Image-Based Visual Servoing for Nonholonomic Mobile Robots Using Epipolar Geometry

Gian Luca Mariottini, *Member, IEEE*, Giuseppe Oriolo, *Senior Member, IEEE*, and Domenico Prattichizzo, *Member, IEEE*

***Abstract*—We present an image-based visual servoing strategy for driving a nonholonomic mobile robot equipped with a pinhole camera toward a desired configuration. The proposed approach, which exploits the epipolar geometry defined by the current and desired camera views, does not need any knowledge of the 3-D scene geometry. The control scheme is divided into two steps. In the first, using an approximate input–output linearizing feedback, the epipoles are zeroed so as to align the robot with the goal. Feature points are then used in the second translational step to reach the desired configuration. Asymptotic convergence to the desired configuration is proven, both in the calibrated and partially calibrated case. Simulation and experimental results show the effectiveness of the proposed control scheme.**

***Index Terms*—Epipolar geometry, image-based visual servoing (IBVS), input–output feedback linearization, nonholonomic mobile robots.**

I. Introduction

THIS PAPER presents an image-based visual servoing (IBVS) method for driving a nonholonomic mobile robot to a desired configuration (*set-point*), which is specified through a desired image previously acquired by an on-board pinhole camera.

Differently from position-based visual servoing (PBVS), in IBVS, both the control objective and the control law are directly expressed in the image feature parameter space [15]. An image feature parameter is a real-valued quantity that can be calculated from one or more image features (points, lines, moments, etc.). As a consequence, IBVS schemes do not need any *a priori* knowledge of the 3-D structure of the observed scene. In addition, IBVS is more robust than PBVS with respect to uncertainties and disturbances affecting the model of the robot, as well as the calibration of the camera [10], [15]. However, robot convergence can typically be guaranteed only in a neighborhood of the desired configuration.

For articulated manipulators, error convergence to zero for the whole task space has been obtained via a hybrid approach between IBVS and PBVS, called 2-1/2-D visual servoing, and introduced in [22] and [23]. Besides image features, this strategy uses in the control law part of the camera displacement (rotation and a scaled translation) between the current and the target view, obtained via a decomposition of the homography matrix.

Recently, there has been an increasing interest in the visual servoing of mobile robots, which are typically subject to nonholonomic kinematic constraints. Among the first PBVS approaches in this area, we mention [26] and [30], where the authors propose to use a pan-tilt camera to add more degrees of freedom to the vision sensor. In [14], a piecewise-smooth PBVS for mobile robots is presented, based on the observation of environment features. All these methods, however, need metrical information about the feature position with respect to the camera-robot frame in order to guarantee convergence to the desired configuration.

The IBVS method for nonholonomic robots proposed in [8] does not need any *a priori* 3-D knowledge of the scene, and uses an adaptive control law to estimate the feature positions with respect to the camera-robot frame. A visual servoing strategy based on the estimation of the height of features on the plane of motion was proposed in [4] for mobile robots equipped with different types of vision sensors, such as pan-tilt heads or panoramic cameras. However, the last two visual servoing schemes suffer from the same potential drawback, i.e., the control law uses the inverse of the image Jacobian, and can, therefore, become singular for certain configurations of the mobile robot or feature points.

The 2-1/2-D approach has also been used to design a visual servoing scheme for unicycle mobile robots in [11]. In addition to the typical requirements of 2-1/2-D techniques, such a scheme requires perfect knowledge of the camera calibration parameters and the adaptive estimation of a constant depth-related parameter during servoing. Since the proposed controller is Lipschitz-continuous and time-varying, convergence to the desired configuration is not exponential.

The IBVS strategy for set-point stabilization of nonholonomic robots proposed in this paper is the outgrowth of [24], which was the first visual regulation scheme for nonholonomic mobile robots based on epipolar geometry. In the literature, this tool has been successfully applied to design visual-control laws for articulated manipulators [3], [29], [31] and omnidirectional mobile robots [7].

Given two views (in our case, desired and actual) of the same scene, the epipoles are the 2-D points where the line joining the two camera centers intersects the corresponding image planes [13]. Since the epipoles only encode the relative orientation between the actual and the desired views, our control algorithm

Manuscript received January 18, 2006; revised July 14, 2006. This paper was recommended for publication by Associate Editor P. Rocco and Editor L. Parker upon evaluation of the reviewers' comments. Color versions of Figs. 19 and 20 are available online at http://ieeexplore.org.

G. L. Mariottini and D. Prattichizzo are with the Dipartimento di Ingegneria dell'Informazione, Università di Siena, 53100 Siena, Italy (e-mail: gmariottini@dii.unisi.it; prattichizzo@dii.unisi.it).

G. Oriolo is with the Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza," 00184 Roma, Italy (e-mail: oriolo@dis.uniroma1.it).

Digital Object Identifier 10.1109/TRO.2006.886842

is based on the execution of two sequential steps. The first, which compensates the orientation error so as to align the robot with the desired configuration, is based on the use of approximate input–output linearization. The second step, which actually leads the system to the target, zeroes the translational displacement using feature points. The resulting visual servoing strategy guarantees asymptotic convergence, with exponential rate, of the nonholonomic mobile robot to the desired configuration.

The epipoles can be classified as image feature parameters, since they can be directly computed from many kinds of image features, such as points [18], apparent contours of curved surfaces [7], [33], or optical flows [37]. Thus, according to [15], our two-step visual servoing algorithm can be classified as an IBVS scheme.

The main advantages of our approach are as follows.

- Using the epipoles as outputs, the problems arising from local minima and singularities, which are typically encountered in IBVS schemes based on the image Jacobian, are avoided.
- No metrical knowledge of the 3-D scene is needed, because our IBVS strategy is based on the epipolar geometry, which is independent from the scene geometry. In particular, differently from 2-1/2-D techniques, the proposed algorithm does not estimate the relative camera displacement via homography matrix decomposition, a process which requires manual intervention (in fact, such decomposition is not unique, and to avoid this ambiguity, it is necessary to discriminate between two possible normal vectors to the target plane) and could also increase the sensitivity to image noise [6].
- Since the proposed visual servoing strategy uses the epipolar geometry, our strategy can be applied in the case of a 3-D scene exhibiting not only image feature points, but also other kinds of features, such as smooth surfaces (e.g., nonplanar object contours) or optical flows.
- Asymptotic convergence to the desired configuration can be proven also in the case of a partially calibrated pinhole camera, and, in particular, in the case of unknown focal length.

From a control viewpoint, it should be emphasized that the set-point regulation of nonholonomic mobile robots considered in this paper is more difficult than the trajectory-tracking problem addressed, e.g., in [19], where the robot tracks an arbitrarily shaped ground curve observed on the camera image plane. This is due to the peculiar nature of nonholonomic systems, for which posture stabilization cannot be achieved via smooth time-invariant feedback [5].

The paper is organized as follows. Section II introduces the basic notations and definitions of epipolar geometry for pinhole cameras. In Section III, the nonholonomic visual servoing problem is formulated, and the two-step control strategy is outlined. The first step is analyzed in Section IV, where the epipole kinematics is derived and used to design a control law based on feedback linearization. Section V describes the feature-based control law which implements the second step. The proposed method is extended to the case of a partially calibrated camera in Section VI. Simulation and experimental results are presented in Sections VII and VIII, respectively, to show the effectiveness and practical applicability of the proposed approach. In Section IX, we discuss some possible extensions to the basic problem. In Section X, we provide some concluding remarks highlighting the main contributions of the paper.

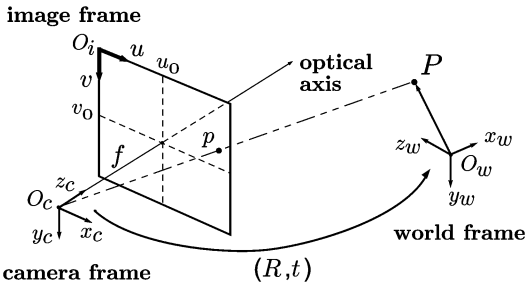

Fig. 1. Formation of a perspective projection of the 3-D point P in the pinhole camera model. The optical center of the camera is in O_c.

II. Basic Epipolar Geometry for Pinhole Camera

In this section, we quickly review the fundamentals of the perspective camera model and multiple-view geometry. The reader is referred to [13] and [20] for a detailed treatment.

With reference to Fig. 1, consider a pinhole camera with optical center in O_c, and the associated camera frame $\{O_c, x_c, y_c, z_c\}$. The full perspective model describes the relationship in homogeneous coordinates between a 3-D point P (expressed in the world frame $\{O_w, x_w, y_w, z_w\}$) and its projection p on the image plane (expressed in the image frame $\{O_i, u, v\}$) as

$$p = K[R \mid t]P \tag{1}$$

where R, t are the *extrinsic* camera parameters (i.e., respectively, the rotation and the translation between the world and the camera frame, defined in the camera frame), and K is the *intrinsic* camera parameter matrix

$$K = \begin{bmatrix} fk_u & fk_u \cot\phi & u_0 \\ 0 & \frac{fk_v}{\sin\phi} & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Here, (u_0, v_0) represent the image frame coordinates, in pixels, of the *principal point* (i.e., the intersection point between the image plane and the optical axis z_c), k_u and k_v are the number of pixels per unit distance in image coordinates, f is the focal length, and ϕ is the angle between the image axes u and v. In the following, we suppose that the camera intrinsic parameters have been determined through a preliminary calibration phase (although we shall partially relax this assumption later in the paper), and that $\phi = \pi/2$. Without loss of generality, it is then possible to assume $u_0 = v_0 = 0$ (i.e., the origin of the image frame is at the principal point) and $k_u = k_v = 1$. Hence, matrix K is known, and takes the form

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

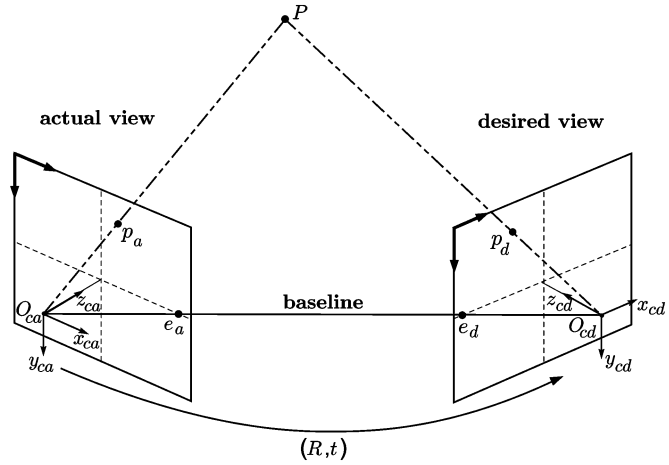

Fig. 2. Basic epipolar geometry setup defined by two different views (distinct camera placements) of the same 3-D point, i.e., the actual view (on the left) and the desired view (on the right). For later use, the figure refers to the particular case in which a planar motion occurs between the two cameras, i.e., the planes $x_{ca} - z_{ca}$ and $x_{cd} - z_{cd}$ coincide.

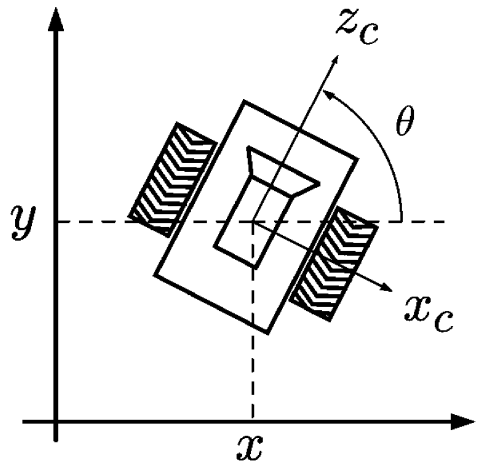

Fig. 3. Mobile robot with unicycle kinematics carrying a camera.

Consider now two distinct camera frames $\{O_{ca}, x_{ca}, y_{ca}, z_{ca}\}$ and $\{O_{cd}, x_{cd}, y_{cd}, z_{cd}\}$, corresponding, respectively, to the *actual* and *desired* views of the same 3-D point P (see Fig. 2). Without loss of generality, we assume that the world frame coincides with the desired view frame (i.e., $\{O_{cd}, x_{cd}, y_{cd}, z_{cd}\} \equiv \{O_w, x_w, y_w, z_w\}$).

The two intersections $e_a = [e_{au}\ e_{av}\ 1]^T$, $e_d = [e_{du}\ e_{dv}\ 1]^T$ between the segment $O_{ca}O_{cd}$ (the *baseline*) and the image planes are called *epipoles*. Note that each epipole is expressed (in homogeneous coordinates) in the coordinate frame of the corresponding image plane (see Fig. 2).

It is easy to understand that the epipoles are invariant with respect to a translation of the image planes along the baseline. This indicates that only the relative orientation between the planes can be retrieved from the knowledge of e_a, e_d (and of the camera parameters). This point, which will be further illustrated later, has precise consequences on the design of our servoing strategy.

The value of the epipoles can be directly computed by comparing the desired and actual views. A common approach [18] is to use the *fundamental matrix* F, an important tool in epipolar geometry. F is defined as

$$F = K^{-T}[t]_{\times}RK^{-1}$$

where $[t]_{\times}$ is the 3×3 skew-symmetric matrix associated with the translation t. Given at least eight generic correspondences (i.e., images of the same points in the two views), F can be computed up to a scale factor, without any knowledge of the 3-D structure of the observed scene. In the presence of image noise, F can be robustly estimated by means of well-known algorithms; see, for example, [13] and [18]. The epipoles e_a and e_d are then readily computed as vectors lying in the left and right null-spaces, respectively, of F

$$Fe_a = 0 \quad F^T e_d = 0. \tag{3}$$

Other techniques are available to estimate the epipoles, e.g., from apparent contours of generically shaped objects [7], [33], from optical flow [37], or using virtual parallax methods [21]. Also, algorithms based on the so-called *homography matrix* can be used to compute the epipoles when a planar scene is observed [13], [20].

III. The Nonholonomic Visual Servoing Problem

The objective of this paper is to present a visual servoing strategy that uses the epipole positions (reconstructed through an on-board camera) in order to drive a nonholonomic robot toward a desired configuration.

A. Problem Formulation

The nonholonomic mobile robot considered in this paper is a unicycle moving on a plane (see Fig. 3). Its configuration vector is defined as $[x\ \ y\ \ \theta]^T$, where x, y are the Cartesian coordinates of the center of the robot in a reference planar frame $\{O, x, y\}$, and θ is its orientation with respect to the x axis (see Fig. 3). The nonholonomic kinematic model is

$$\begin{aligned} \dot{x} &= u_1 \cos\theta \\ \dot{y} &= u_1 \sin\theta \\ \dot{\theta} &= u_2 \end{aligned} \tag{4}$$

where u_1 and u_2 are, respectively, the translational and angular velocity, assumed to be the available control inputs. Without loss of generality, we suppose that the desired configuration is $q_d = [0\ \ 0\ \ \pi/2]^T$, which corresponds to the robot being centered at the origin and aligned with the positive y axis.

The pinhole camera is fixed to the robot body in such a way that the optical center O_c is in $[x\ \ y]^T$ (i.e., at the center of the robot), the optical axis z_c is aligned with the robot forward axis, and the x_c axis is parallel to the motion plane. This means that the desired camera frame $\{O_{cd}, x_{cd}, z_{cd}\}$ coincides with the planar reference frame $\{O, x, y\}$.

Since we are pursuing a visual servoing approach, it is assumed that the desired camera view (i.e., the view acquired with the robot at q_d) has been gathered in advance, and that a

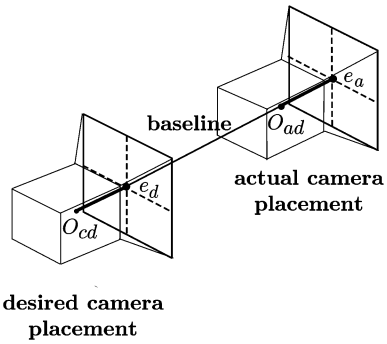

Fig. 4. When the actual and desired configurations are aligned, both epipoles coincide with the origin of the corresponding image frame.

sufficient number n ($n \geq 8$) of corresponding feature points $\{(p_{ai}, p_{di}),\ i = 1, \ldots, n\}$ is observed in the actual and desired views (as already mentioned, other methods can be used to reconstruct the epipoles). Therefore, the epipoles e_a and e_d can be computed at each sampling instant from the fundamental matrix, as in (3).

On the other hand, the configuration q of the mobile robot is assumed to be completely unknown.

B. Sketch of the Control Method

First we note that, since only planar movements of the camera are allowed, only the u-coordinates of the epipoles (i.e., e_{au} and e_{du}) change during the motion (see Fig. 2). With a slight abuse of terminology, we will, henceforth, refer to e_{au} as the *actual* epipole, and e_{du} as the *desired* epipole.

The design of the proposed visual servoing method is inspired by the particular situation shown in Fig. 4, in which a pure translation separates the actual and desired configurations of the robot. In this case, both the epipoles coincide with the origin (i.e., the principal point) of the corresponding image frame. Hence, their coordinates (in particular, e_{au} and e_{du}) are zero. Moreover, the baseline, the actual optical axis, and the desired optical axis coincide.

The proposed visual-control algorithm drives the nonholonomic mobile robot to the desired configuration q_d in two steps (see Fig. 5).

1) From the initial robot configuration q_0, apply a control law that brings the epipole coordinates e_{au} and e_{du} to zero. As shown in Section IV, such a control may be computed through input–output feedback linearization. At the end of this step, the situation is exactly as in Fig. 4: the intermediate configuration q_i is aligned with the desired configuration (see Fig. 5, left).
2) From the intermediate robot configuration q_i, apply a control law producing a translation of the robot from q_i to q_d (see Fig. 5, right). Clearly, the epipoles cannot be used in this phase, as they are identically zero (we already noticed in Section II that a relative translation of the image planes along the baseline cannot be derived from the epipoles). However, this step is easily realized on the basis of corresponding points in the images.

Roughly speaking, the first step is aimed at zeroing the orientation error (and placing the robot along the y axis),

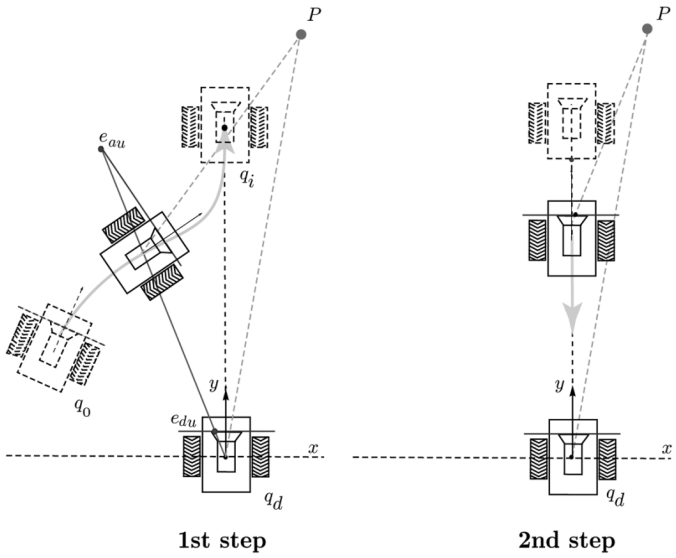

Fig. 5. The two steps of the proposed visual servoing strategy. Left: The nonholonomic robot is first driven to the y axis by zeroing the epipole coordinates e_{au} and e_{du}. Right: A feature-based controller is then used to recover the translation error.

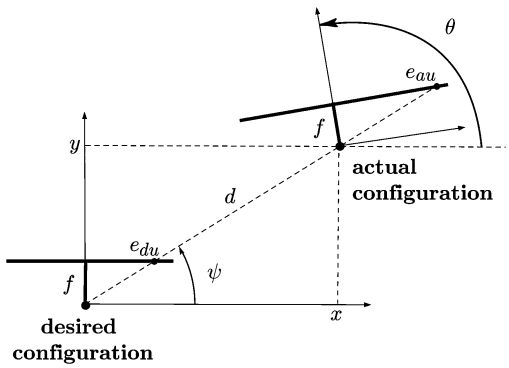

Fig. 6. Planar geometric setup used to compute the epipole kinematics.

while the second step compensates the translation error. In Sections IV–IX, we will analyze in detail the structure of the two steps.

IV. First Step: Zeroing the Epipoles

In this section, we will design a visual feedback for the mobile robot in order to control the motion of the epipoles and realize the first step of our visual-control strategy. To this end, after deriving the epipole kinematics, we adopt an approximate input–output linearization approach.

A. Epipole Kinematics

A detailed derivation of the epipole kinematics is now presented. First, we will derive the expression of the epipoles as a function of the unicycle configuration q. Then, their differential kinematics will be written with respect to the unicycle velocities, i.e., the available control inputs. This last step is instrumental in the feedback-linearization step of our control method, in which the epipole positions will be taken as system outputs.

Let us consider the situation in Fig. 6, which shows the planar setup associated with two views of the same scene: the

actual view, taken with the robot in a generic configuration $q = [x\ y\ \theta]^T$, and the desired view, corresponding to the robot in the desired configuration $q_d = [0\ 0\ \pi/2]^T$. Denote by ψ the angle between the x axis and the line joining the desired and actual camera centers (i.e., the baseline), and by d the distance between these two points.

First of all, recall that the y-coordinates of both epipoles are identically zero, because the motion of the camera is planar. Since $f = e_{du}\tan\psi$ and $y = x\tan\psi$, we readily obtain

$$e_{du} = f\frac{x}{y}. \tag{5}$$

Similarly, as

$$e_{au} = f\tan(\theta-\psi) = f\frac{\tan\theta-\tan\psi}{1+\tan\theta\tan\psi}$$

we get

$$e_{au} = f\frac{x\sin\theta - y\cos\theta}{x\cos\theta + y\sin\theta}. \tag{6}$$

One may easily verify that these expressions hold for any position of the actual camera, i.e., for any sign of e_{au} and e_{du}. As expected, e_{du} and e_{au} are clearly invariant if the robot translates along the baseline. Also, if $x = y = 0$ (i.e., if there is a pure rotation between the desired and actual views), the epipoles are not defined.

For feedback-linearization purposes, we will also need the differential kinematics of the epipoles with respect to the unicycle velocities u_1 and u_2. From (4) and (5), the time derivative of e_{du} is obtained as

$$\dot{e}_{du} = f\frac{u_1 y\cos\theta - u_1 x\sin\theta}{y^2}.$$

Since from Fig. 6, it results that $y = d\sin\psi$ and $x = d\cos\psi$, we may also write

$$\dot{e}_{du} = u_1\frac{f}{d}\frac{\sin(\psi-\theta)}{\sin^2\psi}. \tag{7}$$

The time derivative of e_{au} may be conveniently computed starting from $e_{au} = f\tan(\theta-\psi)$, which yields

$$\dot{e}_{au} = \frac{f}{\cos^2(\theta-\psi)}(\dot{\theta}-\dot{\psi}).$$

From $\psi = \arctan(y/x)$ we obtain $\dot{\psi} = (\dot{y}x - \dot{x}y)/d^2$; using this in conjunction with (4) yields

$$\dot{e}_{au} = u_1\frac{f\sin(\psi-\theta)}{d\cos^2(\psi-\theta)} + u_2\frac{f}{\cos^2(\psi-\theta)}. \tag{8}$$

Equations (7) and (8) may be also written as a function of the epipoles. From Fig. 6, simple geometry gives

$$\cos(\psi-\theta) = \operatorname{sign}(e_{au}e_{du})\frac{f}{\sqrt{e_{au}^2+f^2}}$$
$$\sin(\psi-\theta) = -\operatorname{sign}(e_{au}e_{du})\frac{e_{au}}{\sqrt{e_{au}^2+f^2}}$$

where the sign functions guarantee that these expressions hold for any position of the actual camera. Substituting these in (7) and (8), we finally obtain

$$\dot{e}_{au} = -u_1\frac{\operatorname{sign}(e_{au}e_{du})}{d}\frac{e_{au}\sqrt{e_{au}^2+f^2}}{f} + u_2\frac{e_{au}^2+f^2}{f} \tag{9}$$

$$\dot{e}_{du} = -u_1\frac{\operatorname{sign}(e_{au}e_{du})}{d}\frac{e_{au}\left(f^2+e_{du}^2\right)}{f\sqrt{e_{au}^2+f^2}}. \tag{10}$$

Equations (9) and (10) express the rate of variation of the actual and desired epipole coordinates as a function of the robot input velocities.

B. Control Based on Approximate Input–Output Linearization

Since the epipole coordinates e_{au} and e_{du} are reconstructed in real time via the on-board camera, they can be considered as outputs of the complete camera-robot system

$$\begin{aligned}\dot{x} &= u_1\cos\theta\\ \dot{y} &= u_1\sin\theta\\ \dot{\theta} &= u_2\\ e_{au} &= f\frac{x\sin\theta - y\cos\theta}{x\cos\theta + y\sin\theta}\\ e_{du} &= f\frac{x}{y}.\end{aligned}$$

The standard procedure to compute an input–output linearizing control law is to differentiate the output functions and invert, if possible, the resulting map (see [16]). From the epipole differential kinematics (9) and (10), the relationship between the control inputs and the output time derivatives is expressed as

$$\begin{bmatrix}\dot{e}_{au}\\ \dot{e}_{du}\end{bmatrix} = E\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

with

$$E = \begin{bmatrix} -\frac{\operatorname{sign}(e_{au}e_{du})}{d}\frac{e_{au}\sqrt{e_{au}^2+f^2}}{f} & \frac{e_{au}^2+f^2}{f}\\ -\frac{\operatorname{sign}(e_{au}e_{du})}{d}\frac{e_{au}\left(e_{du}^2+f^2\right)}{f\sqrt{e_{au}^2+f^2}} & 0\end{bmatrix}.$$

Here, we are faced with a major difficulty, i.e., the inverse of the *decoupling matrix* E cannot be computed, because the parameter $d = (x^2+y^2)^{1/2}$ (the distance between the actual and desired robot positions) is unknown in an image-based control setting. In a control terminology, input–output linearization cannot be achieved, in general, using *output* feedback (as opposed to

state feedback). In our case, however, it is possible to perform an *approximate* input–output linearization by setting

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \hat{E}^{-1} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \tag{11}$$

with

$$\hat{E}^{-1} = \begin{bmatrix} 0 & -\text{sign}(e_{au}e_{du})\frac{\hat{d}f\sqrt{e_{au}^2+f^2}}{e_{au}(e_{du}^2+f^2)} \\ \frac{f}{e_{au}^2+f^2} & -\frac{f}{e_{du}^2+f^2} \end{bmatrix}$$

where an estimate $\hat{d}$ of d has been used. The resulting epipole velocities are

$$\begin{bmatrix} \dot{e}_{au} \\ \dot{e}_{du} \end{bmatrix} = E\hat{E}^{-1} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} 1 & \left(\frac{\hat{d}}{d}-1\right)\frac{e_{au}^2+f^2}{e_{du}^2+f^2} \\ 0 & \frac{\hat{d}}{d} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix}.$$

Note that while the second equation is still linear (although time-varying) and decoupled, the first one is not. However, convergence of e_{au} and e_{du} to zero can be achieved by properly choosing the distance estimate $\hat{d}$ and the auxiliary inputs ν_1, ν_2, as shown in the following result.

Proposition 1: Let

$$\begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} -k_1 e_{au} \\ -k_2 e_{du}^{\beta/\gamma} \end{bmatrix} \tag{12}$$

where $k_1 > 0, k_2 > 0$, and β, γ are positive odd integers, with $\beta < \gamma$. Also, update the distance estimate $\hat{d}$ according to the following equation:

$$\dot{\hat{d}} = k_2 f^2 \hat{d} \frac{e_{du}^{\beta/\gamma}}{e_{au}(e_{du}^2+f^2)} \tag{13}$$

initialized at $\hat{d}_0 \geq d_0$, where d_0 is the initial value of d. Then, for sufficiently small k_2, the approximately linearizing control (11) drives the epipole coordinates e_{au} and e_{du} to zero for any initial condition such that $e_{au}(0) \neq 0$, with exponential convergence rate.

Proof: First of all, note that the closed-loop equations under the proposed control law are

$$\dot{e}_{au} = -k_1 e_{au} - \left(\frac{\hat{d}}{d}-1\right)\frac{e_{au}^2+f^2}{e_{du}^2+f^2} k_2 e_{du}^{\beta/\gamma} \tag{14}$$

$$\dot{e}_{du} = -\frac{\hat{d}}{d} k_2 e_{du}^{\beta/\gamma} \tag{15}$$

while the explicit form of the control inputs in (11) is

$$u_1 = \text{sign}(e_{au}e_{du})\frac{\hat{d}f\sqrt{e_{au}^2+f^2}}{e_{au}(e_{du}^2+f^2)} k_2 e_{du}^{\beta/\gamma} \tag{16}$$

$$u_2 = -\frac{f}{e_{au}^2+f^2} k_1 e_{au} + \frac{f}{e_{du}^2+f^2} k_2 e_{du}^{\beta/\gamma}. \tag{17}$$

The distance $d = (x^2+y^2)^{1/2}$ evolves according to

$$\dot{d} = \frac{x\dot{x}+y\dot{y}}{d} = \frac{xu_1\cos\theta + yu_1\sin\theta}{d}$$

and as $x = d\cos\psi$, $y = d\sin\psi$, we get

$$\dot{d} = u_1\cos(\psi-\theta) = k_2 f^2 \hat{d}\frac{e_{du}^{\beta/\gamma}}{e_{au}(e_{du}^2+f^2)}. \tag{18}$$

Comparing this with (13), it is clear that d and $\hat{d}$ obey the same differential equation. Hence

$$\frac{\hat{d}}{d} = \frac{\hat{d}_0 + \int_0^t \dot{\hat{d}}\,dt}{d_0 + \int_0^t \dot{\hat{d}}\,dt} = 1 + \frac{\hat{d}_0 - d_0}{d} \geq 1 \tag{19}$$

under the assumption $\hat{d}_0 \geq d_0$. As a consequence, the coefficient of $e_{du}^{\beta/\gamma}$ in (15) is negative, and bounded below in modulus. This means that zero is a terminal attractor [38] for e_{du}, which will converge to zero at a finite time instant $\bar{t}$. From $\bar{t}$ on, the differential (14) governing e_{au} reduces to $\dot{e}_{au} = -k_1 e_{au}$, so that e_{au} will converge to zero with exponential rate k_1. Note that according to (16), u_1 is identically zero after $\bar{t}$, so that the robot performs a pure rotation in this phase. Hence, convergence of the epipoles to zero is obtained at a finite distance $d(\bar{t})$.

It remains to be shown that the proposed control law does not become singular. In fact, while the angular velocity given by (17) is always defined, the linear velocity in (16) has a potential singularity when $e_{au} = 0$. As already noticed, after the transient ($t \geq \bar{t}$), we have $e_{du} = 0$, and therefore, u_1 becomes identically zero, thus preventing the singularity occurrence. Before $\bar{t}$, the dynamics of e_{au} as given by (14) includes a "perturbation" term whose effect can be arbitrarily bounded by bounding k_2. Hence, for sufficiently small k_2, the actual epipole e_{au} cannot cross zero during the transient. ■

Note the following points.

- The above control law (14) and (15) is image-based, because it only relies on the measured epipoles. No knowledge of the robot configuration or any other odometric data is assumed.
- The particular form of the exponent of e_{du} in the control law (12) is essential in guaranteeing that the desired epipole converges to zero *in finite time*, and therefore, that the proposed control law is never singular. This kind of control law is also known as a *terminal sliding mode*.
- The *zero dynamics* (i.e., the residual dynamics when the outputs are identically zero [16]) associated with our approximate input–output linearization controller is obtained from (18) as $\dot{d} = 0$. That is, the robot will converge to some point of the y axis at a *finite* distance d from its desired position, consistently with the above proof.
- If e_{au} is zero at $t = 0$ (i.e., the robot is initially aligned with the baseline), before applying the proposed controller, it is necessary to perform a preliminary maneuver in order to displace the actual epipole to a nonzero value. This can be

made in closed-loop in a number of ways. For example, one may enforce a fixed displacement of e_{au} from zero to a certain value $\bar{e}_{au}$, and achieve it by setting $u_1 = 0$ (a pure rotation) and

$$u_2 = -\frac{f}{e_{au}^2 + f^2} k_1 (e_{au} - \bar{e}_{au}), \qquad k_1 > 0.$$

With this choice, we get $\dot{e}_{au} = -k_1(e_{au} - \bar{e}_{au})$ from (10), so that e_{au} converges exponentially to $\bar{e}_{au}$. The maneuver can be stopped at a finite time obtaining a nonzero final value of e_{au} which is arbitrarily close to $\bar{e}_{au}$, and the two-step algorithm can then be used. Note that e_{du} does not change during this maneuver.

- *Proposition 1* applies under the assumption that the epipolar geometry between the desired and actual camera views is defined. This excludes the case in which a pure rotation occurs between the initial and desired robot configurations. However, it is possible to detect such a situation (e.g., by observing the norm of the 9-dimensional vector obtained by stacking the fundamental matrix F [21]), estimate the homography matrix (still well defined in this case), and decompose it to get the rotation matrix R between the two views. Once R has been computed, a simple proportional rotational controller can be used to zero the orientation error.
- With the inclusion of the two above rotational maneuvers (i.e., displacement of the initial actual epipole from zero and recovery of a pure orientation error), the proposed control scheme achieves *global* convergence of the robot to the desired configuration.

A final remark is in order concerning the duration of the first step. Since convergence of the actual epipole e_{au} to zero is exponential, an arbitrarily small error must be accepted in principle before switching to the second step of our control scheme. While in practice this is not a problem, as shown by the simulation and experimental results to be presented later, it is also possible to modify the proposed control law so as to guarantee finite-time convergence of *both* epipoles. To this end, one possibility is to replace (12) with

$$\begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} -k_1 e_{au}^{\beta/\gamma} \\ -k_2 e_{du}^{\beta/\gamma} \end{bmatrix}. \tag{20}$$

It is easy to verify that if k_1 is chosen so as to satisfy

$$k_1 < k_2 \left| \frac{e_{au,0}}{e_{du,0}} \right|^{1-\beta/\gamma}$$

the desired epipole e_{du} reaches zero before e_{au} under the assumptions of *Proposition 1*, and therefore, the singularity is still avoided.

C. Control Parameter Selection

We now present some guidelines for the selection of the control parameters, i.e., $\beta/\gamma, k_1, k_2$, and $\hat{d}_0$.

1) Choice of β/γ: Since the control law (14) and (15) is only aimed at zeroing the epipoles, the Cartesian distance d between the actual and desired robot positions may increase during the first step. An upper bound on the possible increment can be derived as follows. From (13), we have

$$\left| \frac{\dot{\hat{d}}}{\hat{d}} \right| = k_2 f^2 \frac{|e_{du}|^{\beta/\gamma}}{|e_{au}|(e_{du}^2 + f^2)} \leq k_2 \frac{|e_{du,0}|^{\beta/\gamma}}{|e_{au,\min}|}$$

where we have considered that $|e_{du}|$ is monotonically decreasing to zero, and we have denoted by $|e_{au,\min}|$ the minimum absolute value of the actual epipole during the transient phase of the first step (i.e., before $\bar{t}$), which is certainly nonzero under the hypothesis of *Proposition 1*. Hence, the maximum increment Δ of $\hat{d}$ during the first step depends on its maximum duration $\bar{t}_{\max}$, and is derived as

$$\Delta = \exp\left(k_2 \frac{|e_{du,0}|^{\beta/\gamma}}{|e_{au,\min}|} \bar{t}_{\max} \right).$$

Since $\hat{d}/d \geq 1$ from (19), $\bar{t}_{\max}$ can be computed from (15) setting $\hat{d} = d$, which corresponds to the slowest possible dynamics. An easy computation gives

$$\bar{t}_{\max} = \frac{|e_{du,0}|^{(1-\beta/\gamma)}}{k_2(1 - \beta/\gamma)}.$$

As $\hat{d} - \hat{d}_0 = d - d_0$ in view of the fact that $\dot{\hat{d}} = \dot{d}$, we can conclude that the maximum distance increment for d is

$$\Delta = \exp\left(\frac{1}{1 - \beta/\gamma} \frac{|e_{du,0}|}{|e_{au,\min}|} \right).$$

This shows that the fractional exponent β/γ should be chosen close to zero in order to reduce Δ. According to (15), this will also result in a faster convergence of e_{du}, i.e., in a quicker approach of the robot to the y axis.

2) Choice of k_1, k_2: In principle, *Proposition 1* requires that a sufficiently small k_2 is used in order to guarantee that the control law does not become singular (i.e., that e_{au} does not cross zero) during the transient. To select k_2 in practice, one may take into account the following arguments, which are based on the observation that with the proposed control law, the epipoles e_{au}, e_{du} never change sign.

- If the initial epipole values $e_{au,0}, e_{du,0}$ are discordant in sign, the perturbation term in (14) pulls e_{au} away from the singularity. Hence, any value of k_2 is legal, while k_1 only determines the speed of convergence of e_{au} to zero after $\bar{t}$.
- If the initial epipole values $e_{au,0}, e_{du,0}$ are concordant in sign, the perturbation term in (14) pushes e_{au} toward the singularity. Hence, k_2 should be sufficiently small. The analytic derivation of an upper bound on k_2 appears to be very difficult, due to the complexity of the dynamics (14). One possibility is to set $k_1 = 0$ until $\bar{t}$, or even setting it to a negative value (a *destabilizing* action), so as to keep e_{au} away from zero. From $\bar{t}$ on, k_1 can be set back to a positive value. In our experience, however, no special strategy of this kind

was necessary. Both in simulation and experiments, constant values of k_1 and k_2 were sufficient to achieve singularity avoidance.

3) Choice of $\hat{d}_0$: According to *Proposition 1*, it is necessary to initialize $\hat{d}$ at a value $\hat{d}_0 \geq d_0$. To this end, one may use an upper bound on d_0 derived from the knowledge of the environment where the robot moves. If d_0 is available (possibly through 3-D reconstruction techniques or *a priori* information) and used to initialize $\hat{d}$, then any positive value of k_1 and k_2 will guarantee global convergence of the epipoles and singularity avoidance.

V. Second Step: Matching the Features

At the end of the first step, both the desired and actual epipoles are zero and the intermediate robot configuration q_i is aligned with the desired configuration (see Fig. 5). We now present an image-based control law which uses feature points to realize the second step of our visual servoing strategy, i.e., moving the robot from q_i to q_d so as to recover the translation error. As the epipole-based controller of the previous section, also the second step controller works directly in the camera image plane.

The basic idea consists in translating the robot until each feature p_{ai} $(i = 1, \ldots, n)$ in the actual image plane matches the corresponding feature p_{di} in the desired image plane. A similar approach has been proposed in [3] and [34]. In principle, only one feature is needed to implement this idea, and therefore, we will present the controller for the case $n = 1$. The proposed method can be easily extended to include a larger number of features, a convenient choice in the case of noisy images.

Assuming that the chosen feature point is not located along the optical axis of the desired camera frame, denote by $D = \|p_a\|^2 - \|p_d\|^2$ the difference between the squared norms of the actual feature p_a and the desired feature p_d, a quantity which can be directly computed from the actual and the desired images.

Proposition 2: Let the robot velocities during the second step be defined as

$$u_1 = -k_t D \tag{21}$$

$$u_2 = 0 \tag{22}$$

where $k_t > 0$. Then, the robot configuration converges exponentially from the intermediate configuration q_i to the origin.

Proof: At the beginning of the second step, an unknown translation $t = [0 \;\; 0 \;\; \tilde{d}]^T$ occurs between the actual and desired camera frames (see Figs. 2 and 5). Here, $\tilde{d} = \pm d$, where d is the Euclidean distance between the actual and desired robot Cartesian positions. The minus (plus) sign should be taken when the actual camera frame is ahead (behind) the desired one, as in Fig. 5.

According to the model (1), with the expression of K in (2), the perspective projection of the feature point $P = [X \;\; Y \;\; Z]^T$ on the desired image plane is $p_d = K[I \mid 0]P$ (recall that the desired camera frame coincides with the world frame). In nonhomogeneous coordinates, we have

$$p_d = \begin{bmatrix} f\dfrac{X}{Z} & f\dfrac{Y}{Z} \end{bmatrix}^T.$$

Since the intermediate configuration is aligned with the desired configuration, and the angular velocity is identically zero during the second step, the perspective projection of P on the actual image plane is $p_a = K[I \mid t]P$, which gives in nonhomogeneous coordinates

$$p_a = \begin{bmatrix} f\dfrac{X}{Z+\tilde{d}} & f\dfrac{Y}{Z+\tilde{d}} \end{bmatrix}^T.$$

Hence, it is

$$D = \|p_a\|^2 - \|p_d\|^2 = -f^2(X^2+Y^2)\frac{\tilde{d}(2Z+\tilde{d})}{Z^2(Z+\tilde{d})^2} = -\alpha\tilde{d} \tag{23}$$

with $\alpha > 0$, since $Z > 0, Z + \tilde{d} > 0$ and $2Z + \tilde{d} > 0$ as a consequence of the assumption that the chosen feature point P is observable from both q_i and q_d.

We have thus expressed the difference D between the squared norms of the actual and the desired feature as a function of the signed distance $\tilde{d}$ between the actual and desired configurations. Note that when the feature point is along the optical axis of the desired camera frame, it is $X = Y = 0$, so that $D = 0$ regardless of the feature-point displacement Z along the optical axis. The bijectivity of the d/D map would be lost in that case.

Consider now the Lyapunov candidate

$$V = \frac{1}{2}\tilde{d}^2 = \frac{1}{2}d^2$$

which is positive definite. Under the proposed control law (21) and (22), we have $\dot{\tilde{d}} = -u_1$ and thus

$$\dot{V} = -\tilde{d}u_1 = k_t\tilde{d}D = -\alpha k_t d^2$$

where we have used (23). Since $\dot{V}$ is negative definite, exponential convergence of the robot configuration to the origin is guaranteed for any initial value of d, i.e., for any q_i. ■

A couple of remarks are now in order with respect to the control algorithm resulting from the sequence of the two steps so far presented. First of all, note that if the modified control law (20) is used, the first step has a finite duration. This fact, in addition to the exponential convergence guaranteed by the second step, allows us to claim exponential convergence for the entire two-step algorithm.

From a structural viewpoint, our two-step IBVS controller is time-invariant, but clearly discontinuous with respect to the robot state. However, this is not surprising, for it is well known that no smooth time-invariant stabilizing feedback exists for nonholonomic systems, and that only non-Lipschitz control laws (either time-invariant or not) can achieve exponential convergence [27]. Note finally that the structure of the proposed control law guarantees that switching from the first to the second step occurs only once, i.e., when the robot has reached the y axis.

VI. The Partially Calibrated Camera Case

So far, it has been assumed that all the camera intrinsic parameters are known and available for implementing the proposed two-phase controller. In this section, however, we show that our control scheme is effective even if the focal length f is unknown

(*partially calibrated* camera), a desirable property in any visual servoing strategy.

First of all, note that only the first step (zeroing the epipoles) makes use of the focal length f for computing the control inputs (16) and (17), as well for updating the $\hat{d}$ parameter through (13), while the second step (matching the features) is insensitive to its value, as shown by (21) and (22).

Assume that f is unknown, so that a constant value $\hat{f} \neq f$ is used to perform the approximate input–output linearization of Section IV-B. The obtained control inputs are [compare with (11)]

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \tilde{E}^{-1} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \tag{24}$$

with

$$\tilde{E}^{-1} = \begin{bmatrix} 0 & -\text{sign}(e_{au}e_{du})\frac{\hat{d}\hat{f}\sqrt{e_{au}^2+\hat{f}^2}}{e_{au}(e_{du}^2+\hat{f}^2)} \\ \frac{\hat{f}}{e_{au}^2+\hat{f}^2} & -\frac{\hat{f}}{e_{du}^2+\hat{f}^2} \end{bmatrix}.$$

We have the following result.

Proposition 3: Let the auxiliary controls ν_1, ν_2 be chosen as in *Proposition 1*, and update the distance estimate $\hat{d}$ using $\hat{f}$ for f, i.e., according to

$$\dot{\hat{d}} = k_2 \hat{f}^2 \hat{d} \frac{e_{du}^{\beta/\gamma}}{e_{au}\left(e_{du}^2+\hat{f}^2\right)} \tag{25}$$

initialized at $\hat{d}_0 > d_0$. If $\hat{f} \leq f$ and k_2 is sufficiently small, the approximately linearizing control (24) drives the epipole coordinates e_{au} and e_{du} to zero for any initial condition such that $e_{au}(0) \neq 0$, with exponential convergence rate.

Proof: The closed-loop equations under the modified control law become [compare with (14) and (15)]

$$\dot{e}_{au} = -\frac{\hat{f}}{f}\frac{e_{au}^2+f^2}{e_{au}^2+\hat{f}^2}k_1 e_{au} - \frac{\hat{f}}{f}\frac{\sqrt{e_{au}^2+f^2}}{e_{du}^2+\hat{f}^2} \cdot \left(\frac{\hat{d}}{d}\sqrt{e_{au}^2+\hat{f}^2} - \sqrt{e_{au}^2+f^2}\right) k_2 e_{du}^{\beta/\gamma} \tag{26}$$

$$\dot{e}_{du} = -\frac{\hat{f}}{f}\frac{e_{du}^2+f^2}{e_{du}^2+\hat{f}^2}\sqrt{\frac{e_{au}^2+\hat{f}^2}{e_{au}^2+f^2}}\frac{\hat{d}}{d}k_2 e_{du}^{\beta/\gamma} \tag{27}$$

while the control inputs are computed as in (16) and (17) with $\hat{f}$ in place of f. To guarantee that e_{du} still goes to zero in finite time, it is necessary to show that the coefficient of $e_{du}^{\beta/\gamma}$ in (27) is negative and bounded below in modulus.

It is easy to prove the following inequalities:

$$\frac{e_{du}^2+f^2}{e_{du}^2+\hat{f}^2} \geq \min\left(1, \frac{f^2}{\hat{f}^2}\right) = 1$$

$$\sqrt{\frac{e_{au}^2+\hat{f}^2}{e_{au}^2+f^2}} \geq \min\left(1, \frac{\hat{f}}{f}\right) = \frac{\hat{f}}{f}$$

where we have used $\hat{f} \leq f$. Hence

$$\frac{\hat{f}}{f}\frac{e_{du}^2+f^2}{e_{du}^2+\hat{f}^2}\sqrt{\frac{e_{au}^2+\hat{f}^2}{e_{au}^2+f^2}}\frac{\hat{d}}{d}k_2 \geq \frac{\hat{f}^2}{f^2}\frac{\hat{d}}{d}k_2$$

and it only remains to be shown that $\hat{d}/d$ does not become or tend to zero. The distance d evolves according to

$$\dot{d} = u_1 \cos(\psi - \theta) = k_2 f \hat{f} \hat{d} \sqrt{\frac{e_{au}^2+\hat{f}^2}{e_{au}^2+f^2}} \frac{e_{du}^{\beta/\gamma}}{e_{au}(e_{du}^2+\hat{f}^2)}. \tag{28}$$

Comparing this with (25), we have

$$\dot{d} = \frac{f}{\hat{f}}\sqrt{\frac{e_{au}^2+\hat{f}^2}{e_{au}^2+f^2}}\dot{\hat{d}} = \mu\dot{\hat{d}}. \tag{29}$$

Simple manipulations allow verifying that the time-varying coefficient μ is bounded above and below

$$1 \leq \mu \leq \frac{f}{\hat{f}}. \tag{30}$$

Therefore:

- $\hat{d}/d$ cannot become zero at a finite time instant, because $\hat{d}_0 > d_0$, and if $\hat{d}$ decreases ($\dot{\hat{d}} < 0$), then d decreases faster ($\dot{d} = \mu\dot{\hat{d}} < \dot{\hat{d}} < 0$);
- $\hat{d}/d$ cannot tend to zero because (29) and (30) imply that: 1) if $\hat{d}$ tends to zero, then d tends to zero with the same speed; 2) if d tends to infinity, then $\hat{d}$ tends to infinity with the same speed.

Having proven that zero is still a terminal attractor for e_{du}, which will converge to zero at a finite time instant $\bar{t}$, the rest of the proof is exactly the same as that of *Proposition 1*. In particular, note that from time $\bar{t}$ on, the differential (26) governing e_{au} reduces to

$$\dot{e}_{au} = -\frac{\hat{f}}{f}\frac{e_{au}^2+f^2}{e_{au}^2+\hat{f}^2}k_1 e_{au} \quad \text{with } \frac{e_{au}^2+f^2}{e_{au}^2+\hat{f}^2} > 1$$

so that e_{au} will converge to zero with exponential rate *at least* equal to k_1. ■

VII. Simulations

In this section, we present simulation results of our IBVS strategy for nonholonomic mobile robots in the calibrated and partially calibrated case. Simulations have been performed using Matlab-Simulink and the Epipolar Geometry Toolbox.[1]

Eight pairs of corresponding feature points are identified in the desired and actual images. They are used for epipole reconstruction in the first step, and directly in the second step.

In the first simulation, it is assumed that the camera has been calibrated in advance (see Section II) and that, in particular, $f = 0.03$ m. The unicycle robot moves under the action of the proposed two-step visual strategy from its initial configuration $q_0 = [1\ \ 1\ \ \pi/2]^T$. First, the control law (16) and (17) is applied, with $k_1 = 0.5$, $k_2 = 0.5$, and $\beta/\gamma = 5/9$. The initial estimate of the robot distance has been set to $\hat{d}_0 = 2$ m. As shown in Fig. 7, both the epipoles coordinates e_{au} and e_{du} are driven to zero (note that, as expected, the second is zeroed in finite time). The resulting control inputs are in Fig. 8, while the

[1]The Epipolar Geometry Toolbox, developed at the University of Siena, Siena, Italy, allows the creation of single and multicamera scenarios as well as the computation of visual information and the estimation of epipolar geometry. See http://egt.dii.unisi.it for additional information.

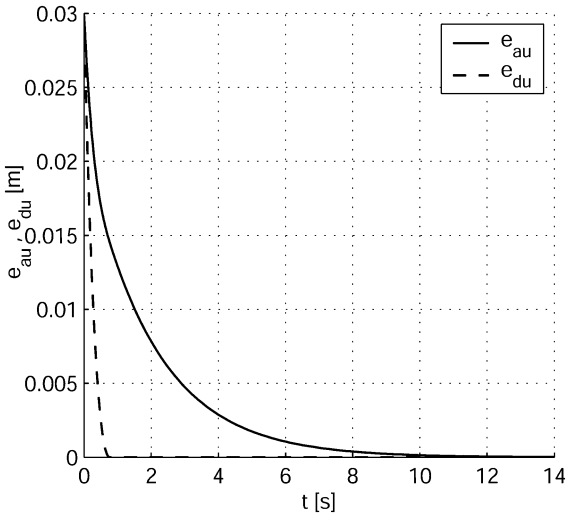

Fig. 7. Simulation results, calibrated camera. First step: Epipole behavior. Note how the desired epipole coordinate e_{du} reaches zero in finite time.

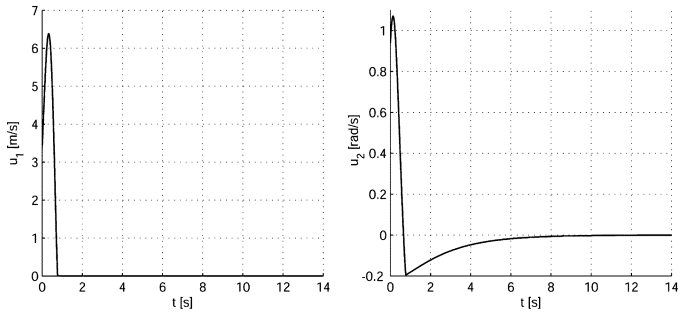

Fig. 8. Simulation results, calibrated camera. First step: Linear (left) and angular (right) velocity of the mobile robot. Note how the linear velocity goes to zero with e_{du}.

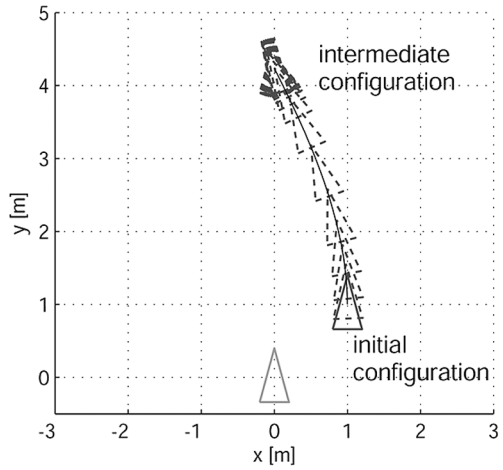

Fig. 9. Simulation results, calibrated camera. First step: Robot trajectory.

robot trajectory for the first step is given in Fig. 9. To guarantee a finite time duration for the first step, a tolerance of 10^{-9} m has been used for e_{au}.

The second step is then executed under the action of the control law (21) and (22), with $k_t = 100$. The exponential decrease of the distance d between the actual and desired robot positions

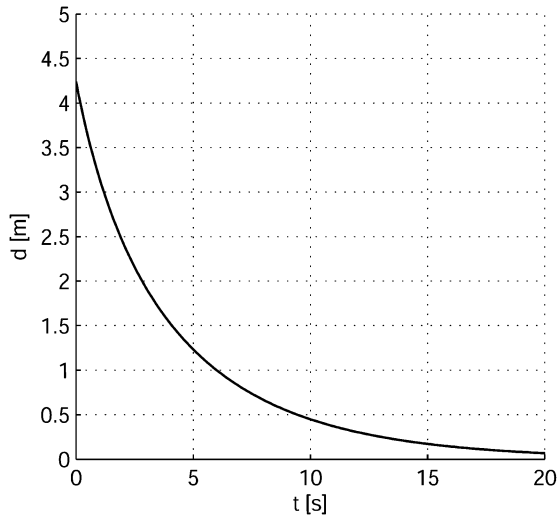

Fig. 10. Simulation results, calibrated camera. Second step: Exponential decrease of the distance d between the actual and the desired robot position.

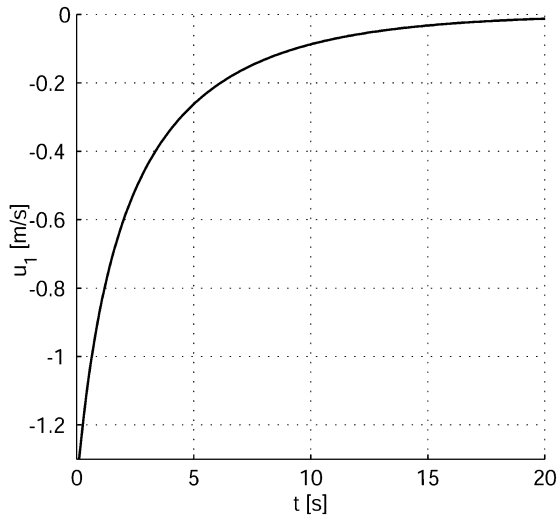

Fig. 11. Simulation results, calibrated camera. Second step: Linear velocity.

is shown in Fig. 10. The linear velocity u_1 and the robot trajectory are given in Figs. 11 and 12, respectively.

We also report in Fig. 13 the overall motion of the feature points in the actual image plane. Note how during the initial phase of the first step, the distance D between the actual and desired features tends to increase as the robot moves toward the y axis. As the actual epipole coordinate e_{au} goes to zero (i.e., the robot rotates to align with the y axis), this distance decreases. Then, during the second step, D is brought to zero.

The second simulation is executed with exactly the same parameters and initial conditions of the first. However, the actual value $f = 0.03$ m of the focal length is assumed to be unknown, and a very imprecise estimate $\hat{f} = 0.005$ m has been used in the control law (and in the update of $\hat{d}$). As shown by the epipole behavior in Fig. 14 and the Cartesian trajectory in Fig. 15, the first step of the proposed control strategy still works as expected (recall that the second-step controller does not need the focal length).

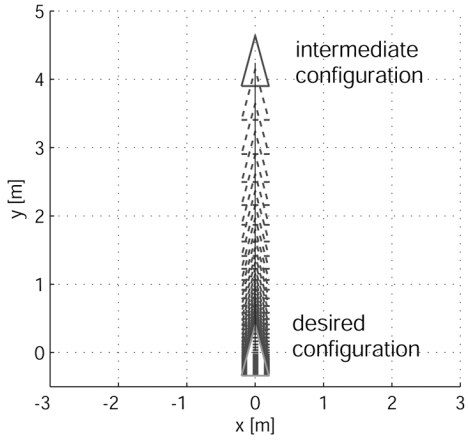

Fig. 12. Simulation results, calibrated camera. Second step: Robot trajectory.

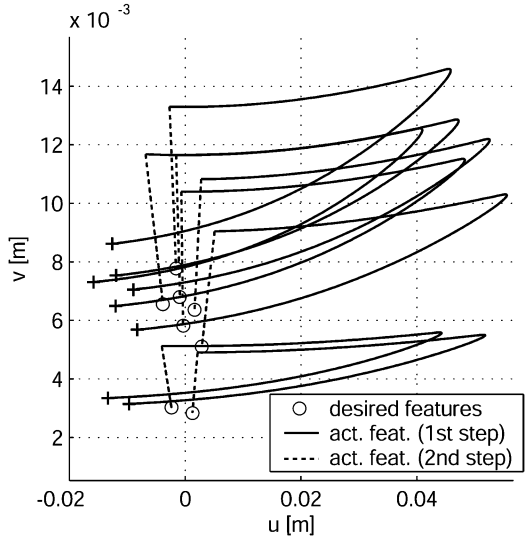

Fig. 13. Simulation results, calibrated camera. Motion of the feature points on the image plane from the initial position $(+)$ to the desired one $(*)$.

VIII. Experimental Results

To further validate the proposed visual servoing strategy, experimental results are presented next. The vision sensor is the pinhole camera LU175C by Lumenera. The mobile robot is a Pioneer 3X-DE by ActivMedia (see Fig. 16), connected to a notebook with a 2-GHz Pentium 4 processor and 640 MB of RAM. The pinhole camera has been coarsely calibrated, obtaining $f = 0.004$ m, $u_0 = 635$, $v_0 = 512$ pixels. The image resolution has been fixed at 1280×1024 pixels.

The desired image has been acquired in advance from the target configuration, while the current image is available at each sampling time. A set of 26 corresponding feature points (Fig. 19) was chosen in the two images and then tracked in real time by means of a Lucas–Kanade pyramidal algorithm [1].

It is well known that epipolar geometry computation should not be performed directly from raw pixel data, due to the uncertainties given by huge numbers. Therefore, feature points are normalized using Hartley's method [12]. The normalized features are then used to compute the epipoles by means of the LMedS estimator [2], which is known to be computationally efficient and robust to image noise and outliers. The overall sampling time of the servoing scheme is about 0.09 s.

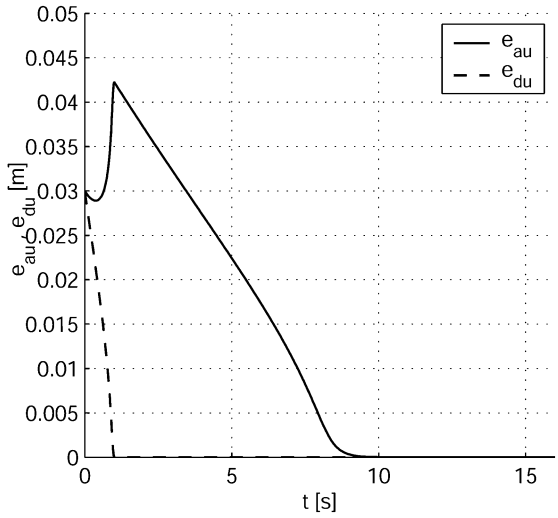

Fig. 14. Simulation results, partially calibrated camera. First step: Epipole behavior.

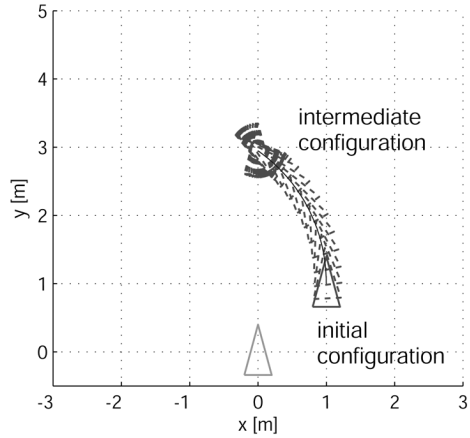

Fig. 15. Simulation results, partially calibrated camera. First step: Robot trajectory.

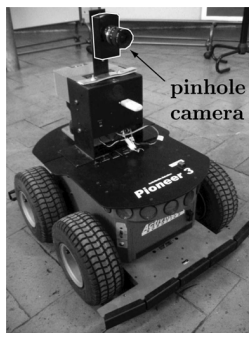

Fig. 16. Mobile robot Pioneer 3X-DE used for the experiments, equipped with the LU175C pinhole camera by Lumenera.

The robot moves under the action of the two-step visual strategy. First, control law (16) and (17) is applied, with $k_1 = 0.08$, $k_2 = 1$ and $\beta/\gamma = 7/9$. The initial estimate of the robot distance is $\hat{d}_0 = 12$ m. Both the epipoles e_{au} and e_{du} are driven to zero, as shown in Fig. 17. As expected, e_{du} is zeroed in finite time. The same is true for e_{au}, due to the fact that angular velocities below a certain threshold are set back to the threshold to avoid the robot actuator deadzone.

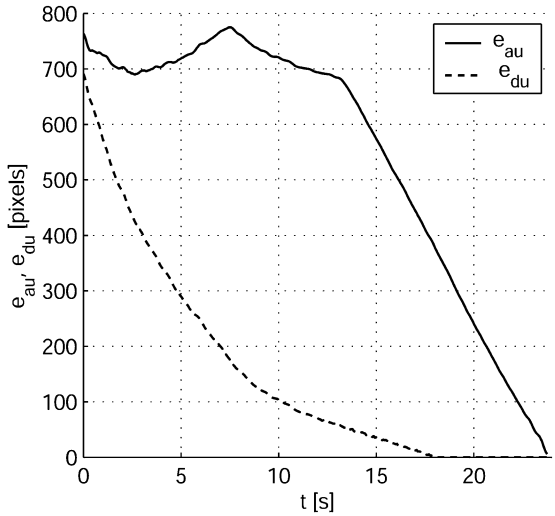

Fig. 17. Experimental results. First step: Epipole behavior.

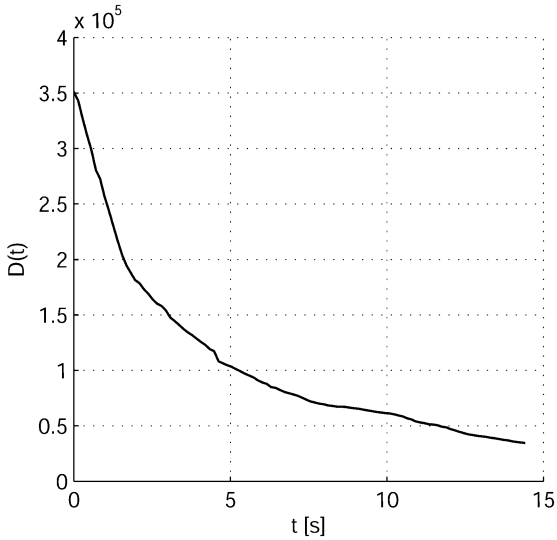

Fig. 18. Experimental results. First step: Exponential decrease of the average difference D between the square norms of the actual and desired features.

Then, the second step is executed under the action of the control law (21) and (22), with $k_t = 4 \cdot 10^{-4}$ (note that the distance between features are measured in pixels in the experiments). The exponential decrease of D (the average difference between the square norms of the actual and desired features) is shown in Fig. 18. Note that, due to floor irregularities and wheel slippage, D does not decrease exactly to zero.

Figs. 19 and 20 collect six snapshots of the robot motion during the first and second steps, respectively. Note the "ghost" robot representing the target. On the right of each snapshot, the actual (red cross) and desired (green circle) feature points are shown superimposed to the desired image.

During the first step (Fig. 19), the epipoles are driven to the principal point. Note that the epipoles are outside the desired image at the initial position. The second step (Fig. 20) is then executed, and the actual feature points converge to their target. The overall servoing performance is satisfactory, resulting in a positioning error of about 4 cm with respect to the target position.

Video clips of this experiment are available at http://sirslab.dii.unisi.it/vision/research.html.

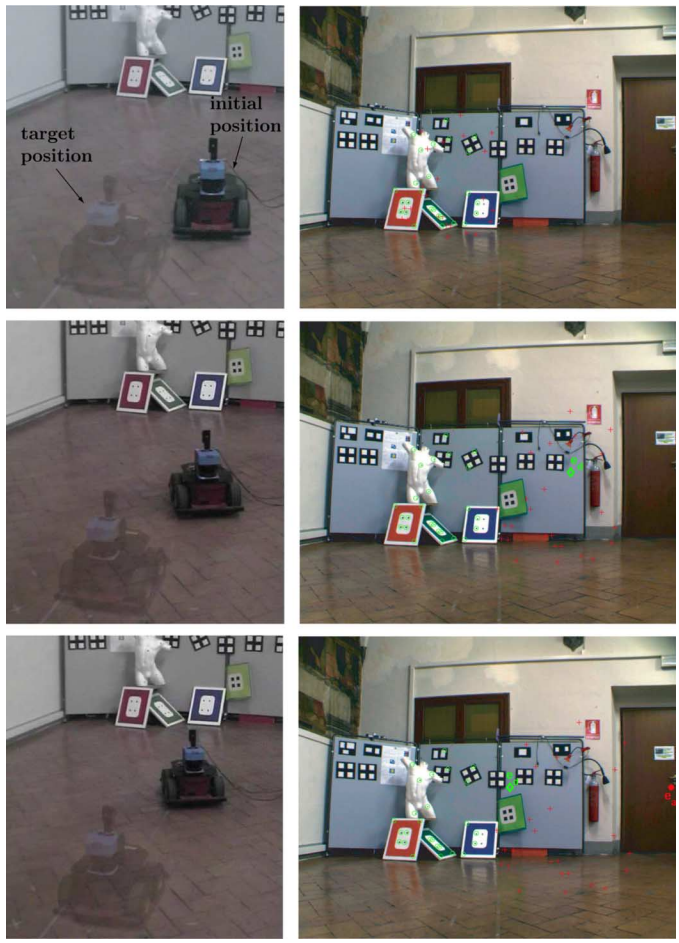

Fig. 19. Experimental results. Snapshots of the robot motion (left) during the first step. Also shown is the epipole/feature motion (right) superimposed on the desired image. In particular, red crosses denote the actual features, while green circles identify the desired features.

Note finally that degeneracies in epipole estimation, which may occur when the actual and desired views are very close (small baseline), were never met in our experiments. This is essentially due to the fact that our two-step IBVS scheme estimates and uses the epipoles only in the first phase, in which the translation error between the two views is typically relevant. In any case, one way to address the degeneracy problem is to use more than two views, as proposed in [35].

IX. Some Extensions to the Basic Problem

We now mention some directions of possible extension of the problem considered in this paper. All of them are the subject of our current work.

1) Keeping the Features in the Field of View: This important issue is not addressed in this paper. One way to alleviate the problem is to exploit the possibility of estimating the epipolar geometry from other image data (e.g., contours or optical flow) than feature points. In principle, one could conceive an algorithm that selects different image data at each sampling interval to guarantee a continuous estimation. To address more effectively the field-of-view issue, however, two possible strategies can be envisaged. The first consists of using panoramic instead of pinhole cameras, and the second of adding a pan-tilt head. While the first case involves the analysis of epipolar geometry

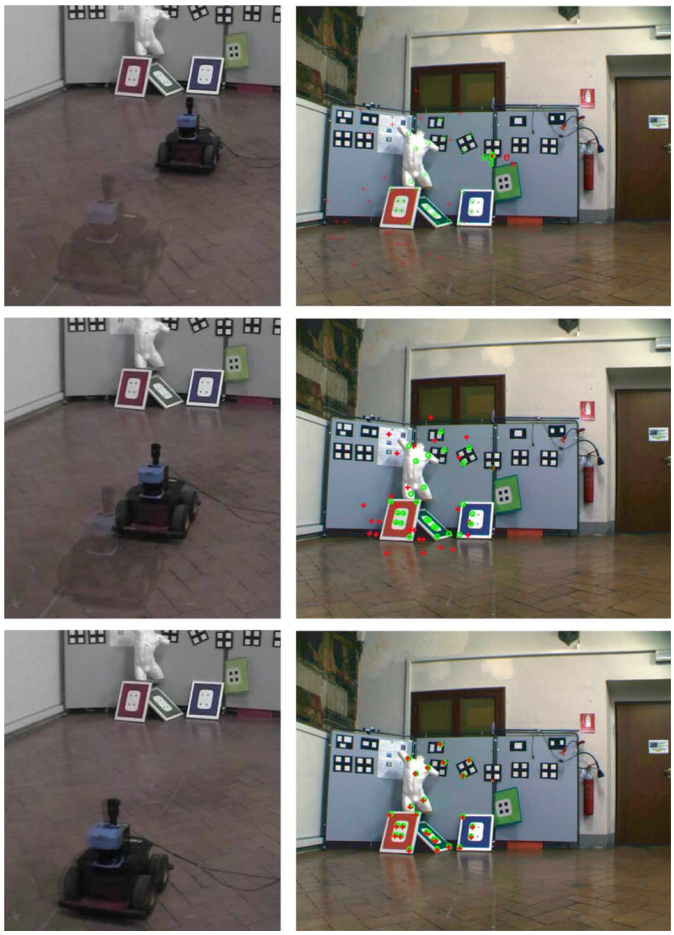

Fig. 20. Experimental results. Snapshots of the robot motion and epipole/feature motion during the second step.

for panoramic cameras [25], [36], a camera-robot system with a pan-tilt head can be easily treated within the control framework proposed in this paper. The key observation here is that from the features/epipoles in the pan-tilt camera, it is possible to compute the equivalent features/epipoles in a virtual fixed camera, provided that the pan and tilt angles are known. Hence, the pan-tilt head can be controlled so as to keep the features in the field of view, while the equivalent features/epipoles are used to implement the control law proposed in this paper.

Different Camera Arrangements and/or Robot Kinematics: A further extension to the basic problem is the case where the camera center is displaced with respect to the center of instantaneous rotation of the unicycle, either due to incorrect alignment or to a design choice. It is interesting to note that such an arrangement would lead to a simpler control design. In fact, a point displaced from the center of the unicycle can be driven along arbitrary Cartesian trajectories [32]. This means that the camera center can follow arbitrary velocities (in particular, those orthogonal to the robot forward axis), making it possible to decrease instantaneously[2] the value of e_{du} even if $e_{au} = 0$. The implication of this fact is that no singularity would occur in the feedback linearization design. Since the possibility of driving displaced points along arbitrary Cartesian trajectories holds also for more general robot kinematics [9], it should be possible to extend the proposed method, e.g., to car-like robots.

2) Robustness to Model Perturbations: Our two-step visual-regulation scheme was designed and analyzed using the nominal kinematic model (4). While the experimental results of the previous section have shown that the performance is satisfactory in real-world conditions, an analytical study of the performance of the method in the presence of model perturbations has not been conducted in this paper. For example, in a differential-drive kinematic arrangement, unequal wheels or asymmetric actuation could result in a nonnegligible rotation of the robot, even when the nominal angular velocity is zero (as in the second step of our control scheme). This situation can be formalized and studied by introducing an input perturbation in model (4). One possible way we are considering to make our strategy more robust is to apply the two-step controller in an iterative fashion, following the approach proposed in [17] and specialized to general nonholonomic robots in [28].

X. Conclusions

A novel IBVS strategy has been presented for nonholonomic mobile robots. The control scheme, which is divided in two independent and sequential steps, drives the robot to a desired configuration specified through the corresponding view, previously acquired by the on-board camera. A key point is the use of multiple-view epipolar geometry during the first step in order to compensate the rotational error and align the current view to the desired one. An approximate input–output linearizing feedback is used to cope with the nonholonomy of the camera-robot system.

The advantages of our approach are summarized hereafter.

- Differently from PBVS techniques, there is no need to assume any *a priori* knowledge of the 3-D scene structure. In fact, feature points, image contours, or optical flow can be used to robustly estimate the epipoles.
- In comparison with 2-1/2-D visual servoing techniques, our algorithm does not need any partial camera pose-estimation phase.
- Global exponential convergence is obtained for any constant overestimation of the robot Cartesian initial distance to the desired configuration.
- The result also holds in the case of a partially calibrated pinhole camera (unknown focal length).

Simulations and experiments on a real robot have shown the practical applicability and effectiveness of the proposed visual servoing algorithm, and a number of possible extensions have been discussed.

[2]This is not possible when the camera displacement is zero, for in this case, the only feasible instantaneous motions when $e_{au} = 0$ are translations along the baseline and rotations about the robot center. None of these motions moves the baseline, and thus, the desired epipole e_{du} would not change.

References

[1] Intel's OpenCV [Online]. Available: http://sourceforge.net/projects/opencvlibrary/
[2] X. Armangue and J. Salvi, "Overall view regarding fundamental matrix estimation," *Image Vis. Comput.*, vol. 21, no. 2, pp. 205–220, 2003.
[3] R. Basri, E. Rivlin, and I. Shimshoni, "Visual homing: Surfing on the epipoles," *Int. J. Comput. Vis.*, vol. 33, no. 2, pp. 117–137, 1999.
[4] D. Burschka and G. Hager, "Vision-based control of mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2001, pp. 1707–1713.

[5] G. Campion, B. d'A. Novel, and G. Bastin, "Modeling and state feedback control of nonholonomic mechanical systems," in *Proc. 30th IEEE Conf. Decision, Control*, 1991, pp. 1184–1189.
[6] F. Chaumette, "Image moments: A general and useful set of features for visual servoing," *IEEE Trans. Robot.*, vol. 20, no. 4, pp. 713–723, Aug. 2004.
[7] G. Chesi, G. Mariottini, D. Prattichizzo, and A. Vicino, "Epipole-based visual servoing for mobile robots," *Adv. Robot.*, vol. 20, no. 2, pp. 255–280, 2006.
[8] F. Conticelli, B. Allotta, and P. K. Khosla, "Image-based visual servoing of nonholonomic mobile robots," in *Proc. 38th IEEE Conf. Decision, Control*, 1999, pp. 3496–3501.
[9] A. De Luca, G. Oriolo, and C. Samson, "Feedback control of a nonholonomic car-like robot," in *Robot Motion Planning and Control*, J. P. Laumond, Ed. New York: Springer, 1998, vol. 229, LNCIS, pp. 171–253.
[10] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," in *Proc. 3rd Int. Symp. Exp. Robot.*, 1993, pp. 182–192.
[11] Y. Fang, W. E. Dixon, D. M. Dawson, and P. Chawda, "Homography-based visual servo regulation of mobile robots," *IEEE Trans. Syst., Man, Cybern. B: Cybern.*, vol. 35, no. 5, pp. 1041–1050, Oct. 2005.
[12] R. Hartley, "In defence of the 8-point algorithm," in *Proc. IEEE Int. Conf. Comput. Vis.*, 1995, pp. 1064–1070.
[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[14] K. Hashimoto and T. Noritsugu, "Visual servoing of nonholonomic cart," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1997, pp. 1719–1724.
[15] S. A. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.
[16] A. Isidori, *Nonlinear Control Systems*. New York: Springer, 1995.
[17] P. Lucibello and G. Oriolo, "Robust stabilization via iterative state steering with an application to chained-form systems," *Automatica*, vol. 37, pp. 71–79, 2001.
[18] Q. T. Luong and O. D. Faugeras, "The fundamental matrix: Theory, algorithms and stability analysis," *Int. J. Comput. Vis.*, vol. 17, no. 1, pp. 43–76, 1996.
[19] Y. Ma, J. Košecká, and S. S. Sastry, "Vision-guided navigation of nonholonomic mobile robot," *IEEE Trans. Robot. Autom.*, vol. 15, no. 3, pp. 521–536, Jun. 1999.
[20] Y. Ma, S. Soatto, J. Košecká, and S. S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*. New York: Springer, 2003.
[21] E. Malis and F. Chaumette, "2-1/2-D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *Int. J. Comput. Vis.*, vol. 37, no. 1, pp. 79–97, 2000.
[22] ——, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," *IEEE Trans. Robot. Autom.*, vol. 18, no. 2, pp. 176–186, Apr. 2002.
[23] E. Malis, F. Chaumette, and S. Boudet, "2-1/2-D visual servoing," *IEEE Trans. Robot. Autom.*, vol. 15, no. 2, pp. 238–250, Apr. 1999.
[24] G. L. Mariottini, G. Oriolo, and D. Prattichizzo, "Epipole-based visual servoing for nonholonomic mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2004, pp. 497–503.
[25] ——, "Image-based visual servoing for nonholonomic mobile robots with central catadioptric camera," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2006, pp. 497–503.
[26] Y. Masutani, M. Mikawa, N. Maru, and F. Miyazaki, "Visual servoing for non-holonomic mobile robots," in *Proc. IEEE Conf. Intell. Robots Syst.*, 1994, vol. 2, pp. 1133–1140.
[27] R. T. M'Closkey and R. M. Murray, "Exponential stabilization of driftless control systems using homogeneous feedback," *IEEE Trans. Autom. Control*, vol. 42, no. 5, pp. 614–628, May 1997.
[28] G. Oriolo and M. Vendittelli, "A framework for the stabilization of general nonholonomic systems with an application to the plate-ball mechanism," *IEEE Trans. Robot.*, vol. 21, no. 2, pp. 162–175, Apr. 2005.
[29] J. Piazzi, D. Prattichizzo, and A. Vicino, "Visual servoing along epipoles," in *Control Problems in Robotics*, A. Bicchi, Ed. *et al.* New York: Springer, 2003, vol. 4 of STAR, pp. 215–232.
[30] R. Pissard-Gibollet and P. Rives, "Applying visual servoing techniques to control a mobile hand-eye system," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1995, pp. 166–171.
[31] P. Rives, "Visual servoing based on epipolar geometry," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2000, vol. 1, pp. 602–607.
[32] C. Samson and K. Ait-Abderrahim, "Feedback control of a nonholonomic wheeled cart in Cartesian space," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1991, pp. 1136–1141.
[33] J. Sato and R. Cipolla, "Affine reconstruction of curved surfaces from uncalibrated views of apparent contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1186–1198, Nov. 1999.
[34] T. Sato and J. Sato, "Visual servoing from uncalibrated cameras for uncalibrated robots," *Syst. Comput. Jpn.*, vol. 31, no. 14, pp. 11–19, 2000.
[35] O. Shakernia, R. Vidal, and S. Sastry, "Structure from small baseline motion with central panoramic cameras," in *Proc. 4th Workshop Omnidirectional Vis.*, 2003 [Online]. Available: http://www.icm.jhu.edu/publications/vidalpubs.html
[36] T. Svoboda and T. Pajdla, "Epipolar geometry for central catadioptric cameras," *Int. J. Comput. Vis.*, vol. 49, no. 1, pp. 23–37, 2002.
[37] A. Verri and E. Trucco, "Finding the epipole from uncalibrated optical flow," *Image Vis. Comput.*, vol. 17, no. 8, pp. 605–609, 1999.
[38] M. Zak, "Terminal attractors in neural networks," *Neural Netw.*, vol. 2, pp. 259–274, 1989.

Gian Luca Mariottini (S'04–M'06) received the M.S. degree (*cum laude*) in information engineering in 2002, and the Ph.D. degree in robotics and automation in 2006, both from the University of Siena, Siena, Italy.

He was a Visiting Scientist with the GRASP Lab of the University of Pennsylvania, Philadelphia, in 2005, and is again in 2007. Since October 2005, he has been an Associate Researcher with the Department of Engineering Information, University of Siena. His research interests include robotics and computer vision, especially visual servoing, multirobot localization and control, panoramic cameras, and human-robot interfaces for medical applications.

Giuseppe Oriolo (S'89–M'91–SM'02) received the Ph.D. degree in control systems engineering in 1992 from the University of Rome ''La Sapienza,'' Rome, Italy.

Since 1998, he has been an Associate Professor of Automatic Control, Department of Computer and System Science of the same university, where he also is in charge of the Robotics Laboratory. His research interests are in the area of robotics and control, in which he has published more than 100 papers.

Domenico Prattichizzo (M'95) received the M.S. degree in electronics engineering and the Ph.D. degree in robotics and automation from the University of Pisa, Pisa, Italy, in 1991 and 1995, respectively.

Since 2002, he has been an Associate Professor of Robotics and Automation at the University of Siena, Siena, Italy. In 1994, he was a Visiting Scientist with the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge. He is co-editor of the books *Control Problems in Robotics (STAR, Springer Tracks in Advanced Robotics)* (New York: Springer Verlag, 2003, vol. 4) and *Multi-Point Physical Interaction with Real and Virtual Objects, (STAR, Springer Tracks in Advanced Robotics)* (New York: Springer Verlag, 2005, vol. 18). He is co-author of more than 100 papers in the area of robotics and automation. His research interests are in the area of visual servoing, robotic grasping, haptics, and geometric control.

Dr. Prattichizzo was Co-Chair of the 2nd IEEE International Workshop on Control Problems in Robotics and Automation (Las Vegas, NV, 2002) and Program Co-Chair of the 1st World Haptics Conference (Pisa, Italy, 2005). He serves as a Member of the Editorial Board of the IEEE TRANSACTIONS ON ROBOTICS, the IEEE TRANSACTIONS ON CONTROL AND SYSTEMS TECHNOLOGIES, and of the *Journal of Dynamics of Continuous, Discrete and Impulsive Systems (DCDIS), Series B: Application and Algorithms*.