Adaptive Nonlinear Visual Servoing Using Lyapunov-Based Design

Fabio Conticelli* Benedetto Allotta†
PERCRO Lab
Scuola Superiore Sant'Anna
Via Carducci 40, I-56127 Pisa, ITALY
tel +39-50-883445
fax +39-50-883210

Abstract

In this paper, an adaptive nonlinear control scheme is designed to solve the problem of controlling the relative pose between a robot camera and a rigid object. The image-based visual system of the camera-object interaction is expressed in terms of global coordinates fully defined in the image plane, then a discrete time interaction model is derived, since the visual sampling time is not negligible at the actual state of technology. By exploiting nonlinear controllability properties, a nonlinear control law is designed based on Lyapunov's direct method. Moreover, we propose a 3-D estimation procedure based on prediction errors to cope with the unknown depth of the object. Experimental results with a 6-DOF robot manipulator in eye-in-hand configuration validate the theoretical framework.

1 Introduction

Visual servoing systems use camera sensors inside the control loop to accomplish tasks in unstructured environments (Hill and Park, 1979). This approach can be used to improve, using noncontact measurements, the adaptability of robotic systems with respect to both environmental uncertainties in positioning tasks, and inaccuracy of the robot's kinematic model. Typical applications are robotic manipulation, teleoperation and mobile robots. Concerning the control aspects, two main paradigms can be outlined: position-based and image-based servoing (Espiau *et al.*, 1992), (Allotta *et al.*, 1998). In the first one, the error is defined in 3D space based on image feature extraction and relative pose estimation. In image-based servoing, instead, any visual task is described in the image plane as a desired evolution of object appearance towards a goal one, since the error is computed in terms of image features. Although image-based systems have been largely investigated and they are now well established, few approaches have been proposed to study formally controllability properties and 3-D parameters estimation. In (Hashimoto *et al.*, 1996), the controllability of a linearizated visual system is studied, and a visual tracking system is proposed to cope with redundant features based on a linear quadratic (LQ) method. In (Papanikolopoulos and Khosla, 1993), adaptive control schemes are formulated to track image features with an eye-in-hand system using optimal control approach, however, stability analysis of the visual closed-loop system is not carried out. In this paper, adaptive nonlinear design is

*Email: *contice@sssup.it*
†Email: *ben@sssup.it*

applied to image-based visual systems, to solve the problem of robot camera-object relative positioning. The state space manifold in which the system is defined is the Special Euclidean group, the image-based visual system is expressed in global coordinates fully defined in the image plane, and nonlinear controllability analysis is carried out. Then a discrete time interaction model is derived, since the visual sampling time is not negligible at the actual state of technology. In structured environments, the system a priori knows the depth of object's points, in this case no adaptation law is needed, and the designed control system ensures global uniform asymptotic stability of the image plane reference setpoint. In unstructured environments, the depth of object's points is not known, and to find a stabilizing feedback is a difficult task to accomplish, since the system is MIMO (multi input-multi output), nonlinear, and with time-varying unknown parameters. We use a 3-D estimation procedure based on prediction errors (Slotine and Li, 1993) to cope with the unknown depth of the object.

Since the state space coordinates are fully defined in the image space, allowing 3-D parameters adaptation from 2-D information, the system does not require any calibration procedure. The image-based visual system is independent of the number of features (i.e. image points) being tracked, thus allowing acquisition of redundant visual features, which guarantees full-rank condition of the interaction matrix. Experimental results, obtained with a robotic system consisting in a PUMA 560 endowed with a camera on its wrist (eye-in-hand configuration), show that system performance is satisfactory in the positioning with respect to generic objects.

The paper is organized as follows. In Sect. 2, visual modeling issues are addressed, image-based coordinates are introduced, and nonlinear controllability properties of the image-based system are shown. In Sect. 3, the adaptive nonlinear control system is designed and stability analysis is carried out. Sect. 4 reports results of robotic experiments carried out to validate the theoretical framework. Finally, in Sect. 5 the major contribution of the paper is summarized.

Notations $\mathcal{A} \setminus \mathcal{B}$ denotes subtraction between sets $\mathcal{A}$ and $\mathcal{B}$. $I_n \in \Re^{n\times n}$ denotes the identity matrix. $0_{n\times m} \in \Re^{n\times m}$ is a matrix of zeros. $diag(x_i) \in \Re^{n\times n}$ is the diagonal matrix with the elements of $\boldsymbol{x} = [x_1, \ldots, x_n]^{\mathrm{T}} \in \Re^{\mathrm{n}}$. $so(3)$ denotes the vector space of the skew-symmetric matrices, an element of $so(3)$ is indicated with $\boldsymbol{x}^{\wedge}$, where $\boldsymbol{x} \in \Re^3$. Given a vector space V on the field $\Re$, $\forall\, \boldsymbol{v}, \boldsymbol{w} \in V$, $< \boldsymbol{v}, \boldsymbol{w} >$ denotes their euclidean scalar product. $\oplus$ denotes the direct sum of vector spaces. Given a nonlinear input-affine smooth system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \sum_1^m g_i(\boldsymbol{x})u_i$, $\boldsymbol{x} \in \mathcal{X}$ the state space manifold, and $\boldsymbol{u} \in U$ the set of piecewise constant inputs, $T_{\boldsymbol{x}}\mathcal{X}$ denotes the tangent space of $\mathcal{X}$ at $\boldsymbol{x}$. $F_{*\boldsymbol{x}}$ indicates the tangent map at $\boldsymbol{x}$ of $F : \mathcal{M} \to \mathcal{N}$, being $\mathcal{M}$, and $\mathcal{N}$ smooth manifolds. $V^{\infty}(\mathcal{X})$ indicates the Lie algebra of smooth vectors fields on $\mathcal{X}$, $\mathcal{C}$ denotes the accessibility algebra (i.e the smallest subalgebra of $V^{\infty}(\mathcal{X})$ that contains $f, g_1, \ldots, g_m$), $C(\boldsymbol{x}) = \{\boldsymbol{f}(\boldsymbol{x}) : \boldsymbol{f}(\boldsymbol{x}) \in \mathcal{C}\}$ is the associated accessibility distribution.

2 Image-Based Visual Modeling

In this section, the image-based visual system of camera-object interaction is derived in the state space form, and controllability properties are analyzed.

2.1 State Space Representation

Assume that the object of interest is rigid, the relative motion between the camera and the object is a rigid body transformation (Murray *et al.*, 1994), the configuration space of the camera-object relative motion is the Special Euclidean group $SE(3) = \Re^3 \times SO(3)$, where $SO(3)$ indicates the Special Orthogonal group. In the following quantities will be expressed

in terms of the camera frame $< c >= \{\boldsymbol{i}_c, \boldsymbol{j}_c, \boldsymbol{k}_c\}$. Let Γ be the region of the object's surface visible from the camera, and $\boldsymbol{x}_i = [x_i,\ y_i,\ z_i]^{\mathrm{T}}, i = 1, \ldots, n \in \Re^3$ be n points of Γ, which define global coordinates of the configuration space $\mathcal{X} = SE(3)$ (i.e at each $\boldsymbol{x} = [\boldsymbol{x}_1^{\mathrm{T}}, \ldots, \boldsymbol{x}_n^{\mathrm{T}}]^{\mathrm{T}} \in \Gamma$ corresponds only one valid configuration $g(\boldsymbol{x}) \in \mathcal{X}$, and $\forall g \in \mathcal{X}$, $\exists_1 \boldsymbol{x}$ with which can be identified). In the following, to simplify notation, the statement $\boldsymbol{x} \in \mathcal{X}$ indicates that $\boldsymbol{x}$ are global coordinates of an element of $\mathcal{X}$. Consider the relative twist coordinates of camera with respect to the object $\boldsymbol{u} = [\boldsymbol{u}_T^{\mathrm{T}} \boldsymbol{u}_R^{\mathrm{T}}]^{\mathrm{T}}$, where $\boldsymbol{u}_T = [u_1,\ u_2,\ u_3]^{\mathrm{T}}$ is the relative translational velocity, and $\boldsymbol{u}_R = [u_4,\ u_5,\ u_6]^{\mathrm{T}}$ is the relative angular velocity, by stacking the equation of rigid body kinematics: $\dot{\boldsymbol{x}}_i = -\boldsymbol{u}_T + \boldsymbol{x}_i^{\wedge}\, \boldsymbol{u}_R,\ i = 1, \ldots, n$, the following nonlinear, input-affine, and driftless smooth system is derived:

$$\dot{\boldsymbol{x}} = H(\boldsymbol{x})\boldsymbol{u} = \begin{bmatrix} -I_3 & \boldsymbol{x}_1^{\wedge} \\ \ldots & \\ -I_3 & \boldsymbol{x}_n^{\wedge} \end{bmatrix} \boldsymbol{u} \quad , \tag{1}$$

where $\boldsymbol{x} \in \mathcal{X}$ is the state vector, and $\boldsymbol{u} \in U$ is the control input vector. Assume a full perspective camera model with fixed focal length f and optical axis $\boldsymbol{k}_c$, the optical flow equations can be derived from system (1) by a change of state coordinates. Namely, define the diffeomorphism:

$$\begin{aligned} \overline{\boldsymbol{p}} &= \boldsymbol{\Phi}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{\Phi}_1(\boldsymbol{x}) \\ \ldots \\ \boldsymbol{\Phi}_n(\boldsymbol{x}) \end{bmatrix} \quad , \\ \boldsymbol{\Phi}_i(\boldsymbol{x}) &= \begin{bmatrix} p_{x_i} \\ p_{y_i} \\ p_{z_i} \end{bmatrix} = \begin{bmatrix} f\frac{x_i}{z_i} \\ f\frac{y_i}{z_i} \\ z_i \end{bmatrix}, \quad i = 1, \ldots, n \quad , \end{aligned} \tag{2}$$

in the region $\mathcal{X}_c = \mathcal{X} \setminus \{\boldsymbol{x} \in \mathcal{X} : z_i = 0, \quad i = 1, \ldots, n\}$, where the first two equations represent the full perspective camera transformation. The expression of the tangent map $\boldsymbol{\Phi}_{*\boldsymbol{x}}$ in the coordinates $\boldsymbol{x}$ results in the following block diagonal matrix:

$$\begin{aligned} \frac{\partial \boldsymbol{\Phi}(\boldsymbol{x})}{\partial \boldsymbol{x}} &= \begin{bmatrix} A(\boldsymbol{x}_1) & \ldots & 0_{3\times 3} \\ \ldots & & \\ 0_{3\times 3} & \ldots & A(\boldsymbol{x}_n) \end{bmatrix} \quad , \\ A(\boldsymbol{x}_i) &= \begin{bmatrix} \frac{f}{z_i} & 0 & -f\frac{x_i}{z_i^2} \\ 0 & \frac{f}{z_i} & -f\frac{y_i}{z_i^2} \\ 0 & 0 & 1 \end{bmatrix}, \quad i = 1, \ldots, n \quad , \end{aligned} \tag{3}$$

which is nonsingular in the region $\mathcal{X}_c$. The system in the new coordinates has equations:

$$\dot{\overline{\boldsymbol{p}}} == \begin{bmatrix} -A(\boldsymbol{x}_1) & A(\boldsymbol{x}_1)\boldsymbol{x}_1^{\wedge} \\ \ldots & \\ -A(\boldsymbol{x}_n) & A(\boldsymbol{x}_n)\boldsymbol{x}_n^{\wedge} \end{bmatrix}_{\boldsymbol{x}=\boldsymbol{\Phi}^{-1}(\overline{\boldsymbol{p}})} \boldsymbol{u} \quad . \tag{4}$$

Consider the vector of the image points $\boldsymbol{p} = [\boldsymbol{p}_1^{\mathrm{T}}, \ldots, \boldsymbol{p}_n^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{p}_i = [p_{x_i},\ p_{y_i}]^{\mathrm{T}}$, $i = 1, \ldots, n$, in the sequel it is assumed that $n > 3$ (i.e. redundant image features), and $\boldsymbol{p}$ defines global coordinates of the configuration space $\mathcal{X}_c$. Then the expression of the system (4), in the coordinates $\boldsymbol{p}$ results:

$$\dot{\boldsymbol{p}} = G(\boldsymbol{p})\,\boldsymbol{u} = \begin{bmatrix} G_1(\boldsymbol{p}_1) \\ \ldots \\ G_n(\boldsymbol{p}_n) \end{bmatrix} \boldsymbol{u} \quad , \ G_i(\boldsymbol{p}_i) = \begin{bmatrix} \boldsymbol{a}_{x_i}^{\mathrm{T}} & \boldsymbol{b}_{x_i}^{\mathrm{T}} \\ \boldsymbol{a}_{y_i}^{\mathrm{T}} & \boldsymbol{b}_{y_i}^{\mathrm{T}} \end{bmatrix} \quad , \tag{5}$$

where the following vectors have been defined:

$$\begin{aligned}
\boldsymbol{a}_{x_i}^{\mathrm{T}} &= \left[-\frac{f}{z_i},\ 0,\ \frac{p_{x_i}}{z_i}\right] \\
\boldsymbol{b}_{x_i}^{\mathrm{T}} &= \left[\frac{p_{x_i}p_{y_i}}{f},\ -f-\frac{p_{x_i}^2}{f},\ p_{y_i}\right] \\
\boldsymbol{a}_{y_i}^{\mathrm{T}} &= \left[0,\ -\frac{f}{z_i},\ \frac{p_{y_i}}{z_i}\right] \\
\boldsymbol{b}_{y_i}^{\mathrm{T}} &= \left[f+\frac{p_{y_i}^2}{f},\ \frac{p_{x_i}p_{y_i}}{f},\ -p_{x_i}\right],\ i=1,\ldots,n \quad . \qquad (6)
\end{aligned}$$

The system is defined on the smooth manifold $\mathcal{X}_c$, $\boldsymbol{p} \in \mathcal{X}_c$ is the state vector, $\boldsymbol{u} \in U$ is the control input vector, and the depths of object's points $z_i,\ i = 1,\ldots,n$ are considered time-varying parameters of the system. The following assumption will be used in the next sections, and concerns the rank of the interaction matrix $G(\boldsymbol{p})$.

Assumption 1 *The matrix $G(\boldsymbol{p})$ is full-rank (i.e. $\mathrm{rank}(G(\boldsymbol{p})) = 6$), $\forall \boldsymbol{p} \in \mathcal{X}_c$, and $\forall z_i \in \Re \setminus \{0\},\ i = 1,\ldots,n$.*

2.2 Nonlinear Controllability

Controllability properties of the image-based visual system in Eq. (5) are established, by applying nonlinear control theory (Isidori, 1989; Nijmeijer and van der Shaft, 1996).

Proposition 1 *The system in Eq. (5) satisfies the accessibility rank condition, $\forall \boldsymbol{p} \in \mathcal{X}_c$.*

Proof:
See Appendix A.

Proposition 2 *The system in Eq. (5) is completely controllable.*

Proof:
See Appendix A.

Given a reference setpoint $\boldsymbol{p}^{(\mathrm{d})} = \left[\boldsymbol{p}_1^{(\mathrm{d})\mathrm{T}},\ldots,\boldsymbol{p}_n^{(\mathrm{d})\mathrm{T}}\right]^{\mathrm{T}} \in \mathcal{X}_c,\ \boldsymbol{p}_i^{(\mathrm{d})} = \left[p_{x_i}^{(\mathrm{d})},\ p_{y_i}^{(\mathrm{d})}\right]^{\mathrm{T}},\ i = 1,\ldots,n$, define the error vector $\tilde{\boldsymbol{p}} = [\tilde{\boldsymbol{p}}_1^{\mathrm{T}},\ldots,\tilde{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}}$, where $\tilde{\boldsymbol{p}}_i = \boldsymbol{p}_i - \boldsymbol{p}_i^{(\mathrm{d})}\ i = 1,\ldots,n$.

The following result will be useful in the next section.

Lemma 1 *Define $\tilde{\boldsymbol{q}} = G(\boldsymbol{p})^T \tilde{\boldsymbol{p}} \in \Re^6$, then $\tilde{\boldsymbol{q}} = \boldsymbol{0}$ if and only if $\tilde{\boldsymbol{p}} = \boldsymbol{0}$.*

Proof:
See Appendix A.

3 Adaptive Nonlinear Visual Servoing

In this section, the problem of finding a stabilizing feedback control of the image-based visual system is addressed using Lyapunov-based design. The feedback stabilization of the equilibrium $\tilde{\boldsymbol{p}} = [\tilde{\boldsymbol{p}}_1^{\mathrm{T}}, \ldots, \tilde{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}} = \mathbf{0}$ is a difficult task due to the presence of unknown depths $z_i,\ i = 1, \ldots, n$ of the object's points. The sampling time of the visual servoing system is not negligible, since, at the actual state of technology, typical values of control time-cycle are from $40\,ms$ to $80\,ms$. So that a discrete-time Lyapunov-based design of the visual servo system is derived. In order to develop a discrete-time state space model of the image-based system, let us approximate the optical flow by finite-differences equations (Papanikolopoulos and Khosla, 1993):

$$\begin{aligned} \dot{p}_{x_i}(kT) &= \frac{p_{x_i}((k+1)T) - p_{x_i}(kT)}{T} \\ \dot{p}_{y_i}(kT) &= \frac{p_{y_i}((k+1)T) - p_{y_i}(kT)}{T} \quad , \end{aligned} \tag{7}$$

where $i = 1, \ldots, n$, and T is the visual sampling time. The discrete time visual model, hence, results:

$$\boldsymbol{p}((k+1)T) = \boldsymbol{p}(kT) + G_d(\boldsymbol{p}(kT))\,\boldsymbol{u} \quad , \tag{8}$$

where $G_d(\boldsymbol{p}(kT)) = TG(\boldsymbol{p}(kT))$. In the sequel, to simplify notation, it has been used k instead of kT. To stabilize the system (8), the following discrete-time control law has been designed:

$$\boldsymbol{u}(k) = -\sigma \hat{G}_d(\boldsymbol{p}(k))^{+} \tilde{\boldsymbol{p}}(k) \quad ,\ \sigma \in \Re \quad . \tag{9}$$

where $\hat{G}_d(\boldsymbol{p})^{+} = (\hat{G}_d(\boldsymbol{p})^{\mathrm{T}}\, \hat{G}_d(\boldsymbol{p}))^{-1} \hat{G}_d(\boldsymbol{p})^{\mathrm{T}}$ is the Moore-Penrose pseudo-inverse of $\hat{G}_d(\boldsymbol{p})$. The following adaptation law is used:

$$\hat{\theta}_i(k+1) = \hat{\theta}_i(k) - \gamma (C_i(k)\boldsymbol{u}_T(k))^{\mathrm{T}} \boldsymbol{e}_{p_i}(k),\ \gamma > 0 \quad , \tag{10}$$

$\hat{\theta}_i(k),\ i = 1, \ldots, n$ is the estimate of the inverse of uncertain depth $\theta_i(k) = \frac{1}{z_i(k)}$, $\hat{G}_d(\boldsymbol{p}(k))$ is obtained by substituting the estimates in the expression of $G_d(\boldsymbol{p}(k))$, and

$$C_i(k) = T \begin{bmatrix} -f & 0 & p_{x_i}(k) \\ 0 & -f & p_{y_i}(k) \end{bmatrix} \quad . \tag{11}$$

The prediction error $\boldsymbol{e}_{p_i}(k) = \hat{\boldsymbol{d}}_i(k) - \boldsymbol{d}_i(k),\ i = 1, \ldots, n$ uses the following linear parametrization form:

$$\boldsymbol{d}_i(k) = \boldsymbol{p}_i(k+1) - \boldsymbol{p}_i(k) - T \begin{bmatrix} \boldsymbol{b}_{x_i}^{\mathrm{T}} \\ \boldsymbol{b}_{y_i}^{\mathrm{T}} \end{bmatrix} \boldsymbol{u}_R = (C_i(k)\boldsymbol{u}_T(k))\theta_i, \quad i = 1, \ldots, n \quad , \tag{12}$$

and $\hat{\boldsymbol{d}}_i(k)$ is obtained by substituting the estimates $\hat{\theta}_i(k),\ i = 1, \ldots, n$ in the expression of $\boldsymbol{d}_i(k)$.

The following proposition formally establishs the properties of the proposed Lyapunov-based control system design.

If it is assumed to know the depths $z_i,\ i = 1, \ldots, n$ of the object's points, which is true in the case of robot operating in structured environments, the following result holds.

Proposition 3 *If assumption 1 is satisfied, and the depths of the object's points are known (i.e. $\hat{\theta}_i(k) = \theta_i(k),\ i = 1, \ldots, n$), the equilibrium $\tilde{\boldsymbol{p}} = \mathbf{0}$ of the system (8) is globally uniformly asymptotically stable, provided that the control law in Eq. (9) is used with the control gain σ chosen in the open interval $(0,\ 2)$.*

Proof:
See Appendix A.

In the case of unstructured environment, the depths $z_i,\ i = 1, \ldots, n$ are unknown time-varying parameters of the visual system. The ultimate boundedness of the error vector $\tilde{\boldsymbol{p}}$ can be obtained whenever the matrix $G_d \hat{G}_d^+$ remains positive definite along the current trajectory. This property of the image-based visual servoing approach has been pointed out also in (Espiau *et al.*, 1992), where constant 3-D parameters are used to estimate the interaction matrix.

4 Experimental Results

The proposed adaptive nonlinear control design has been implemented on an eye-in-hand robotic system. The phisical system, on which the approach was tested, consists of a PUMA 560 robot arm with a Sony CCD camera mounted on its wrist. Camera optics data-sheets provide a raw value for focal length and pixel dimensions; the remaining intrinsic parameters of the camera are not considered. The robot is commanded by the MARK III controller (implementing the inner loop), and—for the outer loop—a PC MMX 200 Mhz equipped with an Imaging Technology frame grabber. The MARK III controller operates under VAL II programs and communicates with the PC through the ALTER real time protocol using an RS-232 serial interface. All the acquisition and control activities running in the PC are executed under the HARTIK kernel (Buttazzo, 1993), which is specifically designed to support real-time control applications with timing constraints. The control system was implemented as a multitasking application. A task τ_{act} with period $T_{\text{act}} = 28\,ms$, reads commands (i.e. the control velocity screw) from a CAB (Cyclic Asynchronous Buffer), and sends them to the robot controller via ALTER. The control task τ_{ctr} with a period of $T_{\text{ctr}} = 80\,ms$ performs the visual analysis (i.e. active contour tracking), computes the velocity screw according to Eqs. (9), (10), and sends the commands to τ_{act}. Finally, a task τ_{vis} displays the status of the system. A hardware-dependent communication level ensures a correct timing table with the physical devices. Quadratic B-spline active contours are used to track the projection of Γ in the image plane (Colombo and Allotta, 1999). The image points $\boldsymbol{p}$ are computed using directly the 2-D time evolution of the B-spline control points. Any experiment with the system begins by bringing the robot in the desired configuration; there an active contour is initialized with the goal image appearance of the target object. After recording the n control points of the goal contour, the robot is moved to the initial configuration, while the current active contour continues tracking the object appearance. Finally, the robot is moved according to the designed control system.

$n = 12$ control points of the active contour are used as image points $\boldsymbol{p}$, in order to obtain robust visual acquisition of the object. The values of estimated depth parameters are initialized to $\hat{z}_i = \frac{1}{\hat{\theta}_i} = 200\,mm,\ i = 1, \ldots, 12$, in all the following experiments, thus providing a coarse initial estimation with an error up to 50% with respect to the real values. Infinite impulse response digital filters were used for smoothing visual measurements. The following cases are addressed:

1. **2-D displacement**: the initial and desired pose of the camera are in planes parallel to the object's plane;

2. **3-D generic displacement**: the initial and desired pose of the camera are 3-D generic configurations.

In the first experiment (2-D displacement), the eye-in-hand robotic system executes a relative positioning task with the following values of state variables. The initial state of system of Eq. (5) and the desired state are: $\boldsymbol{p}(0) = [-1.90\ 2.22\ -2.07\ 1.35\ -2.14\ 0.64\ -2.14\ -0.17\ -1.50\ -0.27\ -0.98\ -0.33 - 0.37\ -0.43\ -0.25\ 0.30\ -0.16\ 1.08\ -0.08\ 1.92\ -0.74\ 2.06\ -1.36\ 2.19]^{\mathrm{T}}$, and $\boldsymbol{p}^{(d)} = [0.30\ 3.15\ 0.17\ 2.28\ 0.09\ 1.57\ 0.08\ 0.74\ 0.72\ 0.62\ 1.24\ 0.561.86\ 0.44\ 1.99\ 1.18\ 2.09\ 1.97\ 2.19\ 2.81\ 1.52\ 2.96\ 0.90\ 3.11]^{\mathrm{T}}$.

Table 1: Desired and reached configurations (encoder measurements) of the robot end-effector : experiments 1, 2. The quantities "o," "a" and "t" are the Orientation, Altitude and Tool angles (degrees), respectively, used by VAL II to represent the end-effector orientation. End-effector translations "x," "y" and "z" are in mm.

	x	y	z	o	a	t
goal(1)	70.0	751.6	-115.5	-89.9	89.9	0.0
reached(1)	69.7	739.3	-115.0	175	88.5	-94.9
goal(2)	-9.5	701.5	-287.5	89.4	61.2	179
reached(2)	-3.7	706.2	-283.1	87.	60.1	177.6

The image points errors and the estimated depths are plotted in Figs. 1.a and 1.b. Figs. 1.c and 1.d show the components of the requested camera velocity twist. Table 1 reports the errors in terms of reached end-effector's pose. In the second experiment (3-D generic displacement), the initial state is $\boldsymbol{p}(0) = [0.99\ 1.26\ 0.76\ 0.46\ 0.57\ -0.33\ 0.44\ -0.92\ 1.08\ -1.12\ 1.63\ -1.32\ 2.01\ -1.41\ 2.18\ -0.86\ 2.38\ -0.07\ 2.53\ 0.55\ 2.12\ 0.76\ 1.53\ 1.02]^{\mathrm{T}}$, and the desired state is $\boldsymbol{p}^{(d)} = [-2.66\ 2.70\ -2.80\ 1.34\ -2.90\ -0.00\ -2.95\ -0.99\ -1.90\ -1.05\ -0.97\ -1.15\ -0.36\ -1.12\ -0.25\ 0.18\ -0.14\ 1.13\ -0.06\ 2.20\ -0.75\ 2.37\ -1.75\ 2.54]^{\mathrm{T}}$. Figs. 2.a and 2.b show the resulting image points errors and the estimated depths, while the components of the requested camera velocity twist are plotted in Figs. 2.c and 2.d. Table 1 shows the errors in terms of reached end-effector's pose.

Experiment 1 concerns the case of 2-D displacement of the camera, it validates the proposed adaptive nonlinear control approach, ensuring convergence of image points errors towards the equilibrium $\tilde{\boldsymbol{p}} = \mathbf{0}$, and providing the estimates of object points's depth. After $T_c = 100\,s$, the maximum magnitude of the position error in terms of reached end-effector's pose is 12.29 mm (see Table 1). Experiment 2 concerns a 3-D generic camera displacement: in this case there is also a large mismatch between initial and desired configurations. The results show a satisfactory accurancy in the positioning of the camera, since, after $T_c = 100\,s$, the maxiumum magnitude of the position error is 5.89 mm (see Table 1).

5 Conclusions

The problem of camera-object relative positioning is addressed by nonlinear controllability analysis and adaptive nonlinear control system design in discrete time. The image-based visual system is expressed in terms of global coordinates fully defined in the image plane. Then a discrete time model is derived, since the visual sampling time is not negligible. In case of known depth of the object, the designed control law ensures global uniform asymptotic stability of the image plane reference setpoint. 3-D estimation procedure based on prediction errors is used to cope with the unknown depth of the object. Robotic experiments show the effectiveness of the approach, and the satisfactory accuracy in the positioning of the camera.

A Appendix

Proof of Prop. (1):

Proof:
From assumption 1, and the definition of the accessibility algebra, it follows that $\dim C(\boldsymbol{p}) \geq \text{rank}G(\boldsymbol{p}) = dim\mathcal{X}_c \ \forall \boldsymbol{p} \in \mathcal{X}_c$, hence $\dim C(\boldsymbol{p}) = \dim \mathcal{X}_c, \ \forall \boldsymbol{p} \in \mathcal{X}_c$.

Proof of Prop. (2):

Proof:
Let $\mathcal{G} = \{G(\boldsymbol{p})\boldsymbol{u} : \boldsymbol{u} \in U\}$ be the family of the vector fields on $\mathcal{X}_c$. Since $\mathcal{G}$ is symmetric (i.e. $G(\boldsymbol{p})\boldsymbol{u} \in \mathcal{G}$ implies $-G(\boldsymbol{p})\boldsymbol{u} \in \mathcal{G}$), and $\mathcal{X}_c$ is connected, from Prop. (1), it follows the thesis (Nijmeijer and van der Shaft, 1996).

Proof of Lemma (1):

Proof:
The necessity is obvious. Assume that $G(\boldsymbol{p})^{\mathrm{T}}\,\tilde{\boldsymbol{p}} = \boldsymbol{0}$, taking as global coordinates of the system (5) the error vector $\tilde{\boldsymbol{p}}$, yields $G(\tilde{\boldsymbol{p}} + \boldsymbol{p}^{(\mathrm{d})})^{\mathrm{T}}\,\tilde{\boldsymbol{p}} = \overline{G}(\tilde{\boldsymbol{p}})^{\mathrm{T}}\,\tilde{\boldsymbol{p}} = \boldsymbol{0}$, where $\overline{G}(\tilde{\boldsymbol{p}}) = [\overline{m}_1, \ldots, \overline{m}_6]$ is the expression of $G(\boldsymbol{p})$ in the error coordinates. Consider the distribution spanned by the vector fields $\overline{m}_i, \ i = 1, \ldots, 6$, $\Delta(\tilde{\boldsymbol{p}}) = \mathrm{Im}\overline{G}(\tilde{\boldsymbol{p}}) = \mathrm{span}\{\overline{m}_1, \ldots, \overline{m}_6\}$. Given $\tilde{\boldsymbol{p}} \neq \boldsymbol{0} \in \mathcal{X}_c$, since from linear algebra, $\mathrm{Ker}\overline{G}(\tilde{\boldsymbol{p}})^{\mathrm{T}} \oplus \Delta(\tilde{\boldsymbol{p}}) = \Re^{2\mathrm{n}}$, it is sufficient to show that $\tilde{\boldsymbol{p}} \in \Delta(\tilde{\boldsymbol{p}})$. Infact $\tilde{\boldsymbol{p}} \in \mathcal{X}_c \cong T_{\tilde{\boldsymbol{p}}}\mathcal{X}_c$, from Prop (1), it results $C(\tilde{\boldsymbol{p}}) = T_{\tilde{\boldsymbol{p}}}\mathcal{X}$ (Chow's theorem). Finally, from assumption 1, a dimensional argumentation shows that $C(\tilde{\boldsymbol{p}}) = \Delta(\tilde{\boldsymbol{p}})$, hence $\tilde{\boldsymbol{p}} \in \Delta(\tilde{\boldsymbol{p}})$, which concludes the proof.

Proof of Prop. (3):

Proof:
Consider the following Lyapunov function candidate:

$$V(\tilde{\boldsymbol{p}}(k)) = \tilde{\boldsymbol{p}}(k)^{\mathrm{T}}\tilde{\boldsymbol{p}}(k) \quad , \tag{13}$$

which is time-invariant, positive definite and radially unbounded. The first order variation of $V(\tilde{\boldsymbol{p}})$, using Eq. (9), results:

$$\Delta V(\tilde{\boldsymbol{p}}(k)) = \tilde{\boldsymbol{p}}(k+1)^{\mathrm{T}}\tilde{\boldsymbol{p}}(k+1) - \tilde{\boldsymbol{p}}(k)^{\mathrm{T}}\tilde{\boldsymbol{p}}(k) = (-2\sigma + \sigma^2)\tilde{\boldsymbol{p}}(k))^{\mathrm{T}}G_d(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))^{+}\tilde{\boldsymbol{p}}(k) \quad ,(14)$$

where it has been used the fact that

$$\begin{aligned} &G_d(\boldsymbol{p}(k))^{+\,\mathrm{T}}G_d(\boldsymbol{p}(k))^{\mathrm{T}}G_d(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))^{+} = \\ &\quad G_d(\boldsymbol{p}(k))(G_d(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))^{\mathrm{T}})^{-\mathrm{T}}G_d(\boldsymbol{p}(k))^{\mathrm{T}}G_d(\boldsymbol{p}(k)) \\ &\quad (G_d(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))^{\mathrm{T}})^{-1}G_d(\boldsymbol{p}(k))^{\mathrm{T}} = \\ &\quad G_d(\boldsymbol{p}(k)(G_d(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))^{\mathrm{T}})^{-\mathrm{T}}G_d(\boldsymbol{p}(k))^{\mathrm{T}} \quad . \end{aligned} \tag{15}$$

By assumption 1, and lemma (1), $\dot{V}(\tilde{\boldsymbol{p}})$ is negative definite, if the control gain σ is chosen in the open interval $(0,\ 2)$.

References

Allotta, B., F. Conticelli, and C. Colombo (1998). "Robust decoupling control of an image-based visual system," in *4th IFAC Nonlinear Contr. Syst. Des. Symp., Enschede, The Netherlands*.

Buttazzo, G. (1993). "Hartik: a real-time kernel for robotics applications," in *IEEE Proc. of Real-Time System Symposium, Raleigh-Durham, Dec. 1993*.

Colombo, C. and B. Allotta (1999). "Image-based robot task planning and control using a compact visual representation," *IEEE Trans. on Systems, Man and Cybernetics*, **29**, no. 1, pp. 92–100.

Espiau, B., F. Chaumette, and P. Rives (1992). "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, **8**, pp. 313–326.

Hashimoto, K., T. Ebine, and H. Kimura (1996). "Visual servoing with hand-eye manipulator-optimal control approach," *IEEE Trans. Robot. Autom.*, **12**, no. 5, pp. 766–774.

Hill, J. and W. T. Park (1979). "Real time control of a robot with a mobile camera," in *Proc. 9th ISIR, Washington, D.C.*, pp. 233–246.

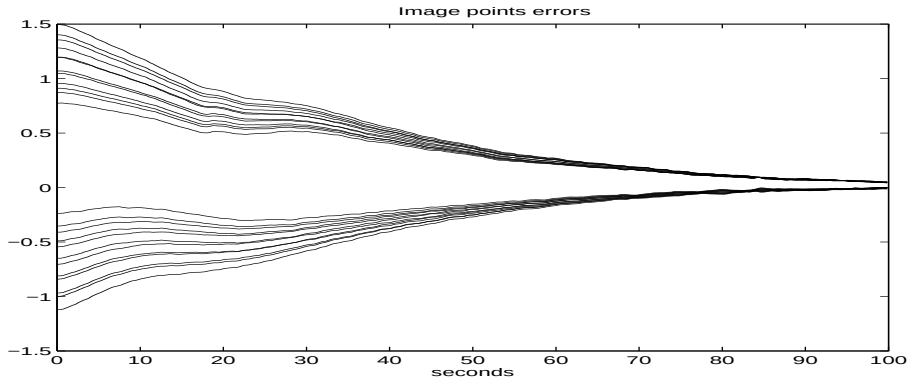
Isidori, A. (1989). *Nonlinear Control Systems: An Introduction*, Springer Verlag.

Murray, R. M., Z. Li, and S. Sastry (1994). *A mathematical introduction to robotic manipulation*, CRC Press, Inc.

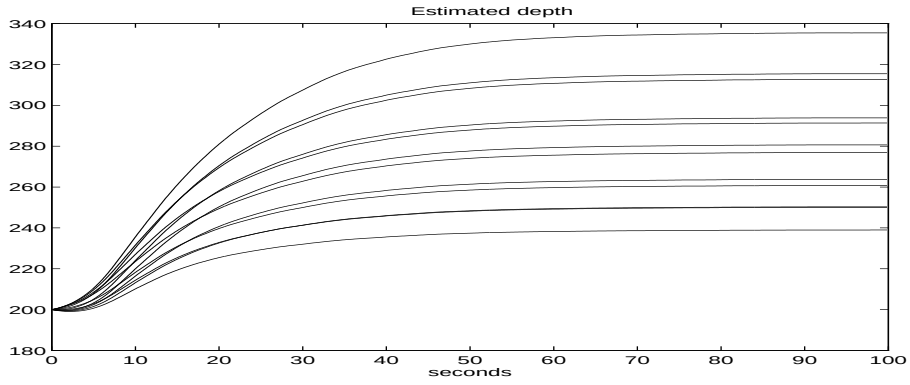
Nijmeijer, H. and A. van der Shaft (1996). *Nonlinear Dynamical Control Systems*, Springer-Verlag, 3rd edn.

Papanikolopoulos, N. P. and P. K. Khosla (1993). "Adaptive robot visual tracking: theory and experiments," *IEEE Trans. Autom. Contr.*, **38**, no. 3, pp. 429–445.

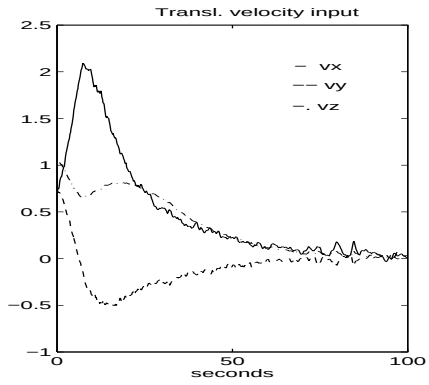
Slotine, J. J. and W. Li (1993). *Applied Nonlinear Control*, Prentice Hall, Englewood Cliffs, New Jersey.



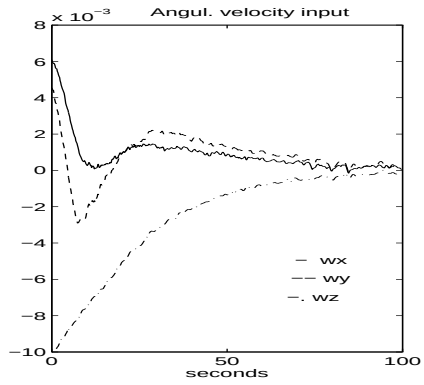
a)



b)



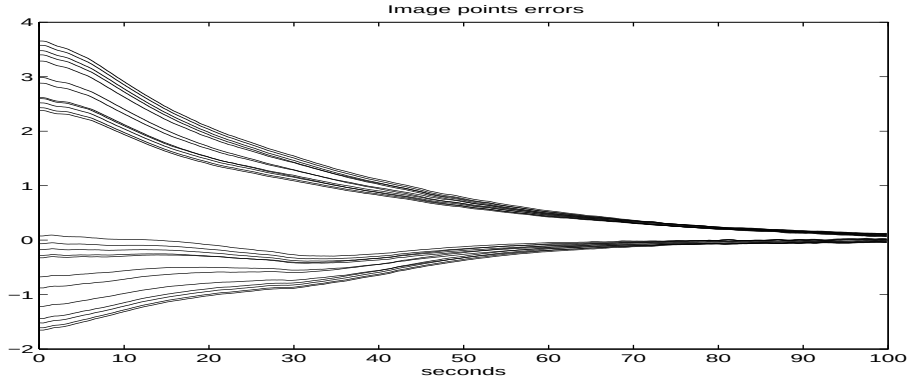
c)



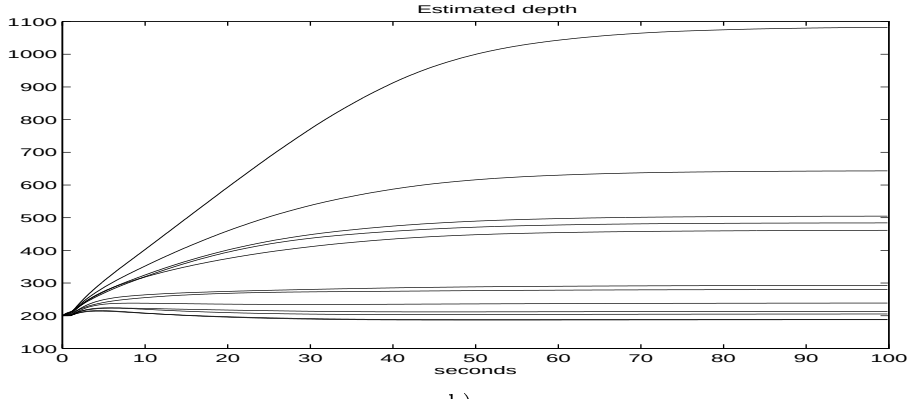
d)

Figure 1: First experiment ($n = 12$ control points of spline used): a) image points errors; b) estimated depths of object's points; c) camera velocity components; d) camera angular velocity components.

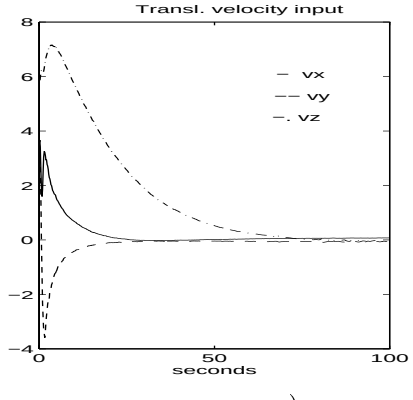
Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999



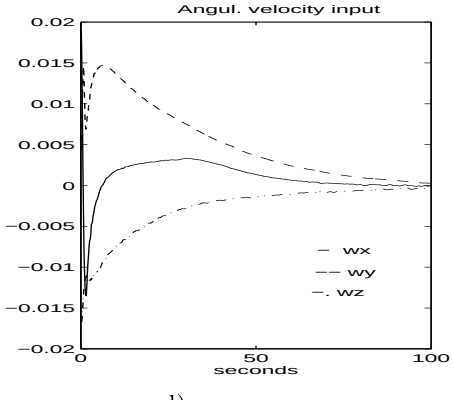
a)



b)



c)



d)

Figure 2: Second experiment ($n = 12$ control points of spline used): a) image points errors; b) estimated depths of object's points; c) camera velocity components; d) camera angular velocity components.

2339