

Epipole-based visual servoing for mobile robots

G. CHESI*, G. L. MARIOTTINI, D. PRATTICHIZZO and A. VICINO

Department of Information Engineering, University of Siena, Via Roma 56, 53100 Siena, Italy

Received 7 December 2004; accepted 14 February 2005

Abstract—This paper proposes a visual servoing algorithm for mobile robot navigation based on the epipolar geometry retrieved by object profiles. The main motivation for this approach is that for unstructured scenes the task of solving correspondences is definitely a harder task than contour detection. An unstructured three-dimensional (3-D) scene consists mainly of objects whose most important 2-D features are their apparent contours in the image plane. Apparent contours are used to estimate the positions of the epipoles and some special symmetry conditions whereby the visual servoing is able to steer the mobile robot to the desired position.

Keywords: Mobile robot; object profile; epipolar geometry; visual servoing.

1. INTRODUCTION

The problem of controlling the pose of a mobile robot with respect to a target object by means of visual feedback is investigated. Visual servoing has been applied recently to mobile robotics, see e.g., Refs [1–4]. In visual servoing, the control goals and the feedback law are directly designed in the image domain. Designing the feedback at the sensor level increases system performance especially when uncertainties and disturbances can affect the robot model and camera calibration [5, 6].

Visual servoing algorithms make use of object cues whose image plane projections are controlled to desired positions through the visual servoing process. Usually, these cues are distinctive textures of objects in the three-dimensional (3-D) scene, like corners or other easily recognizable points, see e.g., Refs [7–13]. However, it may happen that the 3-D scene does not exhibit any appropriate texture, but only smooth surfaces whose main features consist of their apparent contours, defined as the projection of the contour generators of objects surfaces [14]. As pointed out in Ref. [15], if the object surface does not have any noticeable texture,

*To whom correspondence should be addressed. E-mail: chesi@dii.unisi.it

the object profile is the only information available to estimate the structure of the surface and the motion of the camera.

The aim of this work is to exploit object profiles to synthesize a visual servoing algorithm. It is worthwhile to notice that, in general, tracking object profiles instead of textures can be performed in a more robust way since solutions of correspondence problems are not required. Moreover, it appears more appropriate to exploit profiles in outdoor navigation where objects in the scene are highly unstructured (hills, trees, etc.) because, in these contexts, solving correspondences is a difficult task which usually gives rise to poor results. This work builds upon preliminary results presented in Refs [16, 17].

Apparent contours of planar objects have been used as object cues for visual servoing in Refs [18, 19]. Specifically, the camera motion is recovered by matching the initial and final viewed contours, and estimating the homography matrix between them. Then, 2.5-D visual servoing is used to steer the camera to the desired position by exploiting the recovered camera motion. Planar objects are considered also in Ref. [20] which proposes a visual servoing based on 3-D reconstruction. In Ref. [21], a strategy to extract points as object cues from more natural images is presented and used to realize 2.5-D visual servoing.

The epipolar geometry plays a key role in this work. In fact, at each time step the apparent contours, of either planar or non-planar objects, extracted in both actual and target views are used to estimate the position of epipoles in the image for the planar motion case. Then special symmetry conditions on epipoles are used to design the visual servoing algorithm in order to make a holonomic mobile robot able to reach the desired position (where the target view was grabbed).

For a complete analysis of the epipolar geometry the reader is referred to Ref. [22] and references therein. In Ref. [23], a visual servoing strategy for uncalibrated cameras based on epipolar geometry is also presented. The authors assume that the correspondences problem is solved, and use feature points and epipoles (in at least three views) to drive the robot to the target pose.

According to the classification of visual servoing systems, presented in Ref. [6], the approach used in this paper is known as image-based visual servoing, because the error between the robot pose and a target object or a set of target features is computed directly from the image features.

The present paper deals with holonomic mobile robots moving on a plane. Several results are present in the literature on non-holonomic mobile robots. Consider, for instance, Ref. [24] where the task of navigating a non-holonomic mobile robot tracking an arbitrary shaped continuous ground curve is considered or Ref. [2, 4], where a visual servoing controller is proposed for set point stabilization of a bunch of image point features.

However, to the best of our knowledge, the confluence between epipolar geometry and mobile robotics visual servoing is far from being fully understood and the simplicity of results achievable for the holonomic case appears to be important at

this rather early stage of investigation on visual servoing based on epipolar geometry and object contours.

The proposed technique considers only scenes with static objects. In order to analyze dynamic scenes in the epipolar geometry setting, a possible approach could be that of using the multibody epipolar constraints and its associated multibody fundamental matrix as proposed in Ref. [25]. Future work will investigate the extension of the proposed technique to scenes with moving objects. Particular attention will be devoted to the image sampling rate that can affect the performance of the visual-servo controller.

The paper is organized as follows. Section 2 presents the model of the robotic system. Section 3 summarizes the epipolar geometry between two views of the same scene, focusing on the fundamental matrix and epipoles. Symmetric conditions of the epipoles for the planar motion case are analyzed for some special configurations. Estimation of the epipoles is discussed in Section 4 and the epipole-based visual servoing is presented in Section 5. All theoretical results are validated by simulations and experiments in Section 6. Concluding remarks are given in Section 7.

2. VISUAL MODELING

Consider a mobile robot moving on a plane with a fixed pin-hole camera mounted on it whose image plane is perpendicular to the motion plane. Refer to Fig. 1 and let z_a be the optical axis of the camera–robot frame $\langle a \rangle$. The configuration space of the mobile robot (or of the camera) is $R^2 \times SO(2)$, where $SO(2)$ is the special orthogonal group of 2×2 rotation matrices. Let $(X_a \ Y_a \ 0)^T$ be the camera center position in the base frame $\langle b \rangle$ and α_a be the rotation angle of the camera–robot with respect to the x -axis of the base frame.

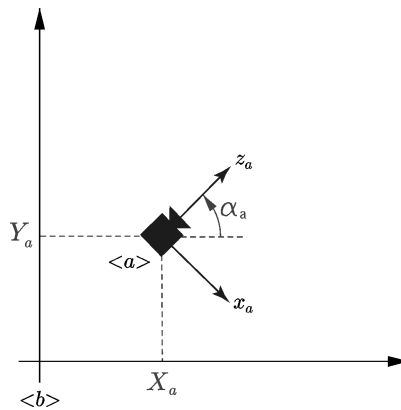


Figure 1. Mobile robot with a fixed camera.

A fully actuated holonomic mobile robot is considered:

$$\begin{cases} \dot{X}_a = u_x \\ \dot{Y}_a = u_y \\ \dot{\alpha}_a = \omega, \end{cases} \quad (1)$$

where $(u_x, u_y)^T$ is the linear velocity of the camera–robot on the plane and ω is the angular velocity.

At each time instant a couple of images are available to the visual servo loop, i.e., the actual and the desired images. Recall that the only data available to autonomously move the robot is the visual information contained in these two images.

The proposed visual servoing algorithm has the aim of moving the robot from an actual to the desired (target) pose, and is based on the epipolar geometry existing between the pair of images grabbed at the actual and the target positions.

Henceforth we will refer to the actual and the desired camera as the robot–camera in the actual and desired position, respectively. A pin-hole camera model is assumed for the camera–robot whose intrinsic matrix is:

$$K = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

where (u_0, v_0) is the principal point, $f_u = f\kappa_u$ and $f_v = f\kappa_v$ being f the focal length and κ_u and κ_v the CCD scaling factors.

3. EPIPOLAR GEOMETRY

The epipolar geometry plays a key role in this work [22, 26]. The proposed visual servoing technique uses object apparent contours to estimate the epipole coordinates in order to detect special symmetry conditions of the epipolar geometry for the planar motion case, whereby the mobile robot is steered to the final target.

The main ideas of epipolar geometry [14, 27] are recalled here. Refer to Fig. 2 and consider the actual and desired cameras with optical centers c_a and c_d , optical axes z_a and z_d , respectively. The segment $\overline{c_a c_d}$ is referred to as the baseline and its intersections with both the image planes define the epipoles e_a and e_d . Any plane containing the baseline is referred to as an epipolar plane.

Given a pair of views of a scene and a set of corresponding image points representing the projection of the same point P in the space, u_a, u_d in homogeneous coordinates (see Fig. 2), there exists a matrix $F \in \mathbb{R}^{3 \times 3}$, referred to as the fundamental matrix [27], such that:

$$u_d^T F u_a = 0.$$

The fundamental matrix has rank 2 and is defined up to an arbitrary scaling factor.

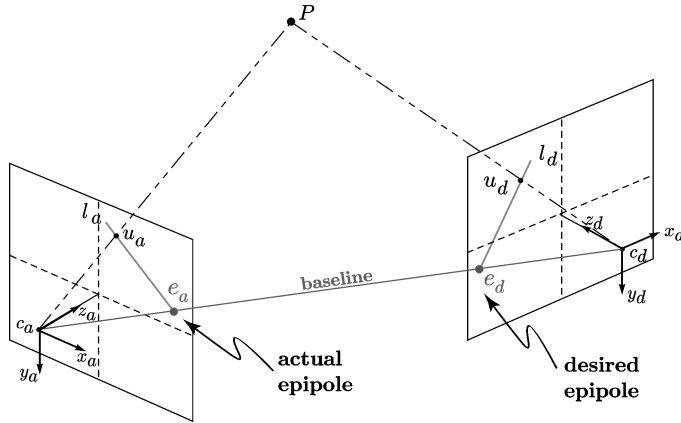


Figure 2. The epipolar geometry between two views.

For any point u_a (u_d) in one view the equation $l_d = F u_a$ ($l_a = F^T u_d$) defines a line in the other view (represented by a homogeneous 3×1 vector), called an epipolar line, such that the corresponding point u_d (u_a) belongs to this line. Moreover, the epipole e_a (e_d) lies on the null right 1-D space of F (F^T).

The fundamental matrix depends on the camera parameters and relative pose (translation and rotation) of the two views and is given by:

$$F = K^{-T} E K^{-1},$$

K being the matrix in (2) and E the essential matrix:

$$E = [t]_{\times} R,$$

with $[t]_{\times}$ the skew matrix of translation vector t and R the rotation matrix between the two views [28].

It is useful to recall that the geometric distance between an image point u_a and the epipolar line, $l_a = F u_d$, is given by:

$$\sqrt{\frac{(u_a^T F^T u_d)^2}{(F^T u_d)_1^2 + (F^T u_d)_2^2}}, \tag{3}$$

$(F^T u_d)_1$ and $(F^T u_d)_2$ being the first and the second components of the epipolar vector $(F^T u_d)$ [26]. Note that if points u_a and u_d correspond to the same feature in the tridimensional scene then this distance is zero. Most of the algorithms used to estimate the epipolar geometry are based on distance (3) and an exhaustive analysis of these techniques is reported in Ref. [22].

The analysis of the epipole positions when the actual and target cameras are in some special relative configurations is paramount for the design of the epipole-based visual servoing.

3.1. Symmetric case

Refer now to the planar motion case as described in Section 2. Assume that two images are taken by the same camera, which undergoes a rotation θ about the axis O as shown in Fig. 3.

The optical centers c_a and c_d are displaced at the same distance $r_{act} = r_{des}$ from the intersection point o of the two optical axes. Moreover, image planes and camera rotation axes are perpendicular to the epipolar plane containing o .

For pure rotations of the camera about the O -axis, the essential matrix becomes (up to a scale factor):

$$E = \begin{bmatrix} 0 & \cos \theta - 1 & 0 \\ \cos \theta - 1 & 0 & \sin \theta \\ 0 & -\sin \theta & 0 \end{bmatrix},$$

and if the calibration matrix is equal to the identity ($u_0 = v_0 = 0, f_u = f_v = 1$), it ensues that $F = E$ where θ is the angle between the optical axes z_a and z_d . In this case epipoles e_a and e_d , which lie in the nullspace of the fundamental matrix and its transpose, respectively, are computed as:

$$\begin{aligned} e_a &= \left[\frac{1}{\tan(\theta/2)}, 0, 1 \right]^T, \\ e_d &= \left[-\frac{1}{\tan(\theta/2)}, 0, 1 \right]^T, \end{aligned} \quad (4)$$

and the following remark applies.

Remark 1. For circular planar displacements of actual and desired camera positions, as in Fig. 3, a special symmetry condition holds: the u -coordinate of the two epipoles in (4) exhibits the same modulus and opposite sign. Moreover, the angular coefficients of the two epipolar lines tangent to the contour have the same modulus and opposite sign.

Such symmetry condition will play a key role in designing the visual servoing algorithm.

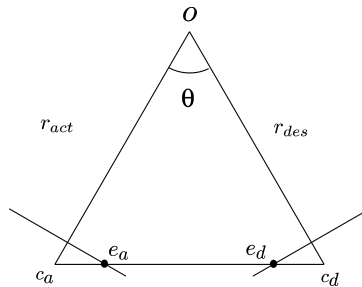


Figure 3. Symmetric camera displacement.

3.2. Asymmetric case

Symmetry is not preserved in the general configuration shown in Fig. 4 where a general rotation and translation occurs between the actual camera position c_a and the desired camera position c_d . To fix notation, assume that, with respect to the configuration of Fig. 3, the camera c_a has been shifted along its optical axis by a distance $\tilde{r} = r_{act} - r_{des}$. In this case, the fundamental matrix F becomes (for $K = I$):

$$F = \begin{bmatrix} 0 & -\beta & 0 \\ \beta \cos \theta - \gamma \sin \theta & 0 & \gamma \cos \theta + \beta \sin \theta \\ 0 & -\gamma & 0 \end{bmatrix},$$

where:

$$\begin{aligned} p &\triangleq \frac{\tilde{r} + r_{des}}{r_{des}} = \frac{r_{act}}{r_{des}}, \\ \beta &\triangleq \frac{1}{p} - \cos \theta, \\ \gamma &\triangleq \sin \theta, \end{aligned}$$

and the epipoles are given by:

$$\begin{aligned} e_a &= (\alpha_a, 0, 1)^T, \\ e_d &= (\alpha_d, 0, 1)^T, \end{aligned} \quad (5)$$

with:

$$\begin{aligned} \alpha_a &= \frac{-\sin(\theta)}{\frac{1}{p} - \cos(\theta)}, \\ \alpha_d &= \frac{\sin(\theta)}{p - \cos(\theta)}. \end{aligned} \quad (6)$$

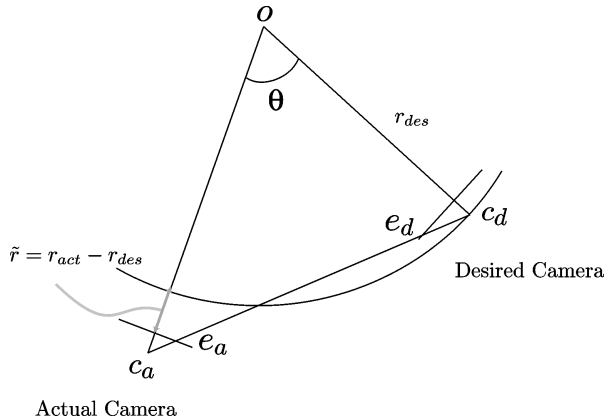


Figure 4. General camera displacement.

For the general roto-translation displacement between the desired and actual camera positions, the symmetric properties stated in Remark 1 does not apply anymore. Detecting symmetric and asymmetric configurations will be the key feature of the proposed visual servoing algorithm.

The assumption of $K = I$ has been introduced for clarity of presentation and does not affect the generality of results. In fact, symmetry and asymmetry can be easily extended to general cameras as shown in the following remarks.

Remark 2. For a general intrinsic matrix K in (2) with any scaling factors f_u and f_v and principal point (u_0, v_0) , the epipoles are obtained scaling α_a and α_d in (6) by the positive factor f_u and translating the first two coordinates by the principal point (u_0, v_0) as:

$$\begin{aligned} e_a &= (f_u \alpha_a + u_0, v_0, 1)^T, \\ e_d &= (f_u \alpha_d + u_0, v_0, 1)^T. \end{aligned} \quad (7)$$

Remark 3. For general intrinsic camera matrices K , symmetry still holds for the u -coordinates of the epipoles, provided that they are referred to the principal point. In other terms, symmetry holds for epipole displacements so defined:

$$\begin{aligned} \delta e_a &= (e_a)_1 - u_0 = f_u \alpha_a, \\ \delta e_d &= (e_d)_1 - u_0 = f_u \alpha_d. \end{aligned} \quad (8)$$

Henceforth, the camera will be supposed to be partially uncalibrated. The visual servoing procedure will be designed for partially unknown intrinsic camera matrices K . The only parameters which are supposed to be known are the coordinates (u_0, v_0) of the principal point. In fact, once the first coordinate of the epipole is known, only the principal point is needed to evaluate the epipoles displacements in (8), which are the key parameters of the visual servoing design.

4. ESTIMATION OF THE EPIPOLES

Estimation of epipoles for the current and desired images is the core procedure of the proposed visual servoing strategy. Note that, for planar camera–robot motions, the second coordinate of the epipoles is always equal to that of the principal point which is assumed to be known.

Many procedures are available in the literature to estimate the epipolar geometry [22, 26]. The focus here is on estimating the epipoles starting from two apparent contours of the same object. Many techniques dealing with apparent contour are available in the literature [14, 15, 29, 30]. In Refs [14, 30], the authors propose a reconstruction algorithm from unknown translational camera motion. They use the concept of the frontier point (i.e. the 3D point at the intersection of both contour generators in two views) (see Fig. 5) to retrieve the epipolar geometry *via* the Hough transform. This estimation has been also generalized to other camera motions in Refs [15, 29].

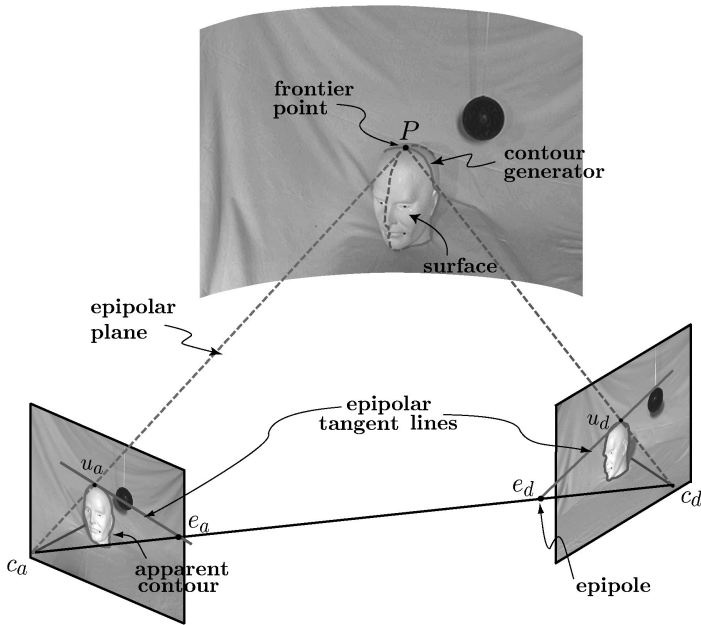


Figure 5. Frontier point, apparent contours and epipolar tangent lines.

This section shows how these techniques can be employed to design visual servoing for mobile robots according to the general parametrization of epipolar geometry proposed in Section 3.

The epipole estimation phase is based on contour detection of the objects in the scene. An interesting property of the epipolar geometry is that if the epipolar plane is tangent to the object surface, then the corresponding epipolar lines are tangent to the apparent contours in both the actual and desired images (see Fig. 5), and they are called epipolar tangents. In other terms, finding the epipolar lines tangent to the object contours in the two images allows us to straightforwardly solve the point correspondence problem for a special point P on the surface of the object, i.e., the frontier point.

Moreover, the perspective projection of the frontier point onto the image plane belongs to the apparent contour and satisfies the epipolar tangency constraints [14].

Let obj be an object of the scene whose frontier point neighborhood is visible and not occluded in both the current and desired views. Moreover, assume that the object exhibits smooth apparent contours which are retrieved by means of segmentation and B-splines representation [14]. Two different cases are in order: the symmetric case (Fig. 3) and the asymmetric case (Fig. 4).

4.1. Symmetric case

Assume that only a rotation about the O -axis occurs between the current and the desired views as shown in Fig. 3, so that the distances from the intersection of the

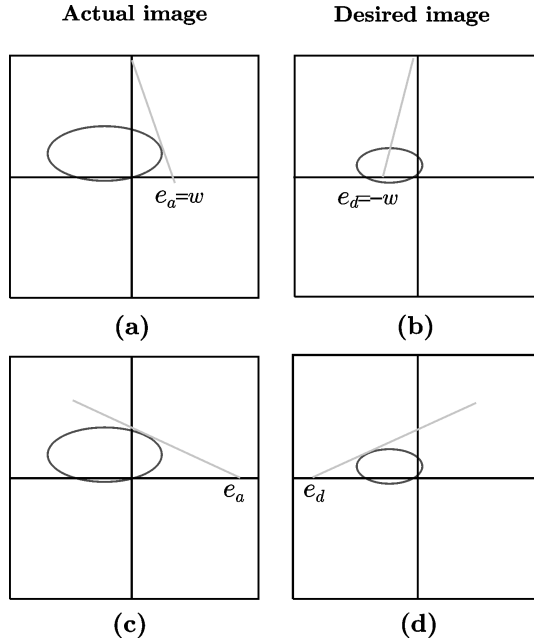


Figure 6. Epipolar tangent lines for the symmetric camera displacement case. Exact (lower) and wrong (upper) epipole estimation.

optical axes to the optical centers are equal, $r_{act} = r_{des}$ (i.e. $p = 1$ in (5)). According to (6) and (7), the displacement of epipoles with $p = 1$ is symmetric with respect to the principal point:

$$\delta e_a = w, \quad \delta e_d = -w$$

being:

$$w = f_u \frac{\sin(\theta)}{1 - \cos(\theta)} = \frac{f_u}{\tan(\theta/2)}.$$

In the symmetric case, the epipole motion exhibits only 1 d.o.f., the parameter w , and Remark 1 suggests a strategy for estimating the epipole displacement w .

Consider the setup depicted in Fig. 6a–d where an object in the scene is being viewed in both actual and desired image planes. Start with an initial guess for the epipole displacement w , then draw a tangent to the object apparent contour passing through this epipole guess in one view, e.g. the current view (Fig. 6a). In the desired view, the line characterized by the opposite angular coefficient and passing through the epipole guess in the second view (Fig. 6b) would be tangent to the object apparent contour in the desired view only if the guess for w was exact, i.e. only if the position of the epipoles were exact as shown in the example of Fig. 6c and d.

The above description suggests estimating the epipole position w by solving the optimization problem:

$$\min_w [\text{dist}(l_a(w), AC_a) + \text{dist}(l_d(-w), AC_d)], \quad (9)$$

where $l_a(-w)$ and $l_d(w)$ are the corresponding epipolar lines depending on the epipole positions, AC_a and AC_d are the apparent contours of the object obj , and $\text{dist}(\cdot, \cdot)$ is the distance between a line and a contour. In other terms, problem (9) is solved at each iteration by w and $-w$ such that the corresponding epipolar tangent in the current view is also an epipolar tangent in the desired view (Fig. 6). In this way, the exact position of actual and desired epipoles is retrieved.

4.2. Asymmetric case

Consider the asymmetric case of Fig. 4 where the optical centers c_d and c_a are displaced at distances r_{des} and $r_{\text{act}} = r_{\text{des}} + \tilde{r}$ from the intersection of the two optical axes, respectively. In this case, the epipole displacements (8) are not symmetric. Unlike the symmetric case, here at least two different tangency conditions are needed to estimate the positions of epipoles.

The relationship between the tangent lines of the current and desired images is more involved than the symmetric case. In the asymmetric case, more than one couple of epipolar tangents must be considered. Refer for instance to the upper and lower epipolar tangents to the object in the scene (Fig. 7c and d). Let γ'_1, γ'_2 be the

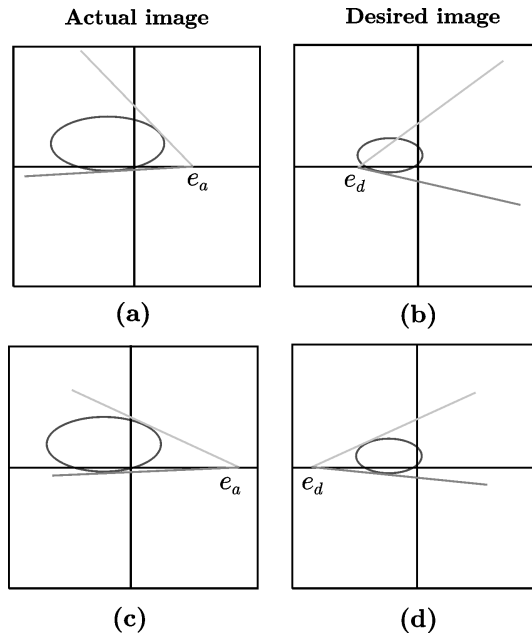


Figure 7. Epipolar tangent lines for the asymmetric displacement case. Exact (lower) and wrong (upper) estimation.

angles between the epipolar tangents and the horizon line in the current view, and let γ_1, γ_2 be the corresponding angles in the desired view. It can be easily shown [14] that, for exact epipole displacements, the following constraint holds

$$\frac{\tan(\gamma'_1)}{\tan(\gamma'_2)} = \frac{\tan(\gamma_1)}{\tan(\gamma_2)}. \quad (10)$$

Therefore, the epipole positions can be estimated by solving the 2-d.o.f. optimization problem:

$$\min_{w', w} [\text{dist}(l'_u(w'), C') + \text{dist}(l'_d(w'), C') + \text{dist}(l_u(w), C) + \text{dist}(l_d(w), C)], \quad (11)$$

where $l'_u(\cdot), l'_d(\cdot)$ are the epipolar lines in the current view, and $l_u(\cdot), l_d(\cdot)$ are the epipolar lines in the desired view. The main idea behind the process is to start from an initial guess of w and w' , then build three of the four tangents (two in one view and one in the other view), and, according to (10), draw the last line and finally compute the distance of this line from the apparent contour. This procedure is then repeated a further 3 times using three different initial tangents in order to compute the cost function in (11). The goal is to find the values for w and w' that minimize the cost function in (11). From a numerical point of view, better results are obtained by using more than two couples of epipolar tangents which means that the scene should consist of more than one object. In the experiment, presented in Section 6, two objects have been considered.

In Ref. [31], the some authors proposed a novel epipolar geometry estimation algorithm from apparent contours that builds upon the basic idea underlying (9) and (11). This algorithm makes use of the same parametrization proposed in Section 3. Experimental results are also provided on general shaped 3D object.

5. EPIPOLE-BASED VISUAL SERVOING

Procedures discussed in Section 4 allow us to estimate the epipole displacements in (8) that are here used to synthesize the proposed visual servoing to steer the robot from the initial position to the goal. It will be shown that detecting the symmetry condition, instead of retrieving the overall epipolar geometry, is sufficient to control the robot motion.

The visual servoing steers the camera–robot from c_a along a trajectory consisting of a translation along the optical axis and a rotation about the axis through O and toward c_d , as shown in Fig. 8. In particular, the visual servoing algorithm steers the robot along this trajectory in two steps:

- (i) The robot starts translating along the optical axis to reach a distance from O equal to r_{des} , i.e. to get $\tilde{r} = 0$ (Fig. 8).
- (ii) The robot moves to the desired position with a circular motion. Actually, the trajectory would be circular only if the radius r_{des} were known, otherwise the trajectory tends to a circular motion as it will be discussed in Section 5.2.

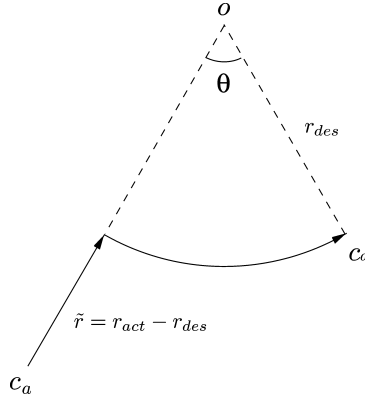


Figure 8. Trajectory followed by the robot in the case of general planar motion.

Observe that the first step does not require any knowledge of the actual or desired radius. In fact, let:

$$m_a = \frac{1}{f_u \alpha_a} + \frac{1}{f_u \alpha_d} = \frac{1}{\delta e_a} + \frac{1}{\delta e_d}$$

be defined as the sum of the inverse of the u -coordinates of the two epipole displacements in (8). The parameter m_a is a measure of the asymmetric part of the displacements between the two views. Indeed, the following relationships hold:

$$\begin{aligned} m_a = 0 &\iff \tilde{r} = 0 \\ m_a \tilde{r} &\leq 0. \end{aligned} \tag{12}$$

Quantity m_a can hence be exploited to steer the robot in the first step, e.g., through the control law:

$$\dot{\tilde{r}}(t) = \lambda m_a(t), \tag{13}$$

where the λ is any positive gain. In fact, from (12) and (13) it turns out that the closed loop dynamic of \tilde{r} satisfies:

$$\dot{\tilde{r}}(t) \begin{cases} > 0 & \text{if } \tilde{r}(t) < 0 \\ = 0 & \text{if } \tilde{r}(t) = 0 \\ < 0 & \text{if } \tilde{r}(t) > 0, \end{cases}$$

that obviously means that $\tilde{r}(t)$ converges to zero.

Note that the parameter m_a can be estimated from apparent contours as discussed in Section 4.

It is important to remark that the parameters m_a cannot be used in (13) in the particular case in which both actual and desired views are aligned without any relative orientation. In fact, it results in $\delta e_a = \delta e_d = 0$ (see (8)) and then $m_a \rightarrow \infty$. Note that in this case many standard techniques can be still used to detect this particular case, such as the epipolar bi-tangency concept [23]. Then a translational

motion based on the area of observed contours can accomplish steering the robot to the target pose.

Moreover, note that in the particular case in which the desired camera is aligned with the actual one ($m_a \rightarrow \infty$) then the so-called epipolar bi-tangents can be applied to design the camera motion strategy proposed in Ref. [30].

The second step brings the robot to the desired position following a circular trajectory. Two cases are in order.

5.1. Case I: known radius

Suppose that the radius r_{des} of the circular motion is *a priori* known. Then the circular trajectory of the camera can be parameterized as follows (see Fig. 9):

$$\begin{aligned} X_a(t) &= r_{des} \cos \varphi_a(t) \\ Y_a(t) &= r_{des} \sin \varphi_a(t) \\ \alpha(t) &= \varphi_a(t) + \pi, \end{aligned} \tag{14}$$

where $\varphi_a(t)$ is the current camera position angle at time t . The differential kinematics of the system is hence described by:

$$\begin{aligned} \dot{X}_a(t) &= -Y_a(t)\dot{\varphi}_a(t) \\ \dot{Y}_a(t) &= X_a(t)\dot{\varphi}_a(t) \\ \dot{\alpha}_a(t) &= \dot{\varphi}_a(t), \end{aligned} \tag{15}$$

where $\dot{\varphi}_a(t)$ is the control parameter steering the linear and angular velocity in (1).

To design the visual servoing algorithm, the control parameter $\dot{\varphi}_a(t)$ must be computable from the image measurement w . Observe that $\theta = \varphi_d - \varphi_a$, thus when

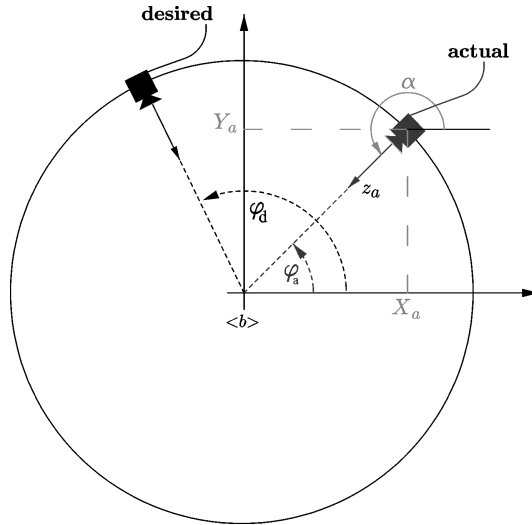


Figure 9. Parametrization of a circular motion centered on the origin of frame $\langle b \rangle$ and with radius r_{des} .

φ_a approaches the desired value φ_d , then epipoles go to infinity, and $1/w$ decreases and tends to zero for $\varphi_a = \varphi_d$ as shown in (4) (see also Fig. 9).

Then, a simple proportional control law of the visual measurement $1/w$ is given by:

$$\dot{\varphi}(t) = \frac{\lambda}{w(t)} \tag{16}$$

for some $\lambda > 0$.

5.2. Case II: unknown radius

Unlike the previous case, suppose here that the only *a priori* knowledge of the motion of the camera–robot is that a circular displacement occurs between the desired and the initial positions about an axis perpendicular to the motion plane and passing through an unknown point of the optical axis z_a . The trajectory radius r_{des} is unknown.

Let the initial configuration c_i and desired camera position c_d be as shown in Fig. 10. Starting from an initial guess \hat{r}_0 for the trajectory radius, apply controls ω , u_x and u_y (angular and linear velocities) as in (15) and (16). If $\hat{r}_0 \neq r$ (where $r = r_{des}$), the camera leaves the circular trajectory of radius r and reaches, after some amount of time, the new configuration c' as shown in Fig. 10. In this new configuration, the desired image and the current one (that taken by the camera in c') do not exhibit the property of symmetry discussed in Remarks 1 and 3. The epipoles are not symmetric with respect to the rotation axis and their positions are given by (7).

Let m_s be defined as:

$$m_s = \frac{1}{f_u \alpha_a} - \frac{1}{f_u \alpha_d} = \frac{1}{\delta e_a} - \frac{1}{\delta e_d}.$$

Parameter m_s accounts for the angle θ and, as in Section 5.1, will steer the camera along the circular trajectory with known radius. In fact, in some neighborhood of

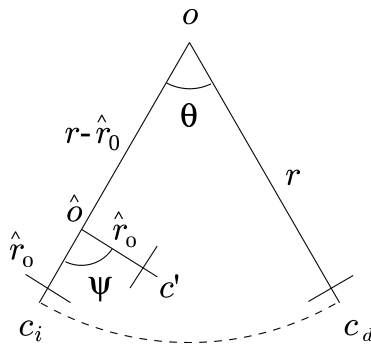


Figure 10. Robot motion under a (circular) control law with a wrong estimation of the circular radius.

the desired position the following properties hold:

$$\begin{aligned} m_s = 0 &\iff \varphi_a = \varphi_d \\ m_s(\varphi_a - \varphi_d) &\geq 0. \end{aligned} \quad (17)$$

In order to define the visual servoing procedure, the camera position c' , obtained by rotating the camera the angle ψ about an estimate of the rotation center \hat{o} as in Fig. 10, must be computed.

The general camera position and orientation c' along the (unknown radius) trajectory with respect to the initial position c_i can be written as:

$$\begin{aligned} X_a(t) &= (r - \hat{r}(t)) \cos \varphi_i + \hat{r}(t) \cos(\varphi_i + \psi(t)) \\ Y_a(t) &= (r - \hat{r}(t)) \sin \varphi_i + \hat{r}(t) \sin(\varphi_i + \psi(t)) \\ \alpha_a(t) &= \varphi_i + \psi(t) + \pi, \end{aligned}$$

where $\hat{r}(0) = \hat{r}_0$, $\psi(0) = 0$ and φ_i identifies the initial camera position c_i on the plane. Note that the camera orientation at c' is such that the optical axis intersects \hat{o} . The corresponding differential kinematics are obtained by differentiating:

$$\begin{aligned} \dot{X}_a(t) &= \dot{\hat{r}}(t) [\cos(\varphi_i + \psi(t)) - \cos \varphi_i] \\ &\quad - \hat{r}(t) \dot{\psi}(t) \sin(\varphi_i + \psi(t)) \\ \dot{Y}_a(t) &= \dot{\hat{r}}(t) [\sin(\varphi_i + \psi(t)) - \sin \varphi_i] \\ &\quad + \hat{r}(t) \dot{\psi}(t) \cos(\varphi_i + \psi(t)) \\ \dot{\alpha}_a(t) &= \dot{\psi}(t). \end{aligned} \quad (18)$$

Then, it turns out that control law:

$$\begin{aligned} \dot{\hat{r}}(t) &= \lambda_r m_a(t) \\ \dot{\psi}(t) &= -\lambda_a m_s(t), \end{aligned} \quad (19)$$

where λ_r and λ_a are positive gains, is locally asymptotically stable around the camera desired position. In fact, since (12) holds, $\hat{r}(t)$ converges to r analogously to $\tilde{r}(t)$ that converges to zero under (13). Then, there exists a neighborhood of the equilibrium point (the desired camera position) where (17) holds and, hence, where φ_a converges to φ_d since:

$$\dot{\varphi}_a(t) = \dot{\psi}(t) \begin{cases} > 0 & \text{if } \varphi_a(t) < \varphi_d \\ = 0 & \text{if } \varphi_a(t) = \varphi_d \\ < 0 & \text{if } \varphi_a(t) > \varphi_d. \end{cases}$$

Note that the parameters m_a and m_s can be estimated from apparent contours as discussed in Section 4.

Figure 11 shows the trajectories followed by the camera for four different initial estimates \hat{r}_0 of the unknown circular radius $r = 1$. Control parameters were set to $\lambda_r = 1$ and $\lambda_a = 0.1$.

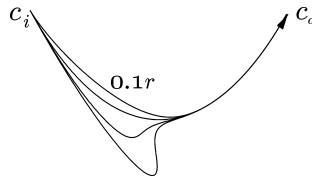


Figure 11. Trajectories followed for different initial estimates \hat{r}_0 : $0.1r$, r , $2r$ and $3r$ (decreasing order).

Remark 4. The roto-translational trajectory in Fig. 8 enjoys the property that the initial configuration can be chosen such that the objects in the scene are close to the image center. This allows us to easily keep the apparent contours of the objects in the field of view during the overall motion. Keeping the feature in the field of view is paramount for successfully executing the visual servoing procedure. Moreover, note that a different approach as performing a rotation first and a pure translation as the second and last action would fail since for pure translated cameras the epipoles go to infinity, thus vanishing the proposed control actions.

6. SIMULATIONS AND EXPERIMENTS

Simulation results are reported to validate the proposed visual servoing algorithm. General planar motion was tested. The initial and final robot–camera configurations are

$$\begin{pmatrix} Z_{c_i} \\ X_{c_i} \\ \alpha_{c_i} \end{pmatrix} = \begin{pmatrix} -0.866 \\ -0.5 \\ \pi/3 \end{pmatrix}; \quad \begin{pmatrix} Z_d \\ X_d \\ \alpha_d \end{pmatrix} = \begin{pmatrix} -0.866 \\ 0.5 \\ 2\pi/3 \end{pmatrix}.$$

The pure translation moves the robot along the optical axis from c_a to the intersection c_i with the circle passing through c_d and centered on o . From this point the robot–camera starts to rotate as described in Section 5.2. This second part of the trajectory, which steers the mobile robot to c_d , strongly depends on the initial guess of the unknown radius r . Simulations are reported in Fig. 12: the solid line (dashed line) corresponds to an initial guess which is 3 (0) times the true value. Control parameters are set to $\lambda = 1$, $\lambda_r = 1$ and $\lambda_a = 0.1$.

A second simulation was run to show system behavior for different control parameters.

Circular motions of unknown radius were considered. Figure 13 shows the trajectories followed for $\hat{r}_0 = 3r$, $\lambda_a = 1$ and different values of λ_r .

6.1. Experiments

Two experiments were run to validate the proposed visual servoing procedure. The experimental testbed consists of a mobile robot mounted with a fixed camera mounted (see Fig. 14). The robot is the Nomad XR4000, by Nomadic. The system

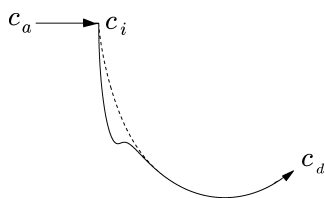


Figure 12. General planar motion: $\hat{r}_0 = 3r$ (solid line) and $\hat{r}_0 = 0$ (dashed line).

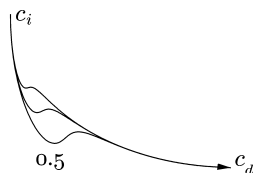


Figure 13. Trajectories for $\hat{r}_0 = 3r$ and different values of λ_r (0.5, 1 and 2).



Figure 14. The robot XR4000 with a fixed camera mounted on top.

is steered by four independently powered wheels which allow full control of the 3 d.o.f. of the camera–robot. The motion control of the XR4000 is left to three DSP's and a dedicated 32-bit micro-controller. The camera is by Hitachi. As far as internal parameters are concerned, recall that only the principal point, i.e. the image center, is needed: $(u_0, v_0) = (160, 120)$ pixels.

The scenario is that reported in Fig. 15. The objects in the scene, used to perform the visual servoing experiment, are a football and a sphere; both of them do not exhibit any special textures, only smooth surfaces whose main image feature consists of their apparent contours.

6.1.1. Circular motion. As a first experiment, the epipole-based visual servoing has been tested for the circular motion case with known radius. This preliminary experiment was useful to validate the estimation procedures and the control perfor-



Figure 15. The experimental setup consisting of the camera-robot and two objects with no special image features other than their apparent contours.

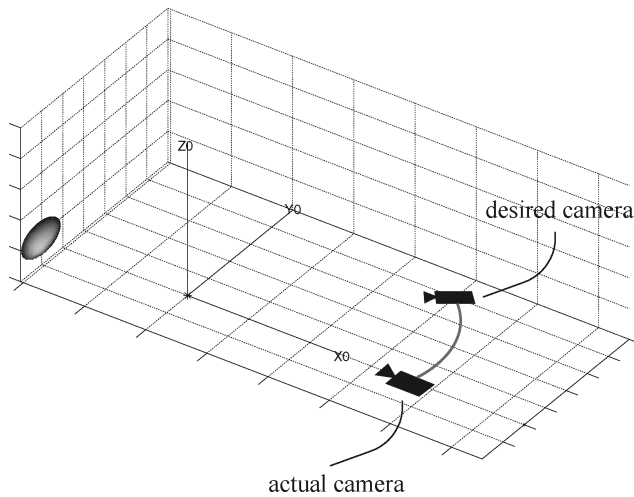


Figure 16. The experimental setup configuration for the circular motion. Only the football object is used.

mances in real cases. For the circular case only one object, i.e. the football, with its upper epipolar tangent has been used.

The experimental layout is that reported in Fig. 16. The initial position of the camera-robot is rotated 60° about the center of a circular trajectory whose radius, equal to 1 m, is assumed to be known. The object is 0.8 m from the center of rotation.

During the image pre-processing phase, the contour extraction and the closed contours detection are carried out in real-time. The visual servoing algorithm, described in Section 5.1, estimates the epipole position w in (4) from the current

images acquired with a frequency of about 2.5 frames/s, and moves the robot according to (15) and (16) with $\lambda = 20$ (pixel rad)/s.

After about 20 s, the visual servoing algorithm leads the control variable in (16), practically to 0. The experiment stops with an error of a few degrees. Figure 17 shows the superimposed final and desired apparent contours. Note that the steady-state error is caused by the chosen proportional control law. Better results would be obtained by means of a combined proportional–integral control action.

6.1.2. General motion. The second experiment deals with a general roto-translational motion. Both the objects, i.e. the football and the sphere, are used. The experimental layout is that reported in Fig. 18. The initial position of the camera–robot is rotated 60° about the center of the circular trajectory whose radius is equal to 1 m and translated 1 m along the optical axis.

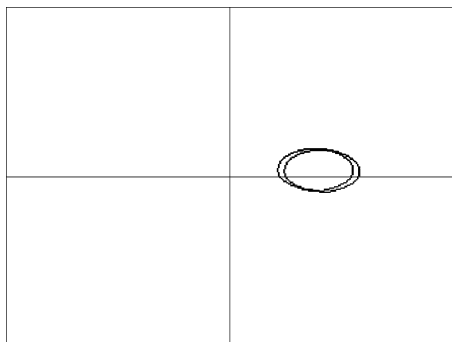


Figure 17. Superimposed final and desired image contour for the circular motion experiment.

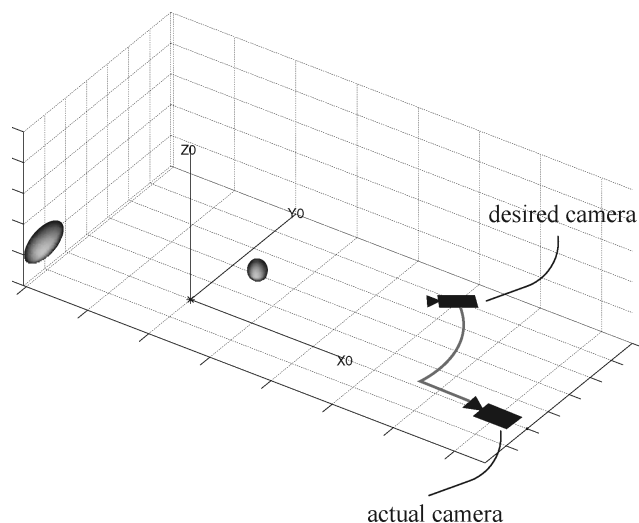


Figure 18. The experimental setup for the general roto-translational motion.

During the image pre-processing phase, the apparent contour extraction and the closed contours detection are completed by a simple procedure to solve the objects correspondence problem.

The optimization epipole estimation phase, discussed in Section 4, is computationally more demanding than the circular case since, here, the optimization problem (11) has 2 d.o.f., which are the two epipole displacements.

All the four couples of tangent lines to the objects have been used to perform the optimization procedure. Figure 19 shows the active contours of the two objects and the tangents obtained for displacements w and w' which minimize the cost function in (11) at a certain sampling time of the visual servoing experiment. The optimization problem is non-linear and non-convex, and the techniques used are classical numerical global optimization methods [32].

For the general motion case, it is very important to use B-splines for apparent contours description for two main reasons: the first one is that B-splines guarantee

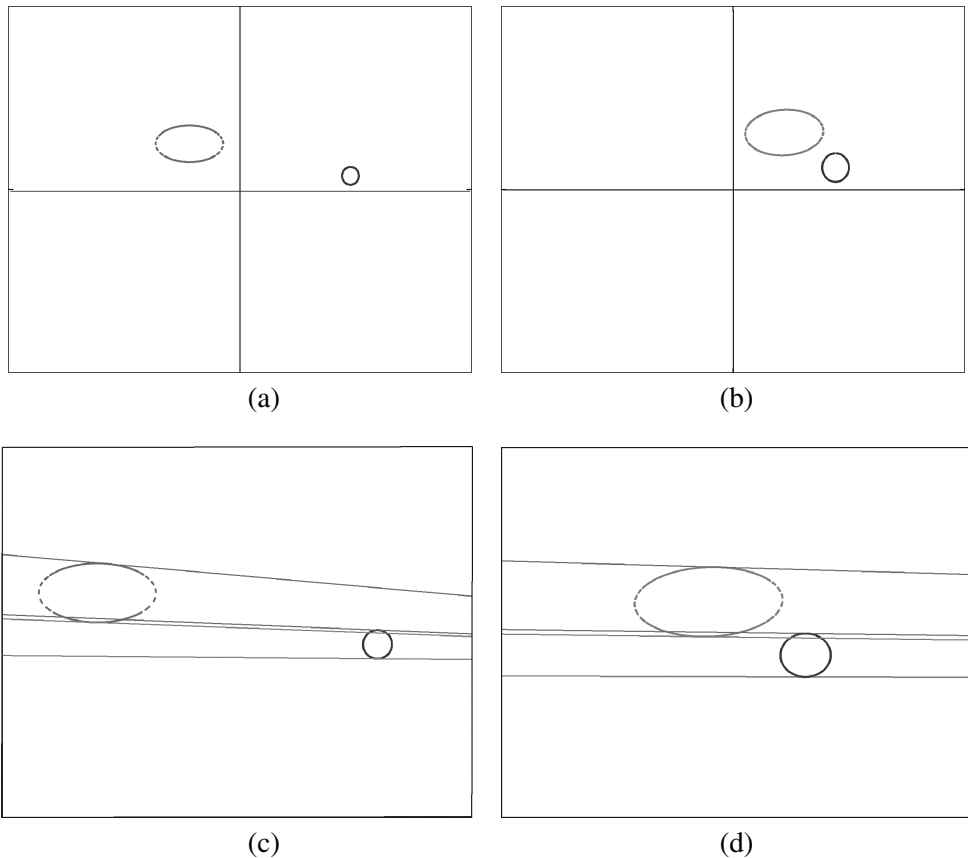


Figure 19. Actual (a) and the desired (b) image views at a certain sampling time. Details of the optimization procedure which ends by finding the four epipolar lines (two for every object) in the actual view (c) and desired view (d).

the smoothness of the contour and the second is that the contour parametrization can be exploited in the estimation procedure.

The image pre-processing and the overall computational burden brings the sampling time close to less than 1 s on a 650-MHz Pentium III.

The visual servoing algorithm is performed in two steps. During the first step, the symmetry condition for the epipole displacement is gained through a translation along the optical axis according to the control law (13) with $\lambda = 30$ (pixel m)/s. At the end of the pure translation, the mobile robot approaches the circular trajectory whose radius is unknown (see Fig. 18). Then, the second step of the visual servoing algorithm starts. The control law is that for an unknown radius discussed in Section 5.2. The control parameters in (19) are chosen as $\lambda_r = 50$ (pixel m)/s and $\lambda_a = 10$ (pixel rad)/s and the initial conditions of (19) are set to $\hat{r}_0 = 2$ m and $\psi_0 = 0$ rad.

Figure 20 shows positions of the camera–robot at different sampling times during the translation phase. Figure 21 shows the position of the robot–camera while tracking the circular trajectory. Note that the initial guess for the trajectory radius was 2 m, i.e. twice the real value.

Details on the optical axis orientation measured at seven sampling times are reported in Fig. 22. Figure 22a shows the optical axis orientation at the first four sampling times while the last three optical axis orientations are shown in Fig. 22b. Note how the optical axis orientation tends to the orientation of the target camera–robot.

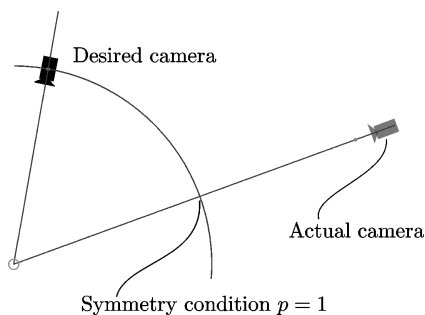


Figure 20. Position of the camera–robot (dots) during the translation.

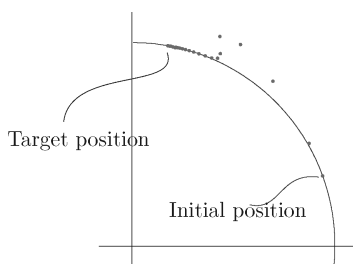


Figure 21. Positions (dots) of the robot–camera while tracking the circular trajectory.

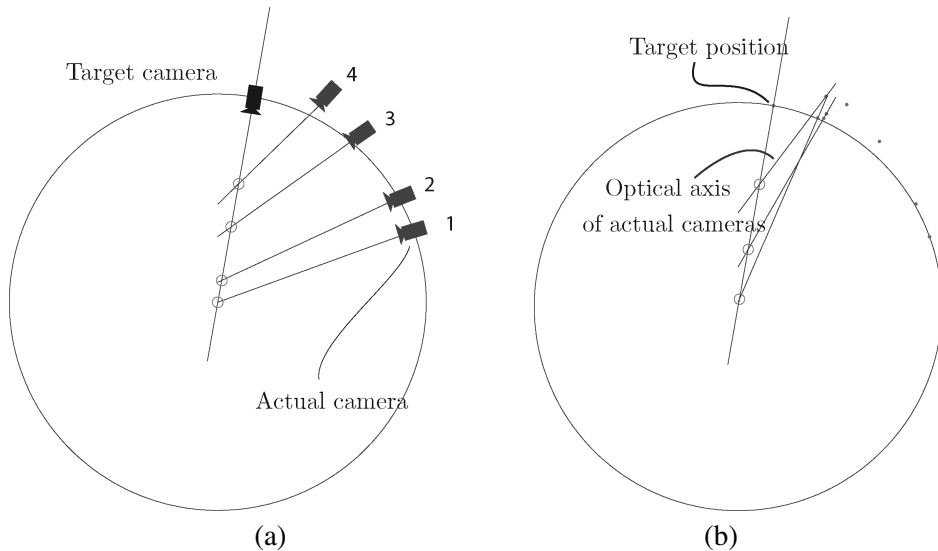


Figure 22. During the circular trajectory phase, the camera-robot's optical axis assumes different orientations. These seven samples (the first four in (a) and the last three in (b)) show how the actual orientation tends to the target orientation of the camera-robot.

7. CONCLUSIONS

Epipolar geometry was exploited to design an image-based visual servoing algorithm for a mobile robot moving on a plane with a fixed camera mounted on the robot. The visual servoing algorithm is based on a measure of the symmetry of the epipolar geometry which is retrieved using image contours and tangency constraints, but without solving any point correspondence problem. Exploiting profiles in visual feedback is crucial in outdoor navigation where objects in the scene are highly unstructured and solving for correspondences is difficult. As far as the camera calibration is concerned, only the principal point of the intrinsic parameters is assumed to be known.

For the sake of simplicity, holonomic planar mobile robots have been considered. To the best of our knowledge the confluence between epipolar geometry and mobile robotics visual servoing is far from being fully understood, and the simplicity of results achievable for the holonomic case appears to be important at this rather early stage of investigation.

Simulations and experiments have been executed to validate the epipole-based visual servoing.

Future work will investigate the extension of the proposed technique to the case of 6-d.o.f. motion. Research will also focus on non-holonomic mobile robots and dynamic scenes by means of the multibody fundamental matrix.

Acknowledgements

We wish to thank Jacopo Piazzzi for his help in designing and performing the experiments.

REFERENCES

1. D. Burschka and G. Hager, Vision-based control of mobile robots, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Seoul, pp. 1707–1713 (2001).
2. F. Conticelli, D. Prattichizzo, A. Bicchi and F. Guidi, Vision-based dynamic estimation and set-point stabilization of non-holonomic vehicles, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, pp. 2771–2776 (2000).
3. G. D. Hager, D. J. Kriegman, A. S. Georghiades and O. Ben-Shahar, Toward domain-independent navigation: dynamic vision and control, in: *Proc. IEEE Conf. on Decision and Control*, Tampa, FL, pp. 3257–3262 (1998).
4. K. Hashimoto and T. Noritsugu, Visual servoing of non-holonomic cart, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Albuquerque, NM, pp. 1719–1724 (1997).
5. K. Hashimoto, *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*. World Scientific, Singapore (1993).
6. S. Hutchinson, G. D. Hager and P. I. Corke, A tutorial on visual servo control, *IEEE Trans. Robotics Automat.* **12**, 651–670 (1996).
7. G. Chesi and A. Vicino, Visual servoing for large camera displacements, *IEEE Trans. Robotics* **20** 724–735 (2004).
8. P. I. Corke and S. A. Hutchinson, A new partitioned approach to image-based visual servo control, *IEEE Trans. Robotics Automat.* **17**, 507–515 (2001).
9. K. Hashimoto, A review on vision-based control of robot manipulators, *Advanced Robotics* **17**, 969–991 (2003).
10. E. Malis, Visual servoing invariant to changes in camera-intrinsic parameters, *IEEE Trans. Robotics Automat.* **20**, 72–81 (2004).
11. E. Malis, F. Chaumette and S. Boudet, 2 1/2 D visual servoing, *IEEE Trans. Robotics Automat.* **15**, 238–250 (1999).
12. Y. Mezouar and F. Chaumette, Path planning for robust image-based control, *IEEE Trans. Robotics Automat.* **18**, 534–549 (2002).
13. C. J. Taylor and J. P. Ostrowski, Robust vision-based pose control, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, pp. 2734–2740 (2000).
14. R. Cipolla and P. J. Luong, *Visual Motion of Curves and Surfaces*. Cambridge University Press, Cambridge (2000).
15. P. R. S. Mendonca, K.-Y. K. Wong and R. Cipolla, Circular motion recovery from image profiles, in: *Proc. Workshop on Vision Algorithms: Theory and Practice*, Corfu, pp. 119–126 (1999).
16. G. Chesi, D. Prattichizzo and A. Vicino, A visual servoing algorithm based on epipolar geometry, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Seoul, pp. 737–742 (2001).
17. J. Piazzzi, D. Prattichizzo and A. Vicino, Visual servoing along epipoles, in: *Control Problems in Robotics*, A. Bicchi, E. Christensen and D. Prattichizzo (Eds), pp. 215–232. Springer, Berlin (2003).
18. G. Chesi, E. Malis and R. Cipolla, Automatic segmentation and matching of planar contours for visual servoing, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, pp. 2753–2758 (2000).
19. E. Malis, G. Chesi and R. Cipolla, 2.5-D visual servoing with respect to planar contours having complex and unknown shapes, *Int. J. Robotics Res.* **22**, 841–853 (2003).

20. C. Collewet, A. Alhaj and F. Chaumette, Model-free visual servoing on complex images based on 3D reconstruction, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, New Orleans, LA, pp. 751–756 (2004).
21. F.-X. Espiau, E. Malis and P. Rives, Robust features tracking for robotic applications: Towards 2 1/2 D visual servoing with natural images, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Washington, DC, pp. 574–579 (2002).
22. R. Hartley and A. Zisserman, *Multiple View in Computer Vision*. Cambridge University Press, Cambridge (2000).
23. T. Sato and J. Sato, Visual servoing from uncalibrated cameras for uncalibrated robots, *Systems Comp. Japan* **31**, 11–19 (2000).
24. Y. Ma, J. Kosecka and S. S. Sastry, Vision guided navigation of non-holonomic mobile robot, *IEEE Trans. Robotics Automat.* **15**, 521–536 (1999).
25. R. Vidal, S. Soatto, Y. Ma and S. S. Sastry, Segmentation of dynamic scenes from the multibody fundamental matrix, in: *ECCV Workshop on Vision and Modelling of Dynamic Scenes* (2002).
26. Q.-T. Luong and O. D. Faugeras, The fundamental matrix: theory, algorithmism and stability analysis, *Int. J. Comp. Vision* **17**, 43–76 (1996).
27. O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, MA (1993).
28. H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, *Nature* **293**, 133–135 (1981).
29. P. R. S. Mendonça and R. Cipolla, Estimation of epipolar geometry from apparent contours: affine and circular motion cases, in: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Fort Collins, CO, pp. 1019–1024 (1999).
30. J. Sato and R. Cipolla, Affine reconstruction of curved surfaces from uncalibrated views of apparent contours, *IEEE Trans. Pattern Analysis Machine Intell.* **21**, 1186–1198 (1999).
31. G. L. Mariottini and D. Prattichizzo, Epipolar geometry estimation for contour-based visual servoing, in: *World Automation Congr.–Int. Symp. on Robotics and Applications*, Sevilla, pp. 529–534 (2004).
32. D. G. Luenberger, *Linear and Nonlinear Programming*. Addison-Wesley, Reading, MA (1989).

ABOUT THE AUTHORS



Graziano Chesi received the MS degree in Information Engineering from Firenze University (1997) and PhD degree in Systems Engineering from Bologna University (2001). He was a Visiting Scientist at the Vision and Robotics Group of Cambridge University (UK, 1999–2000), Ishikawa Hashimoto Laboratory of Tokyo University (2001–2004), and Research Associate (since 2000) and Assistant Professor (since 2004) in Control Systems at Siena University. His research interests include optimization in control systems and visual servoing.



Gian Luca Mariottini received the Laurea degree (cum laude) in Information Engineering from the University of Siena in 2002. From October 2004 to April 2005 he has been a Visiting Scientist at the GRASP Lab, University of Pennsylvania, USA. He is currently a PhD degree candidate at the University of Siena. His research interests include visual servoing for single and multiple robots, with applications to multiple view geometry, panoramic cameras and active vision for localization.



Domenico Prattichizzo received the MS degree in Engineering (1991) and PhD degree in Robotics and Automation (1995) from Pisa University. He was Associate Professor at Siena University (since 2002), Visiting Scientist at the MIT AI Lab (1994), Research Associate in the Robotics and Automation group at the Centro 'E. Piaggio' of Pisa University (since 1992) and Research Associate in the 'Center for Complex Systems Studies' of Siena University (since 2000). He was Co-chair of the IEEE International Workshop on Control Problems in Robotics and Automation (2002) and Co-organizer of the ICRA workshop on Multi-point Interaction in Robotics and Virtual Reality (2004). He is Co-editor of *Multi-point Interaction in Robotics and Virtual Reality* (in press), *Control Problems in Robotics*, vol. 4 (2003), both published by Springer. He is a member of the Editorial Board of *IEEE Transactions on Robotics*, *IEEE Transactions on Control Systems Technologies* (since 2003) and *Journal of Dynamics of Continuous, Discrete and Impulsive Systems (DCDIS) Series B: Application and Algorithms* (since 2001). He is also a member of the Editorial Board and a Member of the Conference Editorial Board of the IEEE Control System Society (since 2001). He was a member of the Program Committee of IEEE/RSJ International Conference on Intelligent Robots and Systems (2003) and the International Program Committee of IEEE International Conference on Robotics and Automation (2004). He has also been Co-organizer and Chairman of many sessions on robotics and control theory in many international conferences, and has been invited to give talks in many international conferences and research centers. He is the co-author of more than 130 papers in the area of robotics and automation.



Antonio Vicino received the Laurea in Electrical Engineering from the Politecnico di Torino, Torino, Italy, in 1978. From 1979 to 1982 he held several Fellowships at the Dipartimento di Automatica e Informatica of the Politecnico di Torino. He was assistant professor of Automatic Control from 1983 to 1987 at the same Department. From 1987 to 1990 he was Associate Professor of Control Systems at the Università di Firenze. In 1990 he joined the Dipartimento di Ingegneria Elettrica, Università di L'Aquila, as Professor of Control Systems. Since 1993 he has been is with the Università di Siena, where he founded the Dipartimento di Ingegneria dell'Informazione and covered the position of Head of the Department from 1996 to 1999. From 1999 he has been Dean of the Engineering Faculty. In 2000 he founded the Center for Complex Systems Studies (CSC) of the University of Siena, where he presently covers the position of Director. He served as Associate Editor for the *IEEE Transactions on Automatic Control* from 1992 to 1996. Presently he serves as Associate Editor for *Automatica* and is Associate Editor at Large for *IEEE Transactions on Automatic Control*. He is Fellow of the IEEE. He is the author of 200 technical publications, Co-editor of 2 books on *Robustness in Identification and Control*, Guest Editor of the Special Issue 'Robustness in Identification and Control' of the *International Journal of Robust and Nonlinear Control*. He has worked on stability analysis of non-linear systems and time series analysis and prediction. Presently, his main research interests include robust control of uncertain systems, robust identification and filtering, mobile robotics and applied system modeling.