_t as functions depending on the tracking error $(e_{cx} - e^r_{cx})$ and $(e_{tx} - e^r_{tx})$. Assuming the control objective to be the output to track the set point $(e^r_{cx},\ e^r_{tx})$, we can make

$$\begin{cases} \nu_c = \dot{e}^r_{cx} - k_c(e_{cx} - e^r_{cx}) \\ \nu_t = \dot{e}^r_{tx} - k_t(e_{tx} - e^r_{tx}) \end{cases}$$

$k_c > 0$ and $k_t > 0$ being the controller gains. The tracking error from the new inputs results in an exponentially stable error dynamics [23].

4.2 Switching Control Scheme

Here we define a switching control law to perform the robot motion towards the target. The motion to the target is divided in three sequential steps. The first one is a rotation to point the robot properly to the target. The second step reduces to zero the lateral and orientation errors of the robot to the target position. Finally, the robot moves towards the target following a straight path. The resultant path for the robot is shown in Fig. 3. Next, the controllers of each step are defined.

4.2.1 Controller of step 1, $t \in [T_0, T_1)$

There is no constraint about the initial orientation of the robot, only that at least a part of the scene must be shared by the initial and target images in order to find enough matches for computing the epipolar geometry. With the first controller the robot performs a pure rotation. The rotation could also be carried out combined with a translation moving forward, for example in case of car-like motion constraints. From (5) and taking into account that $v = 0$ we have that

$$\omega = \frac{\cos^2(\phi - \psi)}{\alpha_x} \nu_c \ , \tag{7}$$

where the reference function e^r_{cx} needed to compute ν_c is defined next. From (6) we obtain the constraint $\dot{e}_{tx} = 0$ and therefore, $e^r_{tx} = constant$, choosing its value as the initial $e^r_{tx} = e_{tx}(0)$. Then we define a smooth parabolic reference function as

$$e^r_{cx} = (\lambda_e + e_{cx}(0)) \left(\frac{t^2}{T_1^2} - 2\frac{t}{T_1} \right) + e_{cx}(0) \ ,$$

where T_1 is the duration time of the first step and λ_e is a constant value which determines the maximum rotation of the robot in this step. If $\lambda_e = 0$ the robot would point towards the target and, in this case, the robot would follow a straight line in the second step; a higher absolute value will produce a higher rotation. More rotation in the first step implies that the lateral error will be corrected faster at the next step. The definition of a proper value for e^r_{tx} requires that $\lambda_e > 0$ if $x < 0$ or $\lambda_e < 0$ if $x > 0$. Note that $e_{tx}(0)$ gives these particular signs (see Fig. 2). Besides, for higher initial position values of x, higher values of lambda (rotation) are required; this matches with $e_{tx}(0)$ because its value increases with $|x|$. Therefore we have chosen the target epipole to define λ_e. So, in our experiments we define $\lambda_e = e_{tx}(0)$. Nevertheless, the rotation can be limited by the camera field of view; in this case, the first step is carried out until the reference value is reached or while a minimum set of matches is available.

4.2.2 Controller of step 2, $t \in [T_1, T_3)$

The robot moves towards the target while correcting the lateral and orientation error between current and target positions. From (6), the velocity is

$$v = -\frac{d \cos^2(\psi)}{\alpha_x \sin(\phi - \psi)} \nu_t \ . \tag{8}$$

The parameter d is the distance to the target, which is unknown. We fix d to a constant value considering it as a control gain. Experiments in Section 5 show that

the feedback control law copes correctly with this approximation.

The angular velocity ω is obtained from (5) as

$$\omega = \frac{\cos^2(\phi-\psi)}{\alpha_x}\,\nu_c + \frac{\sin(\phi-\psi)}{d}\,v\,,$$

and using (8) with the previous expression it follows that

$$\omega = \frac{\cos^2(\phi-\psi)}{\alpha_x}\,\nu_c - \frac{\cos^2(\psi)}{\alpha_x}\,\nu_t\,. \tag{9}$$

The reference functions needed to compute ν_c and ν_t are defined as second order polynomial in time as

$$e^r_{tx} = e_{tx}(0)\left(\frac{t^2}{T_2^2} - 2\frac{t}{T_2} + 1\right)\,,$$
$$e^r_{cx} = -\lambda_e\left(\frac{t^2}{T_3^2} - 2\frac{t}{T_3} + 1\right)\,,$$

where λ_e has been already defined and the time duration of each reference function are T_2 and T_3 respectively. We require the target epipole to be regulated to zero before the current one, so $T_2 < T_3$, in order to avoid the singularity in (8) when $\sin(\phi-\psi) = 0$ at the end of the step. Thus, when the time T_3 is reached both epipoles are zero and the control switches to the next step. The particular case in which the initial position of the robot is aligned with the target, i.e. $(x,z,\phi) = (0,z,0)$, makes the controller to switch automatically to step 3, performing properly the motion to the target following a straight line.

4.2.3 Controller of step 3, $t \in [T_3, T_5)$

When the control switches to the third controller, the robot is pointing to the target with no lateral or orientation error and only a straight forward motion is left. The robot moves with constant velocity v_3 . Here we have $\nu_t = 0$ and then, from (9)

$$\omega = \frac{\cos^2(\phi-\psi)}{\alpha_x}\,\nu_c\,. \tag{10}$$

The previous expression is used during this last motion to correct the orientation error due to noise or drift, which cannot be ignored in real situations. The reference trajectories of the epipoles are set as $e^r_{cx} = 0$ and $e^r_{tx} = 0$.

The behavior of the robot fully depends on the reference trajectories of the epipoles. The trajectories that have been defined for the three controllers of the switching control are summarized in Fig. 3 together with the resultant motion of the robot.

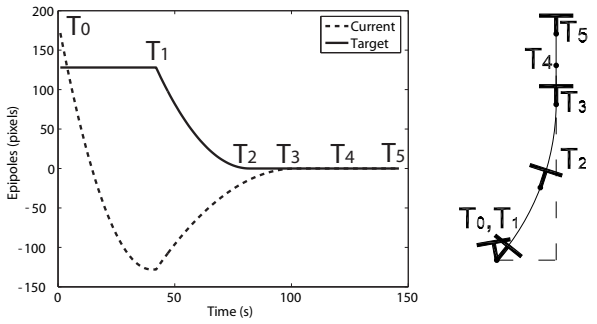


Fig. 3. Reference trajectories of the epipoles (left) and robot motion (right).

During the last step, when the robot comes closer to the target, the epipolar geometry turns out to be not well defined. Therefore, an additional consideration is needed to servo the robot in the final step. We can use the same epipolar control model proposed by taking into account that, at the end of the second step, in T_3 (Fig. 3), the epipoles are zero and the robot has the same orientation as in the target position. Then, the image taken at T_3 is used as an auxiliary target to correct orientation at the final step of the control loop. During the straight motion from T_3 to T_5, two fundamental matrices are computed: one relating the current position with the target (F) and the other one between the current position and the auxiliary target (F_{aux}). The method switches from the fundamental matrix F to F_{aux} in order to avoid the degeneracy of F at the end of the motion because of the short baseline. The switch (T_4) is autonomously done when the estimated epipoles are detected to become unstable.The issue of detecting degenerate situations has been addressed in [24]. That work discusses related methods and present strategies for tackling the problem of degeneracy by using a statistical model selection test. Here we use a simpler approach consisting in checking the evolution of the estimated epipoles considering that they become unstable if their values change suddenly over a threshold. Then, the robot goes on moving towards the target position and it stops when the mean squared error of the feature coordinates between the current and target image is minimum.

Alternatively, a suitable approach with short baseline is to use a homography based control [6,22]. This is a viable solution since, close to the target, all the environment is seen in the plane of the infinity. We have implemented a homography-based control for the last step of the navigation to compare with our method for short baseline. This approach consists of a proportional controller on two elements of the homography matrix computed between current and target image (refer to [12] for details). The control is decoupled as only pure translation is left, having the homography $H = h_{ij}$ with $(i, j) = 1,2,3$, and that the desired final homography is the identity matrix, the velocities of the robot to reach the target are given by

$$\begin{cases} v = k_v\,(1 - h_{33}) \\ \omega = -k_\omega\,h_{13} \end{cases} \tag{11}$$

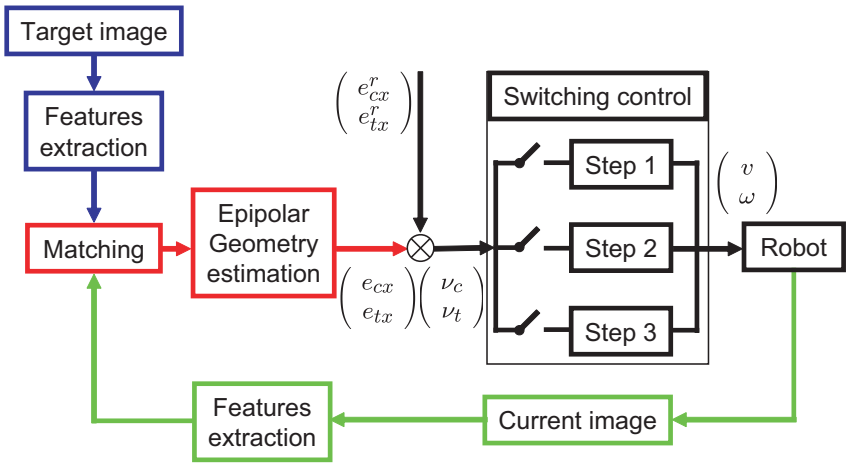


Fig. 4. Diagram of the switching control law. An image in the current position is taken at each loop of the control. The epipolar geometry that links it with the target image is computed from the feature matching. Using the fundamental matrix, the switching control law gives the rotation and translation velocities to carry out the motion.

$k_v > 0$ and $k_\omega > 0$ being the control gains. This controller changes the model used during the navigation from the epipolar geometry to the homography, being one advantage that the last step is now uncalibrated. The robot motion of the proposed control law results in an asymptotic correction of the rotation and lateral distance (the x-coordinate), which is critical when using mobile robots with nonholonomic motion constraints. Using the additional homography based controller the resultant motion is also asymptotic in the correction of depth.

The switching control law presented is summarized in algorithm 1, where k_v and k_ω are positive control gains. A diagram of the switching control law is shown in Fig. 4. When the loop finishes, the robot is in the target position, and current and target images are the same.

Algorithm 1 Switching Control Law

1) $$\begin{pmatrix} v \\ \omega \end{pmatrix} = \begin{bmatrix} k_v & 0 \\ 0 & k_\omega \end{bmatrix} \begin{pmatrix} 0 \\ \cos^2(\phi - \psi)\, \nu_c / \alpha_x \end{pmatrix}$$

2) $$\begin{pmatrix} v \\ \omega \end{pmatrix} = \begin{bmatrix} k_v & 0 \\ 0 & k_\omega \end{bmatrix} \begin{pmatrix} -\frac{d \cos^2(\psi)\, \nu_t}{\alpha_x \sin(\phi - \psi)} \\ \frac{\cos^2(\phi - \psi)\, \nu_c - \cos^2(\psi)\, \nu_t}{\alpha_x} \end{pmatrix}$$

3) $$\begin{pmatrix} v \\ \omega \end{pmatrix} = \begin{bmatrix} k_v & 0 \\ 0 & k_\omega \end{bmatrix} \begin{pmatrix} v_3 \\ \cos^2(\phi - \psi)\, \nu_c / \alpha_x \end{pmatrix}$$

3') $$\begin{pmatrix} v \\ \omega \end{pmatrix} = \begin{bmatrix} k_v & 0 \\ 0 & k_\omega \end{bmatrix} \begin{pmatrix} 1 - h_{33} \\ -h_{13} \end{pmatrix}$$

4.3 Stability Analysis

The analysis of stability for switched systems is developed in [4]. In that work it is defined the *sequence nonincreasing condition* which guarantees the stability of such systems. Our particular control scheme is a deterministic sequential switching control and therefore, if the individual controllers are stable, the global system will be stable as well. In order to analyze the stability of the epipolar-based control law, we define the Lyapunov candidate function expressing the robot position in the coordinate system $(x(t),\, z(t),\, \phi(t))$, as

$$V = V_x + V_z + V_\phi = \frac{(x-x^r)^2}{2} + \frac{(z-z^r)^2}{2} + \frac{(\phi-\phi^r)^2}{2}\,, \tag{12}$$

where $(x^r,\, z^r,\, \phi^r)$ is the desired position at the end of each step. We have that V is positive if the reference state of the system corresponding to the desired position for every step has not been reached. Otherwise $V = 0$ and the corresponding reference state has been reached. Therefore V is positive definite. The scalar function V is differentiable, and its derivative with respect to time is $\dot{V} = \dot{V}_x + \dot{V}_z + \dot{V}_\phi$, where

$$\dot{V}_x = (x-x^r)\,\dot{x} = -(x-x^r)\,v\,\sin\phi\,. \tag{13}$$

$$\dot{V}_z = (z-z^r)\,\dot{z} = (z-z^r)\,v\,\cos\phi\,. \tag{14}$$

$$\dot{V}_\phi = (\phi-\phi^r)\,\dot{\phi} = (\phi-\phi^r)\,\omega\,. \tag{15}$$

Now we analyze these functions in each step of the switching control scheme in order to prove that $\dot{V}$ is negative definite.

4.3.1 Controller of step 1

The robot only performs a rotation in the initial position, so $\dot{V}_x = 0$ and $\dot{V}_z = 0$. Thus, we have to study $\dot{V} = \dot{V}_\phi$. We analyze the sign of this derivative function (15) depending on the initial robot position. All the possible cases in the plane can be divided by the z-axis: $x < 0$ if the robot is in the left semiplane and $x > 0$ if it is in the right semiplane. The initial orientation of the robot is not constrained (up to field-of-view constraints), therefore, in each case we need to consider the different values of the robot orientation. Refer to Fig. 2 to follow the reasoning and (7) for

deducing the sign of the control velocity ω :

$$x<0\rightarrow\left\{\begin{array}{l}(\phi-\phi^r)>0,\ \text{thus}\ \ (e_{cx}<0\ \text{ and }\ e_{cx}>e^r_{cx})\\ \quad\text{or}\ \ (e_{cx}>0)\Rightarrow(e_{cx}-e^r_{cx})>0\Rightarrow \nu_c<0\Rightarrow\omega<0\\ (\phi-\phi^r)<0,(\text{thus}\ \ e_{cx}<0\ \text{ and }\ e_{cx}<e^r_{cx})\\ \qquad\qquad\Rightarrow(e_{cx}-e^r_{cx})<0\Rightarrow \nu_c>0\Rightarrow\omega>0\end{array}\right\}\Rightarrow\dot{V}_\phi<0\,.$$

$$x>0\rightarrow\left\{\begin{array}{l}(\phi-\phi^r)>0,(\text{thus}\ \ e_{cx}>0\ \text{ and }\ e_{cx}>e^r_{cx})\\ \qquad\qquad\Rightarrow(e_{cx}-e^r_{cx})>0\Rightarrow \nu_c<0\Rightarrow\omega<0\\ (\phi-\phi^r)<0,\text{thus}\ \ (e_{cx}>0\ \text{ and }\ e_{cx}<e^r_{cx})\\ \quad\text{or}\ \ (e_{cx}<0)\Rightarrow(e_{cx}-e^r_{cx})<0\Rightarrow \nu_c>0\Rightarrow\omega>0\end{array}\right\}\Rightarrow\dot{V}_\phi<0\,.$$

Thus, for all possible cases depending on the robot position, we obtain that $\dot{V}_\phi<0$. Therefore $\dot{V}<0$ and the first controller is stable in the Lyapunov sense.

4.3.2 Controller of step 2

In the second step, the robot moves towards the target while lateral and orientation errors are corrected. The desired position is (x^r, z^r, ϕ^r)=(0, z^r, 0). The robot moves towards the target and the linear velocity is required to be positive in any case. It can be seen that the velocity *v* given by the control law (8) is positive in any case:

$$v\ \rightarrow\left\{\begin{array}{l}e_{tx}<0\rightarrow(\phi-\psi)>0,\ \nu_t>0\\ e_{tx}>0\rightarrow(\phi-\psi)<0,\ \nu_t<0\end{array}\right\}\Rightarrow v>0\,.$$

Now we analyze each derivative term of the function $\dot{V}=\dot{V}_x+\dot{V}_z+\dot{V}_\phi$. In order to check graphically the signs of the different elements of $\dot{V}$ refer to Fig. 2.

$$\dot{V}_x\ (13)\rightarrow\left\{\begin{array}{l}(x-x^r)=x>0\rightarrow\phi>0\\ (x-x^r)=x<0\rightarrow\phi<0\end{array}\right\}\Rightarrow\ As\ \ v>0,\ \dot{V}_x<0\,.$$

The target has been defined in front of the initial robot position, in such a way that backwards motion is not required. Then, we have $(z-z^r)<0$,

$$\dot{V}_z\ (14)\rightarrow(z-z^r)<0\Rightarrow\ As\ \ v>0,\ \ \dot{V}_z<0\,.$$

In this step, $(\phi - \phi^r) = \phi$ and $(e_{cx} - e_{cx}^r) = e_{cx}$. The sign of the angular velocity ω given by the control is deduced from (9),

$$\dot{V}_\phi\ (15) \rightarrow \left\{ \begin{array}{l} \phi > 0,\ e_{cx} > 0,\ e_{tx} < 0 \rightarrow \nu_c < 0,\ \nu_t > 0 \Rightarrow \omega < 0 \\ \phi < 0,\ e_{cx} < 0,\ e_{tx} > 0 \rightarrow \nu_c > 0,\ \nu_t < 0 \Rightarrow \omega > 0 \end{array} \right\} \Rightarrow \dot{V}_\phi < 0\ .$$

Thus, we have shown in the different cases that $\dot{V}_x < 0$, $\dot{V}_z < 0$ and $\dot{V}_\phi < 0$; and therefore, $\dot{V} < 0$ and the control in step 2 is stable in the Lyapunov sense.

4.3.3 Controller of step 3

After the previous step only a forward motion along z-axis towards the target is left. We define the Lyapunov candidate function $V = V_z + V_\phi$. In the third step the robot moves with linear velocity $v_3 > 0$ and we have that $z < 0$ because the target is in front,

$$\dot{V}_z\ (14) \rightarrow (z - z^r) = z < 0 \Rightarrow\ As\ \ v > 0,\ \ \dot{V}_z < 0\ .$$

Although orientation has been corrected in the previous step, the angular velocity ω (10) is used during this step to compensate for odometry drift,

$$\dot{V}_\phi\ (15) \rightarrow \left\{ \begin{array}{l} \phi > 0 \rightarrow e_{cx} > 0 \rightarrow \nu_c < 0 \Rightarrow \omega < 0 \\ \phi < 0 \rightarrow e_{cx} < 0 \rightarrow \nu_c > 0 \Rightarrow \omega > 0 \end{array} \right\} \Rightarrow \dot{V}_\phi < 0\ .$$

Alternatively, we have defined a homography controller as another solution for the last step where the epipolar geometry is not well defined (11). The homography elements can be defined as a function of the state parameters as [12]:

$$v = k_v(1 - h_{33}) = k_v\,(1 - \cos\phi + (-x\sin\phi + z\cos\phi)n_z/d_\pi)\ ,$$

$$\omega = -k_\omega\, h_{13} = -k_\omega\,\alpha_x\,(\sin\phi + (x\cos\phi + z\sin\phi)n_z/d_\pi)\ ,$$

where the plane that generates the homography is defined by its normal $\mathbf{n} = (n_x, n_y, n_z)$ and the distance to the origin d_π. We define again the candidate Lyapunov function as $V = V_z + V_\phi$,

$$\dot{V}_\phi\ (15) \rightarrow x \simeq 0,\ z < 0,\ n_z < 0 \left\{ \begin{array}{l} \phi > 0 \Rightarrow \omega < 0 \\ \phi < 0 \Rightarrow \omega > 0 \end{array} \right\} \Rightarrow \dot{V}_\phi < 0\ .$$

$$\dot{V}_z\ (14) \rightarrow x \simeq 0,\ \phi \simeq 0,\ z < 0,\ n_z < 0 \Rightarrow v > 0 \Rightarrow \dot{V}_z < 0\ .$$

Thus $\dot{V} < 0$, and this alternative controller is also stable in the Lyapunov sense.

In summary, we have $\dot{V} < 0$ for each controller of the switching control law. We have also asymptotic stability given that $\dot{V}$ is negative definite for the controller defined in each step.

5 Experimental Evaluation

Simulated and real experiments have been performed to demonstrate the validity and performance of the approach. The size of both simulated and real images is 320×240 pixels.

5.1 Simulation Results

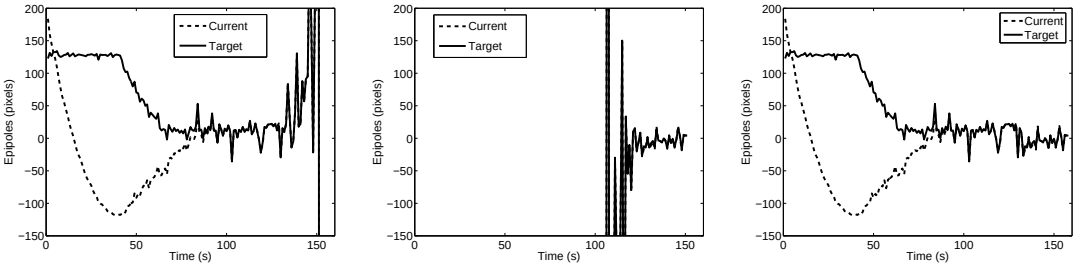
The simulated data consists of a synthetic scene of random points which are generated and projected to the image plane in each iteration of the control loop. In each iteration the epipolar geometry is computed from the point matches and the control law is used to send the velocities (v, ω) to the virtual robot.

The simulation presented in Fig. 5 considers as start position $(x, z, \phi) = (-2, -10, 5^\circ)$ and as target position $(0, 0, 0^\circ)$. This experiment has been performed adding white noise to the image points with a standard deviation of $\sigma = 0.3$ pixels. The evolution of the epipoles from the fundamental matrix between the current and target position is shown in Fig. 5(a). The computed epipoles using the auxiliary target and the values of the epipoles used in the control are shown in Fig. 5(b) and Fig. 5(c) respectively. It can be seen that switching from the fundamental matrix to the auxiliary one avoids the degenerate values of the epipoles due to the short baseline when reaching the target.

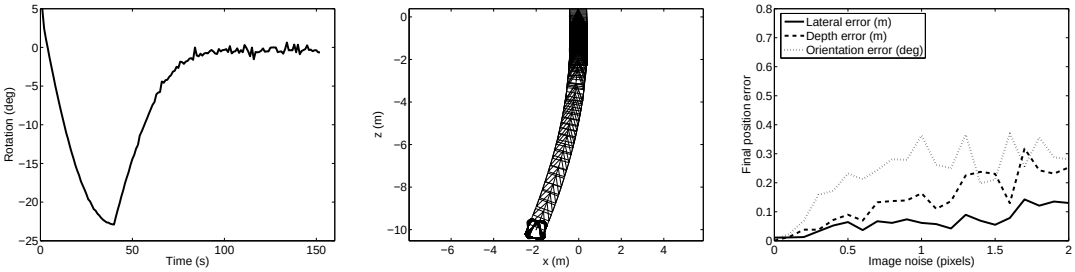
Simulations with different values of image noise have been performed to show the robustness of the control law. The final position errors obtained in the simulations are shown in Fig. 5(f) using standard deviation values from $\sigma = 0$ to 2 pixels. Results show that the control behaves properly with image noise.

The focal length of the camera (α_x) can be considered as a constant gain of each control law of the switching scheme, see (7), (8), (9) and (10). Therefore, due to the benefits of the feedback controller, is more than enough to know a rough approximation of α_x . Simulations show that the control law can cope with high errors in α_x. In the results shown in Fig. 7(a), the focal length used in the controller is fixed to $f = 6$ mm while simulations with different real focal length are carried out.

In our approach we have assumed that the principal point is in the center of the

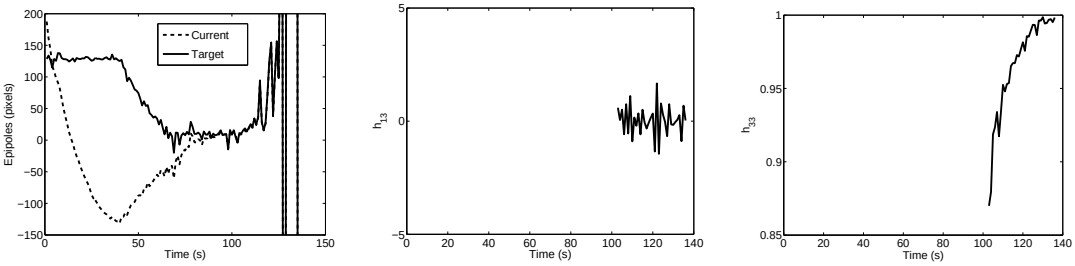


(a) Epipoles from F (b) Epipoles from F_{aux} (c) Epipoles for the control

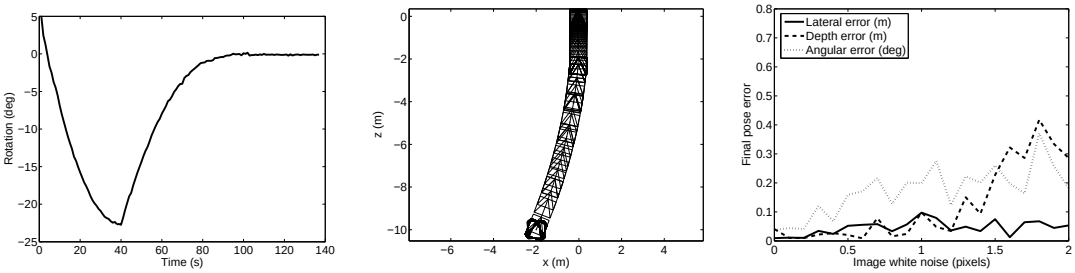


(d) Robot rotation (e) Path (f) Final error vs noise

Fig. 5. (a-e) Simulation with white image noise of $\sigma = 0.3$ pixels. (f) Final position error using different values of noise.

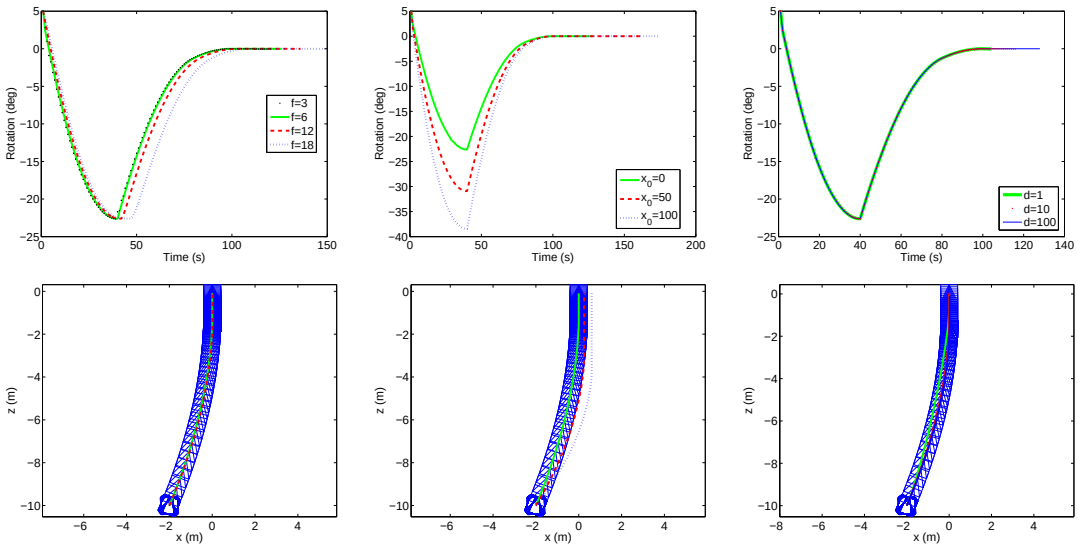


(a) Epipoles from F (b) Evolution of h_{13} (c) Evolution of h_{33}



(d) Robot rotation (e) Path (f) Final error vs noise

Fig. 6. (a-e) Simulation with homography-based control in the last phase carried out with white image noise of $\sigma = 0.3$ pixels. (f) Final position error using different values of noise.



(a) (b) (c)

Fig. 7. Simulations with: different values of the focal length (a), different x-coordinates of the principal point (b) and different values of parameter d (c). In each case, first and second row are the robot rotation and path respectively.

image. Simulations have been performed where the x-coordinate of the principal point is fixed to zero in the controller while its real value is modified (Fig. 7b). It can be seen that calibration error in the principal point produces a lateral error in the final position. This is because the desired epipoles at the end of the motion are the x-coordinate of the principal point; and therefore, an error on x_0, is transferred to the motion. However, the assumption of the principal point to be in the center of the image is a good supposition in practise since small deviations are observed.

In the second step of the controller we use the unknown parameter d (8). It can be considered again as a gain of the control and it is fixed to a constant value. The typical distances in our experiments are about meters, and we have selected $d = 2$ meters in our simulations and real experiments. To check the correct behaviour of the system against the fixed value of the unknown d, we have tested with values of d from centimeters to dozens of meters for the same initial real distance of 10 meters. The only difference obtained is a proportional variation in the time needed for completing the experiment, reaching in all the cases the desired target position (see Fig. 7c).

Experiments by using the homography-based control in the last step (11) are presented in Fig. 6. The homographies computed during the motion in the last phase are less noisy than the epipoles; this can be seen comparing the rotation obtained by both methods, (compare Fig. 5d and Fig. 6d). The evolution of the homography element h_{33} used to compute v is shown in Fig. 6(c), and its value reaches 1 as desired when the robot approaches to the target. The robot trajectory is a straight path and therefore no correction of the rotation is needed up to noise and drift,

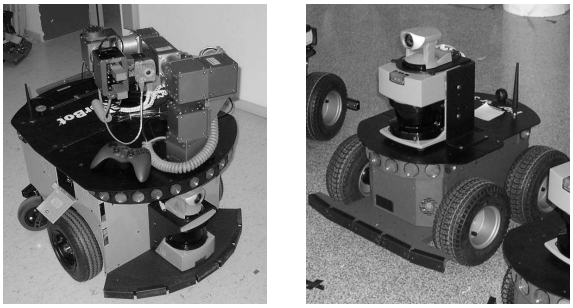


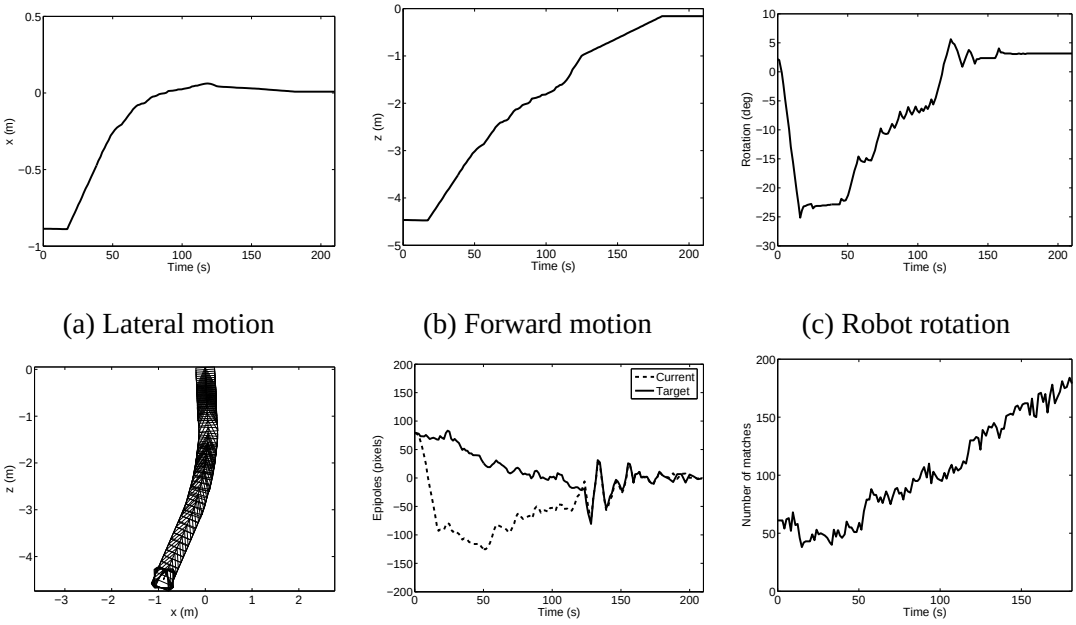
Fig. 8. The experimental platforms: ActiveMedia's PowerBot and Pioneer.

thus the element h_{13} used to compute ω fluctuate around its desired zero value. A disadvantage of performing this third step using the homography is the additional requirement of the homography computation.

5.2 Real World Experiments

We evaluated the proposed control law in two experimental platforms: a PowerBot and a Pioneer from ActivMedia (Fig. 8). Both robots have a nonholonomic differential drive base, with two rear caster wheels for the PowerBot and with the four driving wheels for the Pioneer. The robots are equipped with a Canon VC-C4 pan-tilt-zoom CCD camera mounted on top of a laser scanner. The platforms are controlled by sending translation and rotational velocities (v, ω) given by the switching control law. The vision control loop has not yet been optimized and it is quite slow, therefore, the velocities have to be quite low. The same implementation of the control was used on both platforms. The experiments have been performed in indoor and outdoor realistic settings. It has to be mentioned that no particular knowledge of the environment, other than the image from the desired pose, is provided. The camera calibration parameters have been set without performing the calibration to f = 6 mm for the focal length and the image center for the principal point.

Results from one of the experimental runs are shown in Fig. 9. The lateral distance to the target (a) and depth distance (b) shows the robot position evolution and (c) shows the rotation. The lateral distance to the target, which is specially critical with nonholonomic motion constraints, is corrected in an asymptotic convergence (d). The definition of the desired trajectories of the epipoles in each step requires the selection of the values of T_1 and T_3. The value of T_2 is defined respect T_3 with a few seconds of difference ($T_2 = T_3 - 10$ s). These parameters determine the time for the control to track the desired input trajectories. The minimum time required for the motion is constrained by the maximum velocities of the robot and the control loop time period. The selection of these parameters is not critical having a wide margin. In the experiments we have defined different values for the indoor and outdoor experiments. Outdoor experiments require higher values of these parameters as the distances covered are higher. The evolution of the epipoles is shown



(a) Lateral motion (b) Forward motion (c) Robot rotation

(d) Path (e) Epipoles evolution (f) Number of SIFT matches

Fig. 9. Real experiment with target position at $(0, 0, 0^\circ)$. Position data shown in (a-d) has been obtained from robot odometry.

in (e) with a similar behavior with respect to the simulations. It can be seen that the epipoles have proper values, except noise, avoiding the problem of degeneracy due to the short baseline. The number of SIFT matches found along the motion is shown in (f). As expected, the number of matches increases with the similarity of the images as the robot moves towards the target. If the number of SIFT points is too low, for example because a scene with poor information, the epipole estimation accuracy decreases. If this situation keeps along several iterations the control may fail, for example when the common scene between the current and target image leaves the camera field of view. Fig. 10 shows the evolution of an indoor and outdoor experiment with some of the images taken during the navigation. The distance between initial and target position is 3 m and 5 m respectively in the indoor and outdoor experiments of Fig. 10. The final error given by the odometry is $(-0.003, -0.054, -2.81^\circ)$ and $(-0.047, -0.105, -0.04^\circ)$ respectively. The final image of each experiment, compared with the target image, allows to see the correctness of the motion control.

In the approach presented we consider that there are no obstacles between the initial and target positions. The control itself is not designed for detecting obstacles. If an obstacle appears in the middle of the robot path, an additional procedure would be needed to avoid the obstacle (for example a reactive potential field routine). After the routine ends and the obstacle is avoided the controller takes control of the robot again toward the target.

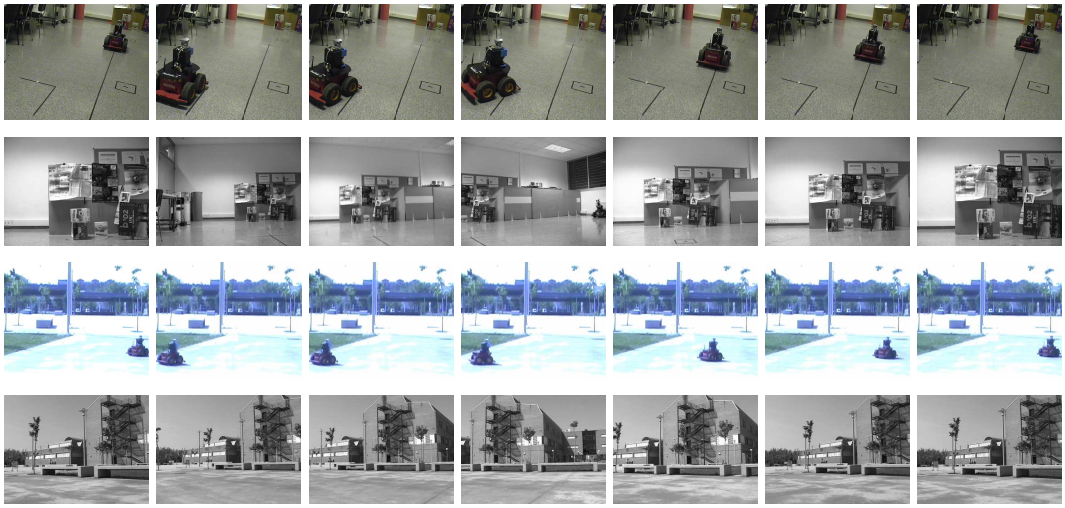


Fig. 10. Sequences of some images taken by the robot camera during an indoor (2^{nd} row) and outdoor (4^{th} row) experiment. Rows 1 and 3 are the corresponding images taken by an external video camera. In each row the first is the target image, the second is the initial and the last is the image at the end of the motion. Videos of these experiments are available in http://webdiis.unizar.es/∼csagues/videos.html/

6 Conclusion

In this paper, we have presented a visual servoing scheme for mobile platforms based on the epipolar geometry. The approach designed consists of a sequential switching control law obtained from the input-output linearization of the system. The resultant path followed by the robot is due to the specific reference trajectories defined for both epipoles in the current and target images. The estimation of the epipolar geometry is performed through an automatic robust feature detection and matching process. The contributions are that the control scheme presented deals with the motion constraints of the platform and the robot is driven in a direct motion towards the target using as image information the epipoles. The presented approach does not require complete camera calibration or any particular knowledge about the environment. The stability of the switching control law in the Lyapunov sense has been analyzed. Real experiments have been performed in realistic indoor and outdoor settings to show the validity of the approach. The epipolar based control is designed to avoid degeneracy due to short baseline when the robot comes closer to the target. Additionally, an alternative solution is proposed to avoid this problem by switching between the fundamental matrix and the homography model.

A Geometric Relations Between the System State and the Epipoles

The derivation of the geometric relations (1) and (2) is detailed next. The goal is to relate the epipoles (e_{cx}, e_{tx}) with the state of the system (x, z, ϕ). The geometric relations can be deduced from Fig. 2 using the idea of similar triangles. The expression for the target epipole e_{tx} is straightforward from Fig. 2: The right-angled triangles with sides $\triangle(x,z)$ and with sides $\triangle(e_{tx}, \alpha_x)$ are similar and then it is possible to deduce proportionalities between corresponding sides of the two triangles:

$$\frac{e_{tx}}{\alpha_x} = \frac{x}{z} \,,$$

which leads to (2). The same procedure is used to obtain the expression for the current epipole e_{cx}. In this case the similar triangles are $\triangle(e_{cx}, \alpha_x)$ and $\triangle(x\cos\phi + z\sin\phi, z\cos\phi - x\sin\phi)$. The sides of the similar triangles used from Fig. 2 are drawn in bold. Then we obtain the expression that leads to (1),

$$\frac{e_{cx}}{\alpha_x} = \frac{x\cos\phi + z\sin\phi}{z\cos\phi - x\sin\phi} \,.$$

B Input-output Linearization of the System

The derivation of the expressions (5) and (6) obtained after the input-output linearization is detailed next. We first differentiate the expressions that relate the epipoles with the state of the system. So, for the epipole in the current image according to (1), we have

$$\dot{e}_{cx} = \frac{de_{cx}}{dt} = \frac{d}{dt}\left(\alpha_x \frac{x\cos\phi + z\sin\phi}{z\cos\phi - x\sin\phi}\right) \,.$$

Denoting $s = \sin\phi$ and $c = \cos\phi$, the result of the derivation is

$$\dot{e}_{cx} = \alpha_x \frac{(\dot{x}c - x\dot{\phi}s + \dot{z}s + z\dot{\phi}c)(zc - xs) - (\dot{z}c - z\dot{\phi}s - \dot{x}s - x\dot{\phi}c)(xc + zs)}{(z\cos\phi - x\sin\phi)^2} \,.$$

Computing products and simplifying we have

$$\dot{e}_{cx} = \alpha_x \frac{\dot{x}z - x\dot{z} + \dot{\phi}(x^2 + z^2)}{(z\cos\phi - x\sin\phi)^2} \,.$$

Using (3), it follows

$$\dot{e}_{cx} = \alpha_x \frac{-v\sin\phi z - xv\cos\phi + \omega(x^2+z^2)}{(z\cos\phi - x\sin\phi)^2} \quad ,$$

and using (4),

$$\dot{e}_{cx} = \alpha_x \frac{-v\sin\phi\, d\cos\psi + d\sin\psi\, v\cos\phi + \omega d^2}{(d\cos\psi\cos\phi + d\sin\psi\sin\phi)^2} = \alpha_x \frac{v(-\sin\phi\cos\psi + \cos\phi\sin\psi)/d + \omega}{(\cos\psi\cos\phi + \sin\psi\sin\phi)^2} \quad .$$

Finally using trigonometry, it turns out the expression (5),

$$\dot{e}_{cx} = -\frac{\alpha_x \sin(\phi-\psi)}{d\cos^2(\phi-\psi)} \; v + \frac{\alpha_x}{\cos^2(\phi-\psi)} \; \omega \quad .$$

Similarly for the epipole in the target image (2) we differentiate as follows

$$\dot{e}_{tx} = \frac{de_{tx}}{dt} = \frac{d}{dt}\left(\alpha_x \frac{x}{z}\right) = \alpha_x \frac{\dot{x}z - x\dot{z}}{z^2} \quad .$$

Using (3) we have

$$\dot{e}_{tx} = \alpha_x \frac{-v\sin\phi}{z} - \alpha_x \frac{xv\cos\phi}{z^2} \quad ,$$

and using (4) it follows

$$\dot{e}_{tx} = \alpha_x \frac{-v\sin\phi}{d\cos\psi} + \alpha_x \frac{d\sin\psi\, v\cos\phi}{d^2\cos^2\psi} = \alpha_x \frac{v}{d}\left(\frac{-\sin\phi\cos\psi + \cos\phi\sin\psi}{\cos^2\psi}\right) \quad .$$

Finally using trigonometry, it turns out the expression (6),

$$\dot{e}_{tx} = -\frac{\alpha_x \sin(\phi-\psi)}{d\cos^2\psi} \; v \quad .$$

Acknowledgment

The authors would like to thank F. Kock for his contribution in the experiments and D. Lowe for the SIFT implementation.

References

[1] R. Basri, E. Rivlin, and I. Shimshoni. Visual homing: Surfing on the epipoles. *International Journal of Computer Vision*, 33(2):117–137, 1999.

[2] 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*, pages 2173–2178, April 2005.

[3] G. Blanc, Y. Mezouar, and P. Martinet. Indoor navigation of a wheeled mobile robot along visual routes. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3365–3370, April 2005.

[4] M.S. Branicky. Multiple lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):475–482, April 1998.

[5] K. Deguchi. Optimal motion control for image-based visual servoing by decoupling translation and rotation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 705–711, 1998.

[6] Y. Fang, W.E. Dixon, D.M. Dawson, and P. Chawda. Homography-based visual servo regulation of mobile robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 35(5):1041–1050, 2005.

[7] N.R. Gans and S.A. Hutchinson. An asymptotically stable switched system visual controller for eye in hand robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 735–742, Las Vegas, Nevada, October 2003.

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

[9] S. Hutchinson, G.D. Hager, and P.I. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5):651–670, 1996.

[10] A. Isidori. *Nonlinear Control Systems*. Springer, 1995.

[11] V. Kyrki, D. Kragic, and H.I. Christensen. New shortest-path approaches to visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 349–354, Sendai, Japan, September 2004.

[12] G. López-Nicolás, C. Sagüés, and J. J. Guerrero. *Vision Systems*, chapter 30, pages 583–596. Edited by Goro Obinata and Ashish Dutta, I-Tech Education and Publishing, Vienna, Austria, 2007.

[13] G. López-Nicolás, C. Sagüés, J.J. Guerrero, D. Kragic, and P. Jensfelt. Nonholonomic epipolar visual servoing. *IEEE International Conference on Robotics and Automation*, pages 2378–2384, 2006.

[14] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[15] Y. Ma, J. Kosecka, and S. Sastry. Vision guided navigation for a nonholonomic mobile robot. *IEEE Transactions on Robotics and Automation*, 15(3):521–537, 1999.

[16] E. Malis and S. Benhimane. A unified approach to visual tracking and servoing. *Robotics and Autonomous Systems*, 52:39–52, 2005.

[17] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 D visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):234–246, April 1999.

[18] E. Marchand and F. Chaumette. Feature tracking for visual servoing purposes. *Robotics and Autonomous Systems*, 52:53–70, 2005.

[19] G.L. Mariottini, D. Prattichizzo, and G. Oriolo. Epipole-based visual servoing for nonholonomic mobile robots. *IEEE International Conference on Robotics and Automation*, pages 497–503, 2004.

[20] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 257–263, June 2003.

[21] P. Rives. Visual servoing based on epipolar geometry. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 602–607, 2000.

[22] C. Sagüés and J.J. Guerrero. Visual correction for mobile robot homing. *Robotics and Autonomous Systems*, 50(1):41–49, 2005.

[23] J.J. E. Slotine and W. Li. *Applied nonlinear control*. Prentice Hall, Englewood Cliffs NJ, 1991.

[24] P. H. S. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):35–61, 2002.