Discrete-Time Robot Visual Feedback Systems with Nonlinear Depth Adaptation: Stability Analysis and Experiments

Fabio Conticelli Benedetto Allotta

Scuola Superiore Sant'Anna
Via Carducci 40, I-56127 Pisa, Italy
E-mail {contice, ben}@sssup.it

Abstract

In this paper, a nonlinear adaptive visual feedback scheme is designed to perform 3-D positioning tasks, consisting in the regulation of the relative pose between a robotic camera and a rigid object of interest. The dynamical system of robotic camera-object interaction is expressed in terms of image features, i.e. 2-D points track-able in the image plane. Since the visual sampling period is not negligible at the current state of technology, a discrete-time representation of the camera-object interacton model is first derived. By exploiting nonlinear controllability properties, a Lyapunov-based discrete-time control law is designed to ensure asymptotic stability in the large of the image reference set-point. Moreover, a 3-D adaptation procedure, in case of unknown object depth, ensures ultimate boundedness of the whole state vector.

1 Introduction

Visual servoing systems use camera sensors inside the control loop to accomplish tasks in unstructured environments [5]. This approach can be used to improve, using non-contact measurements, the adaptability of robotic systems with respect to both environmental uncertainties in positioning tasks, and inaccuracy of the robot's kinematic model. Concerning the control aspects, two main paradigms can be outlined: position-based and image-based servoing [3, 6, 1]. 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 this paper, we propose a novel discrete-time adaptive visual control scheme, to regulate the relative pose between a robotic camera and a rigid object of interest. The configuration space in which the system is defined is the Special Euclidean group, and the dynamic visual system is expressed in terms of image features, i.e. 2-D points trackable in the image plane. A discrete-time representation of the camera-object interaction model is derived, since the visual sampling time is not negligible at the current state of technology. In structured environments, where the system a priori knows the object depth, no adaptation law is needed, and the designed discrete-time control system ensures uniform asymptotic stability of the image reference set-point. In unstructured environments the depth 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 depth parameters. The proposed nonlinear 3-D adaptation procedure, based on Lyapunov-based design and prediction error, ensures ultimate boundedness of the whole state vector. 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 re-called. In Sect. 3, the adaptive discrete-time nonlinear control system is designed and stability analysis is carried out. Sect. 4 reports experiment results carried out to validate the theoretical framework. Finally, in Sect. 5 the major contribution of the paper is summarized.

2 Dynamic Robotic Camera-Object Visual Interaction

In this section, the dynamic visual model of camera-object interaction is derived in state space form, and nonlinear controllability properties are analyzed. Assume that the object of interest is rigid, so the the relative motion between the camera and the object is a rigid body transformation [9], the configuration space

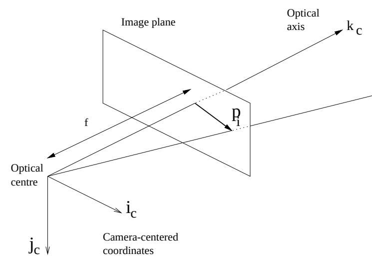

Figure 1: Definition of full perspective camera model. The perspective projection of $\boldsymbol{x}_i = [x_i\ y_i\ z_i]^{\mathrm{T}}$ is $\boldsymbol{p}_i = [p_{x_i}\ p_{y_i}]^{\mathrm{T}}$, defined, formally, by the diffeomorphism $\boldsymbol{\Phi}_i = [p_{x_i}\ p_{y_i}\ p_{z_i}]^{\mathrm{T}} = [f\frac{x_i}{z_i}\ f\frac{y_i}{z_i}\ z_i]^{\mathrm{T}},\ i = 1, \ldots, n$.

of the camera-object relative pose 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 $SE(3)$, i.e at each $\boldsymbol{x} = [\boldsymbol{x}_1^{\mathrm{T}}\ \ldots\ \boldsymbol{x}_n^{\mathrm{T}}]^{\mathrm{T}}$, such that $\boldsymbol{x}_i \in \Gamma,\ i = 1, \ldots, n$, corresponds only one valid configuration $g(\boldsymbol{x}) \in SE(3)$, and $\forall g \in SE(3)$, there exists only one $\boldsymbol{x}$ with which it can be identified. In the following, to simplify notation, the statement $\boldsymbol{x} \in SE(3)$ indicates that $\boldsymbol{x}$ are global coordinates of an element of $SE(3)$. Consider the relative twist coordinates of the 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. Given 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), so that $\boldsymbol{p}$ defines coordinates of the configuration space $\mathcal{X}$. Then the expression of the camera-object visual model in the coordinates $\boldsymbol{p}$ results [2]

$$\dot{\boldsymbol{p}} = G(\boldsymbol{p})\,\boldsymbol{u} = \begin{bmatrix} G_1(\boldsymbol{p}_1) \\ \cdots \\ G_n(\boldsymbol{p}_n) \end{bmatrix} \boldsymbol{u},\ G_i(\boldsymbol{p}_i) = \begin{bmatrix} \boldsymbol{a}_{x_i}^{\mathrm{T}} & \boldsymbol{b}_{x_i}^{\mathrm{T}} \\ \boldsymbol{a}_{y_i}^{\mathrm{T}} & \boldsymbol{b}_{y_i}^{\mathrm{T}} \end{bmatrix} \tag{1}$$

where the following vectors have been defined

$$\begin{aligned} \boldsymbol{a}_{x_i}^{\mathrm{T}} &= \begin{bmatrix} -\frac{f}{z_i} & 0 & \frac{p_{x_i}}{z_i} \end{bmatrix} \\ \boldsymbol{b}_{x_i}^{\mathrm{T}} &= \begin{bmatrix} \frac{p_{x_i} p_{y_i}}{f} & -f - \frac{p_{x_i}^2}{f} & p_{y_i} \end{bmatrix} \\ \boldsymbol{a}_{y_i}^{\mathrm{T}} &= \begin{bmatrix} 0 & -\frac{f}{z_i} & \frac{p_{y_i}}{z_i} \end{bmatrix} \\ \boldsymbol{b}_{y_i}^{\mathrm{T}} &= \begin{bmatrix} f + \frac{p_{y_i}^2}{f} & \frac{p_{x_i} p_{y_i}}{f} & -p_{x_i} \end{bmatrix},\ i = 1, \ldots, n \end{aligned} \tag{2}$$

The matrix $G(\boldsymbol{p})$ is a particular type of interaction matrix (or image Jacobian) [6]. The system is defined on the smooth manifold $\mathcal{X}$, $\boldsymbol{p} \in \mathcal{X}$ is the state vector, $\boldsymbol{u} \in U$ is the control input vector, and the depth values $z_i,\ i = 1, \ldots, n$ are considered as 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.* $\mathrm{rank}(G(\boldsymbol{p})) = 6$*), $\forall \boldsymbol{p} \in \mathcal{X}$, and $\forall z_i \in \Re \setminus \{0\}$, $i = 1, \ldots, n$.*

Notice that the above assumption is not restrictive, since the interaction matrix $G(\boldsymbol{p})$ has $2n$ rows with $n > 3$, i.e. the case of redundant image points is considered. The matrix $G(\boldsymbol{p})$ loses rank only in really particular configurations, for example, when all the n points are collinear [4].

Controllability properties of the image-based visual system in Eq. (1) are re-called, they result from the application of nonlinear control theory [7].

Given a reference set-point $\boldsymbol{p}^{(\mathrm{d})} = \left[\boldsymbol{p}_1^{(\mathrm{d})\mathrm{T}}\ \ldots\ \boldsymbol{p}_n^{(\mathrm{d})\mathrm{T}}\right]^{\mathrm{T}} \in \mathcal{X},\ \boldsymbol{p}_i^{(\mathrm{d})} = \left[p_{x_i}^{(\mathrm{d})}\ p_{y_i}^{(\mathrm{d})}\right]^{\mathrm{T}},\ i = 1, \ldots, n$, 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 lemmas will be useful in the next section.

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

Denote with $||\cdot||$ the usual Euclidean norm for a vector, and the induced spectral norm for a matrix.

Lemma 2 *The following inequality holds: $||\tilde{\boldsymbol{q}}||^2 \geq \lambda_m\, ||\tilde{\boldsymbol{p}}||^2$, where $\sqrt{\lambda_m},\ \lambda_m > 0$ is the minimum singular value of $G(\boldsymbol{p})$.*

For the proofs of the above results see [2].

3 Adaptive Discrete-Time Visual Feedback

In this section, the problem of finding a stabilizing feedback control law of the dynamic visual system is addressed using Lyapunov-based design.

3.1 Asymptotic Stabilization

The sampling time of visual servoing systems is not negligible at the actual state of technology, since typical values of control time-cycle are from 40 ms to 80 ms. As consequence, a discrete-time Lyapunov-based design of the visual servo system is derived here. In order to develop a discrete-time state space model of the visual system, let us approximate the optical flow by forward Euler finite-differences, as in [10]:

$$\begin{aligned} \dot{p}_{x_i}(kT) &\approx \frac{p_{x_i}((k+1)T) - p_{x_i}(kT)}{T} \\ \dot{p}_{y_i}(kT) &\approx \frac{p_{y_i}((k+1)T) - p_{y_i}(kT)}{T} \quad , \end{aligned} \qquad (3)$$

where $i = 1, \ldots, n$, and T is the visual sampling time. This approximation assumes that the image features do not change during the sampling interval, i.e. $\boldsymbol{p}(t) \approx \boldsymbol{p}(kT),\ z_i(t) \approx z_i(kT),\ t \in [kT,\ (k+1)T),\ k \geq 0,\ i = 1, \ldots, n$. In other words, it is required that the camera-object relative twist is "small" in comparison with the sampling rate.

The discrete-time visual model, from Eq. (1), results:

$$\boldsymbol{p}((k+1)T) = \boldsymbol{p}(kT) + G_d(\boldsymbol{p}(kT))\, \boldsymbol{u}(kT) \quad , \qquad (4)$$

where $G_d(\boldsymbol{p}(kT)) = T\, G(\boldsymbol{p}(kT))$. In the sequel, to simplify notation, k will be used instead of kT, and Δ denotes the forward difference operator, i.e. for any discrete scalar quantity $x(k) \in \Re$, $\Delta x(k) = x(k+1) - x(k)$.

If it is assumed to know the depth parameters $z_i,\ i = 1, \ldots, n$, which is true in the case of a robot operating in structured environments, in order to stabilize the system (4), the following discrete-time control law has been designed

$$\boldsymbol{u}(k) = -\sigma G_d^{+}(\boldsymbol{p}(k))\tilde{\boldsymbol{p}}(k) \quad , \sigma \in \Re \quad . \qquad (5)$$

where $G_d^{+}(\boldsymbol{p}) = (G_d^{\mathrm{T}}(\boldsymbol{p})\, G_d(\boldsymbol{p}))^{-1} G_d^{\mathrm{T}}(\boldsymbol{p})$ is the Moore-Penrose pseudo-inverse of $G_d(\boldsymbol{p})$. The stabilizing properties of the above control design are formally summarized in the following proposition.

Proposition 1 *If assumption 1 is satisfied, and the depth parameters are known, the equilibrium* $\tilde{\boldsymbol{p}} = \boldsymbol{0}$ *of the system (4) is uniformly asymptotically stable, provided that the control law in Eq. (5) is used with the control gain* σ *chosen in the open interval* $(0,\ 2)$.

Proof: Consider the following Lyapunov function candidate:

$$V(\tilde{\boldsymbol{p}}) = \tilde{\boldsymbol{p}}^{\mathrm{T}}(k)\tilde{\boldsymbol{p}}(k) \quad , \qquad (6)$$

which is time-invariant, positive definite and radially unbounded. The first order variation of $V(\tilde{\boldsymbol{p}})$ along the system dynamics (4), using Eq. (5), results:

$$\begin{aligned} \Delta V(\tilde{\boldsymbol{p}}) &= \tilde{\boldsymbol{p}}(k+1)^{\mathrm{T}}\tilde{\boldsymbol{p}}(k+1) - \tilde{\boldsymbol{p}}^{\mathrm{T}}(k)\,\tilde{\boldsymbol{p}}(k) \\ &= (-2\sigma + \sigma^2)\tilde{\boldsymbol{p}}^{\mathrm{T}}(k) G_d(\boldsymbol{p}(k)) G_d^{+}(\boldsymbol{p}(k))\tilde{\boldsymbol{p}}(k) \end{aligned} \qquad (7)$$

where it has been used the fact that, at the step k, the following equations hold:

$$\begin{aligned} G_d^{+\,\mathrm{T}} G_d^{\mathrm{T}} G_d G_d^{+} &= G_d (G_d^{\mathrm{T}}\, G_d)^{-\mathrm{T}}\, G_d^{\mathrm{T}}\, G_d (G_d^{\mathrm{T}} G_d)^{-1}\, G_d^{\mathrm{T}} \\ &= G_d (G_d^{\mathrm{T}}\, G_d)^{-\mathrm{T}}\, G_d^{\mathrm{T}} = G_d\, G_d^{+} \quad . \end{aligned} \qquad (8)$$

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

3.2 Nonlinear Depth Adaptation

In the case of an unstructured environment, the feedback stabilization of the equilibrium $\tilde{\boldsymbol{p}} = [\tilde{\boldsymbol{p}}_1^{\mathrm{T}}\ \ldots\ \tilde{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}} = \boldsymbol{0}$ is a difficult task due to the presence of unknown depth parameters $z_i,\ i = 1, \ldots, n$. First, observe that in the system (4) the inverse of each uncertain parameter $\theta_i = \frac{1}{z_i},\ i = 1, \ldots, n$ appears linearly in the same equation of the control inputs, as it can been verified from Eqs. (1), (2); this is the case of matching condition [11], i.e. the level of uncertainty is zero. The control system design is also complicated by the fact that the uncertain parameters are time-varying. In order to obtain fast adaptation, two sources of parameters information are used: the tracking error and the prediction error. The following adaptive control scheme has been designed:

$$\begin{aligned} \boldsymbol{u}(k) &= -\sigma \hat{G}_d^{+}(\boldsymbol{p}(k))\tilde{\boldsymbol{p}}(k) \\ \hat{\theta}_i(k+1) &= \hat{\theta}_i(k) + \gamma_1 T(-\tilde{p}_{x_i}(k) f u_1(k) \\ &- \tilde{p}_{y_i}(k) f u_2(k) + (p_{x_i}(k)\tilde{p}_{x_i}(k) \\ &+ p_{y_i}(k)\tilde{p}_{y_i}(k)) u_3(k)) - \gamma_2 (C_i(k)\boldsymbol{u}_T(k))^{\mathrm{T}} \\ & \quad \boldsymbol{e}_{p_i}(k),\ \sigma \in (0,\ 2),\ \gamma_1 > 0,\ \gamma_2 > 0 \end{aligned} \qquad (9)$$

$\hat{\theta}_i(k),\ i = 1, \ldots, n$ 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 $G_d(\boldsymbol{p}(k))$, $\hat{G}_d^{+}(\boldsymbol{p}(k))$ is its Moore-Penrose pseudo-inverse, $\boldsymbol{e}_{p_i}(k)$ is the prediction error defined below, and

$$C_i(k) = T \begin{bmatrix} -f & 0 & p_{x_i}(k) \\ 0 & -f & p_{y_i}(k) \end{bmatrix} \quad . \qquad (10)$$

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 parameterization form: $\boldsymbol{d}_i(k) = \boldsymbol{p}_i(k+1) - \boldsymbol{p}_i(k) - T \begin{bmatrix} \boldsymbol{b}_{x_i}^{\mathrm{T}} \\ \boldsymbol{b}_{y_i}^{\mathrm{T}} \end{bmatrix} \boldsymbol{u}_R = (C_i(k)\boldsymbol{u}_T(k))\theta_i,\ \ i = 1, \ldots, n,$ and $\hat{\boldsymbol{d}}_i(k)$ is obtained by substituting the estimates $\hat{\theta}_i(k),\ i = 1, \ldots, n$ in the expression of $\boldsymbol{d}_i(k)$.

Let $[0\ T_c]$ be the time interval in which the system (4) is controlled by (9). Indicate with $\mathcal{B}(\boldsymbol{x}, r) = \{\boldsymbol{x} \in \Re^{n} : \|\boldsymbol{x}\| \leq r\},\ r > 0$ a closed ball in $\Re^{n}$.

Defined the vectors:
$\boldsymbol{\theta}(k) = [\theta_1(k)\ \ldots\ \theta_n(k)]^{\mathrm{T}}$, $\tilde{\boldsymbol{\theta}}(k) = \left[\tilde{\theta}_1(k)\ \ldots\ \tilde{\theta}_n(k)\right]^{\mathrm{T}}$,

where $\tilde{\theta}_i(k) = \hat{\theta}_i(k) - \theta_i(k),\ i = 1,\dots,n,\ \Delta\boldsymbol{\theta}(k) = [\Delta\theta_1(k)\ \dots\ \Delta\theta_n(k)]^{\mathrm{T}},\ \Delta\tilde{\boldsymbol{\theta}}(k) = \left[\Delta\tilde{\theta}_1(k)\ \dots\ \Delta\tilde{\theta}_n(k)\right]^{\mathrm{T}}$, and $\tilde{\boldsymbol{\theta}}_r(k) = [\|C_1(k)\boldsymbol{u}_T(k)\|\ \tilde{\theta}_1(k)\ \dots\ \|C_n(k)\boldsymbol{u}_T(k)\|\tilde{\theta}_n(k)]^{\mathrm{T}}$. Define the matrix $H(\boldsymbol{p}(k)) = G_d^{\mathrm{T}}(\boldsymbol{p}(k))G_d(\boldsymbol{p}(k))$, and the perturbative matrix $M(\boldsymbol{p},\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$ such that $H(\boldsymbol{p}(k))\,\hat{H}(\boldsymbol{p}(k))^{-1} = I_6 + M(\boldsymbol{p},\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$, where $\hat{H}$ is obtained by substituting the estimates in the expression of H. Denote with $\mu = \|M\| = \max_i |\lambda_i(M)|$ the spectral norm of the matrix $M(\boldsymbol{p},\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$, where $\lambda_i(M),\ i = 1,\dots,6$ are its eigenvalues. Since assumption 1 is verified $\forall z_i \in \Re\setminus\{0\},\ i = 1,\dots,n$, hence, in particular, if the depth estimates are used in the expression of the interaction matrix, $\hat{H}$ is definite positive; let $\hat{\lambda}_M = \|\hat{H}\| = \max_i \lambda_i(\hat{H})$, and $\hat{\lambda}_m = \min_i \lambda_i(\hat{H})$ be, respectively, the maximum and minimum eigenvalue of $\hat{H}$, and indicate with $\hat{\rho} = \sqrt{\frac{\hat{\lambda}_M}{\hat{\lambda}_m}}$ the 2-norm condition number of matrix $\hat{G}_d(\boldsymbol{p}(k))$. Moreover, since the system in Eq. (4) is smooth, there exist finite constants $\xi_1,\ \xi_2 > 0$, such that $\|\Delta\boldsymbol{\theta}(k)\| < \xi_1$, and $\|\Delta\tilde{\boldsymbol{\theta}}(k)\| < \xi_2,\ \forall k \in [0\ T_c]$.

The stabilizing properties of the adaptive control design (9) are formally summarized in the following proposition.

Proposition 2 *In the above definitions, if assumption 1 is satisfied, the translational velocity input satisfies the condition* $\boldsymbol{u}_T \neq \boldsymbol{0},\ \forall k \in [0\ T_c]$, *and* $\sigma \in (0,\ \frac{2}{1+\mu\hat{\rho}^2})$, *then:*

- *the control system in Eq. (9), applied to (4), guarantees the uniform boundedness of the whole state* $(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\theta}})$ *of the closed-loop system;*
- *if no object point* $\boldsymbol{x}_i,\ i = 1,\dots,n$ *moves along its projection ray, the set-point error* $\tilde{\boldsymbol{p}}(k)$ *converges, at least, to the residual set* $\mathcal{B}(\tilde{\boldsymbol{p}}, \sqrt{\frac{\xi_1^2+2\gamma_2\delta^2\xi_2}{2\sigma^*\hat{\lambda}_m\gamma_1\gamma_2\delta^2}})$, *where* $\delta = \min_i \|C_i\boldsymbol{u}_T\|$, *and* $\sigma^* = -(\hat{\lambda}_M^{-1} + \mu\hat{\lambda}_m^{-1})\sigma^2 + 2\,\hat{\lambda}_M^{-1}\sigma > 0$.

Proof: Consider the time-invariant, positive definite, and radially unbounded function, defined in the whole state $(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\theta}})$: $V(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\theta}}) = \tilde{\boldsymbol{p}}^{\mathrm{T}}(k)\tilde{\boldsymbol{p}}(k) + \frac{1}{\gamma_1}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(k)\tilde{\boldsymbol{\theta}}(k) = \sum_{i=1}^n(\tilde{p}_{x_i}^2(k) + \tilde{p}_{y_i}^2(k) + \frac{1}{\gamma_1}\tilde{\theta}_i^2(k))$. Define the following matrix:

$$\hat{G}_{d_i} = T\begin{bmatrix}\hat{\boldsymbol{a}}_{x_i}^{\mathrm{T}} & \boldsymbol{b}_{x_i}^{\mathrm{T}}\\ \hat{\boldsymbol{a}}_{y_i}^{\mathrm{T}} & \boldsymbol{b}_{y_i}^{\mathrm{T}}\end{bmatrix}\quad, \tag{11}$$

with $\hat{\boldsymbol{a}}_{x_i}^{\mathrm{T}}$, and $\hat{\boldsymbol{a}}_{y_i}^{\mathrm{T}}$ obtained by substituting the estimates $\hat{\theta}_i,\ i = 1,\dots,n$ in the expressions of $\boldsymbol{a}_{x_i}^{\mathrm{T}}$, and $\boldsymbol{a}_{y_i}^{\mathrm{T}}$. The first order variation of $V(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\theta}})$ along the system dynamics (4), using Eq. (9), results:

$$\begin{aligned}
\Delta V(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\theta}}) = \sum_{i=1}^n(\tilde{p}_{x_i}(k+1)^2 + \tilde{p}_{y_i}(k+1)^2 \qquad (12)\\
- &\ \tilde{p}_{x_i}(k)^2 - \tilde{p}_{y_i}(k)^2 + \frac{1}{\gamma_1}(\tilde{\theta}_i(k+1)^2 - \tilde{\theta}_i(k)^2))\\
= &\ \sum_{i=1}^n[(\tilde{p}_{x_i}(k) + T\,\boldsymbol{a}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_T + T\,\boldsymbol{b}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_R)^2\\
+ &\ (\tilde{p}_{y_i}(k) + T\,\boldsymbol{a}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_T + T\,\boldsymbol{b}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_R)^2\\
- &\ \tilde{p}_{x_i}(k)^2 - \tilde{p}_{y_i}(k)^2 + \frac{1}{\gamma_1}(\Delta\tilde{\theta}_i(k)^2\\
+ &\ 2\tilde{\theta}_i(k+1)\tilde{\theta}_i(k) - 2\tilde{\theta}_i(k)^2)]\\
= &\ \sum_{i=1}^n[(T\,\boldsymbol{a}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_T + T\,\boldsymbol{b}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_R)^2\\
+ &\ (T\,\boldsymbol{a}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_T + T\,\boldsymbol{b}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_R)^2\\
+ &\ 2\,T\tilde{p}_{x_i}(k)\,\boldsymbol{a}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_T + 2\,T\,\tilde{p}_{x_i}(k)\boldsymbol{b}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_R\\
+ &\ 2\,T\tilde{p}_{y_i}(k)\,\boldsymbol{a}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_T\\
+ &\ 2\,T\,\tilde{p}_{y_i}(k)\boldsymbol{b}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_R + \frac{1}{\gamma_1}(\Delta\tilde{\theta}_i(k)^2\\
+ &\ 2\tilde{\theta}_i(k+1)\tilde{\theta}_i(k) - 2\tilde{\theta}_i(k)^2)]\\
= &\ \sum_{i=1}^n[\boldsymbol{u}^{\mathrm{T}}G_{d_i}^{\mathrm{T}}(k)G_{d_i}(k)\boldsymbol{u} - 2\,T\,\tilde{p}_{x_i}(k)\,\theta_i(k)\,f\,u_1\\
- &\ 2\,T\,\tilde{p}_{y_i}(k)\,\theta_i(k)\,f\,u_2\\
+ &\ 2\,T\,\theta_i(k)(p_{x_i}(k)\tilde{p}_{x_i}(k) + p_{y_i}(k)\tilde{p}_{y_i}(k))u_3\\
+ &\ 2\,T\,\tilde{p}_{x_i}(k)\boldsymbol{b}_{x_i}^{\mathrm{T}}(k)\boldsymbol{u}_R + 2\,T\,\tilde{p}_{y_i}(k)\boldsymbol{b}_{y_i}^{\mathrm{T}}(k)\boldsymbol{u}_R\\
+ &\ \frac{2}{\gamma_1}(\hat{\theta}_i(k)\Delta\hat{\theta}_i(k) - \tilde{\theta}_i(k)\Delta\theta_i(k))\\
+ &\ \frac{1}{\gamma_1}\Delta\tilde{\theta}_i(k)^2]\\
= &\ \sum_{i=1}^n[2T\,\theta_i(k)(-\tilde{p}_{x_i}(k)fu_1 - \tilde{p}_{y_i}(k)fu_2 + (p_{x_i}(k)\tilde{p}_{x_i}(k)\\
+ &\ p_{y_i}(k)\tilde{p}_{y_i}(k))u_3 - \frac{\Delta\hat{\theta}_i(k)}{T\,\gamma_1})\\
+ &\ 2\tilde{p}_{x_i}(k)T\,\boldsymbol{b}_{x_i}^{\mathrm{T}}\boldsymbol{u}_R + 2\tilde{p}_{y_i}(k)T\boldsymbol{b}_{y_i}^{\mathrm{T}}\boldsymbol{u}_R\\
+ &\ \boldsymbol{u}^{\mathrm{T}}G_{d_i}^{\mathrm{T}}(k)G_{d_i}(k)\boldsymbol{u}\\
+ &\ \frac{2}{\gamma_1}(\hat{\theta}_i(k)\Delta\hat{\theta}_i(k) - \tilde{\theta}_i(k)\Delta\theta_i(k)) + \frac{1}{\gamma_1}\Delta\tilde{\theta}_i(k)^2]\\
= &\ \sum_{i=1}^n(2\tilde{\boldsymbol{p}}_i^{\mathrm{T}}(k)\hat{G}_{d_i}(k)\boldsymbol{u} + \boldsymbol{u}^{\mathrm{T}}G_{d_i}^{\mathrm{T}}(k)G_{d_i}(k)\boldsymbol{u}\\
+ &\ \frac{1}{\gamma_1}(-2\tilde{\theta}_i(k)\,\Delta\theta_i(k) + \Delta\tilde{\theta}_i(k)^2)\\
- &\ 2\frac{\gamma_2}{\gamma_1}\|C_i(k)\boldsymbol{u}_T\|^2\tilde{\theta}_i^2(k))
\end{aligned}$$

where it has been used the fact that

$$\begin{aligned}
\tilde{\theta}_i(k+1)^2 - \tilde{\theta}_i(k)^2 &= (\tilde{\theta}_i(k+1) - \tilde{\theta}_i(k))^2\\
&+ 2\,\tilde{\theta}_i(k+1)\,\tilde{\theta}_i(k) - 2\,\tilde{\theta}_i(k)^2\\
\tilde{\theta}_i(k+1)\,\tilde{\theta}_i(k) - \tilde{\theta}_i(k)^2 &= \hat{\theta}_i(k)\Delta\hat{\theta}_i(k)\\
&- \theta_i(k)\Delta\hat{\theta}_i(k) - \tilde{\theta}_i(k)\Delta\theta_i(k)\quad.
\end{aligned}$$

Define the vector $\hat{\tilde{\boldsymbol{q}}} = \hat{G}^{\mathrm{T}}(\boldsymbol{p})\,\tilde{\boldsymbol{p}} \in \Re^6$, by using Eq. (9), yields

$$\begin{aligned}
&\Delta V(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\theta}}) = && (13)\\
- \;& 2\sigma \tilde{\boldsymbol{p}}^{\mathrm{T}}(k)\, \hat{G}_d(\boldsymbol{p}(k))\, \hat{G}_d^{+}(\boldsymbol{p}(k)) \\
& \tilde{\boldsymbol{p}}(k) + \sigma^2 \tilde{\boldsymbol{p}}^{\mathrm{T}}(k) \hat{G}_d(\boldsymbol{p}(k))^{+\,\mathrm{T}} G_d^{\mathrm{T}}(\boldsymbol{p}(k)) G_d(\boldsymbol{p}(k)) \\
& \hat{G}_d^{+}(\boldsymbol{p}(k))\tilde{\boldsymbol{p}}(k) - \frac{2}{\gamma_1}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(k)\Delta\boldsymbol{\theta}(k) \\
+ \;& \frac{1}{\gamma_1}\|\Delta\tilde{\boldsymbol{\theta}}(k)\|^2 - \frac{2\gamma_2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}_r(k)\|^2 \\
= \;& \hat{\tilde{\boldsymbol{q}}}^{\mathrm{T}}(k)[(-2\sigma + \sigma^2)\hat{H}(\boldsymbol{p}(k))^{-1} \\
+ \;& \sigma^2 \hat{H}(\boldsymbol{p}(k))^{-1} M(\boldsymbol{p}, \boldsymbol{\theta}, \hat{\boldsymbol{\theta}})]\, \hat{\tilde{\boldsymbol{q}}}(k) \\
- \;& \frac{2}{\gamma_1}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(k)\Delta\boldsymbol{\theta}(k) \\
+ \;& \frac{1}{\gamma_1}\|\Delta\tilde{\boldsymbol{\theta}}(k)\|^2 - \frac{2\gamma_2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}_r(k)\|^2 \\
\leq \;& -\sigma^* \|\hat{\tilde{\boldsymbol{q}}}(k)\|^2 + \frac{2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}(k)\|\,\|\Delta\boldsymbol{\theta}(k)\| \\
+ \;& \frac{1}{\gamma_1}\|\Delta\tilde{\boldsymbol{\theta}}(k)\|^2 - \frac{2\gamma_2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}_r(k)\|^2
\end{aligned}$$

where it has been used the fact that

$$\|\sigma^2\, \hat{H}^{-1}\, M\| \leq \sigma^2 \lambda_m^{-1}\mu$$

and

$$\begin{aligned}
\|\hat{\tilde{\boldsymbol{q}}}^{\mathrm{T}}(k)[(-2\sigma \;&+\; \sigma^2)\hat{H}^{-1} + \sigma^2 \hat{H}^{-1} M]\, \hat{\tilde{\boldsymbol{q}}}(k)\| && (14)\\
&\leq ((\hat{\lambda}_M^{-1} + \mu\hat{\lambda}_m^{-1})\sigma^2 \\
&-\; 2\,\hat{\lambda}_M^{-1}\sigma)\|\hat{\tilde{\boldsymbol{q}}}(k)\|^2 = -\sigma^*\|\hat{\tilde{\boldsymbol{q}}}(k)\|^2 \quad .
\end{aligned}$$

Notice that $\sigma^* > 0$, if the control gain σ is chosen in the open interval $(0,\ \frac{2}{1+\mu\hat{\rho}^2})$. Due to the assumption 1, the lemmas (1), (2) also hold if the uncertain parameters are substituted with their estimates $\hat{\theta}_i,\ i = 1, \ldots, n$ in the expression of $G(\boldsymbol{p})$. Since the system in Eq. (4) is smooth, there exist finite constants $\xi_1,\ \xi_2 > 0$, such that $\|\Delta\boldsymbol{\theta}(k)\| < \xi_1$, and $\|\Delta\tilde{\boldsymbol{\theta}}(k)\| < \xi_2,\ \forall k \in [0\ T_c]$.

Using these upper-bounds on the uncertain parameters, the first order variation of $V(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\theta}})$ is negative, if the following inequality holds

$$\frac{2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}(k)\|\xi_1 + \frac{1}{\gamma_1}\xi_2 < \sigma^*\|\hat{\tilde{\boldsymbol{q}}}(k)\|^2 + \frac{2\gamma_2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}_r(k)\|^2 \quad . \qquad (15)$$

If the translational velocity input is nonzero: $\boldsymbol{u}_T \neq \boldsymbol{0},\ k \in [0\ T_c]$, then $\|\tilde{\boldsymbol{\theta}}_r(k)\| \neq \boldsymbol{0}, \forall k \in [0\ T_c]$. Eq. (15) holds, if, at least, $\|\tilde{\boldsymbol{\theta}}\|$, or $\|\hat{\tilde{\boldsymbol{q}}}\|$ are sufficiently large. Since $V(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\theta}})$ is time-invariant, positive definite, and radially unbounded, and using lemma (2), $\tilde{\boldsymbol{\theta}}(k)$ and $\tilde{\boldsymbol{p}}(k)$ are uniformly bounded.

If $\delta = \min_i \|C_i(k)\boldsymbol{u}_T\| > 0$ (i.e. no object point $\boldsymbol{x}_i,\ i = 1, \ldots, n$ moves along its projection ray), then $\|\tilde{\boldsymbol{\theta}}_r(k)\| \geq \delta\|\tilde{\boldsymbol{\theta}}(k)\|$, and using lemma (2), the first order variation of $V(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\theta}})$ is negative, at least, in the region:

$$\begin{aligned}
\mathcal{R} = \{\tilde{\boldsymbol{p}} \in \mathcal{X}, \tilde{\boldsymbol{\theta}} \in \Re^n : \frac{2}{\gamma_1}\|\tilde{\boldsymbol{\theta}}\|\xi_1 \\
+\; \frac{1}{\gamma_1}\xi_2 - \frac{2\gamma_2}{\gamma_1}\delta^2\|\tilde{\boldsymbol{\theta}}\|^2 < \sigma^*\hat{\lambda}_m\|\tilde{\boldsymbol{p}}\|^2\} \quad , \qquad (16)
\end{aligned}$$

where $\sqrt{\hat{\lambda}_m}$ is the minimum singular value of $\hat{G}(\boldsymbol{p})$. Defined the quantity $y = \|\tilde{\boldsymbol{\theta}}\|$, and the function $f(y) = -\frac{2\gamma_2}{\gamma_1}\delta^2\, y^2 + \frac{2}{\gamma_1}\,\xi_1\, y + \frac{1}{\gamma_1}\xi_2$, which has its maximum value in $y_M = \frac{\xi_1}{2\gamma_2\delta^2}$, Eq. (16) indicates that the first order variation of $V(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\theta}})$ is negative if the condition $\sigma^*\hat{\lambda}_m\|\tilde{\boldsymbol{p}}\|^2 > f(y_M)$ holds. It follows that $\tilde{\boldsymbol{p}}(k)$ converges, at least, to the residual set $\mathcal{B}(\tilde{\boldsymbol{p}}, \sqrt{\frac{\xi_1^2 + 2\gamma_2\delta^2\xi_2}{2\sigma^*\hat{\lambda}_m\gamma_1\gamma_2\delta^2}})$. ∎

The assumption $\boldsymbol{u}_T \neq \boldsymbol{0}, k \in [0\ T_c]$ is not restrictive but necessary to obtain an adaptive depth estimation. In fact, in the case $\boldsymbol{u}_T = \boldsymbol{0}$ the system (4) does not depend on depth parameters. In other words, it's impossible to reconstruct the object structure from motion without translational velocity, see, for example, [8]. In the proposed framework, fixed a threshold $\epsilon > 0$, a possible solution is to freeze the estimates $\hat{\theta}_i,\ i = 1, \ldots, n$, when $\|\boldsymbol{u}_T\| < \epsilon$. Finally, since a redundant number of image points is considered, i.e. $n > 3$, it's possible to eliminate from the state vector $\boldsymbol{p}$ the image point $\boldsymbol{p}_j$, $j \in \{1, \ldots, n\}$ such that $C_j\boldsymbol{u}_T = 0$, i.e. the object point $\boldsymbol{x}_j$ is moving along its projection ray. Hence, the condition $\delta > 0,\ \forall k \in [0\ T_c]$, and consequently the convergence towards the residual set $\mathcal{B}(\tilde{\boldsymbol{p}}, \sqrt{\frac{\xi_1^2 + 2\gamma_2\delta^2\xi_2}{2\sigma^*\hat{\lambda}_m\gamma_1\gamma_2\delta^2}})$ are always guaranteed. Notice that the condition $C_i\boldsymbol{u}_T \neq 0$ is easily detected, since all the involved quantities are measured directly.

4 Experimental Results

The proposed adaptive visual control design has been implemented on an eye-in-hand robotic system, consisting 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 f 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 velocity loop), and, for the outer visual loop, a PC MMX 200 Mhz equipped with an Imaging Technology frame grabber. In our system implementation the actuation delay has been estimated to be $56\,ms$, which is less than the sampling interval $T \approx 80\,ms$. In general, if an actuation delay of d samples is present, the stability of the system can be ensured, at the expenses of performances, if the control input is maintained constant for $d + 1$ samples, i.e., if the sampling time of the visual loop is increased to $(d+1)T$. It can be shown that negative definiteness of ΔV with sampling time $(d+1)T$ can be obtained by proper choice of σ, under the assumption that $G_d(\boldsymbol{p}(k)) \simeq G_d(\boldsymbol{p}(k-d))$, again at the expenses of performance. 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 of the following experiments. They provide a coarse initial estimation with an error up to 50% with respect to the real values. The values of the control and adaptive gains have been chosen as: $\sigma = 0.046$, $\gamma_1 = 0.0001$, and $\gamma_2 = 0.001$. Infinite impulse response digital filters were used for smoothing visual measurements, and in the expression of the linear parameterization form, it has been approximated $\boldsymbol{p}_i(k+1) - \boldsymbol{p}_i(k) \approx \boldsymbol{p}_i(k) - \boldsymbol{p}_i(k-1)$, $i = 1, \ldots, n$. The case of **3-D generic displacement** is addressed: the initial and desired pose of the camera are 3-D generic configurations. 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. Notice that the robot arm moves slowly during the visual feedback, performance limits and slow convergence are due mainly to the high value of the control loop period which is in our case of $80\,ms$.

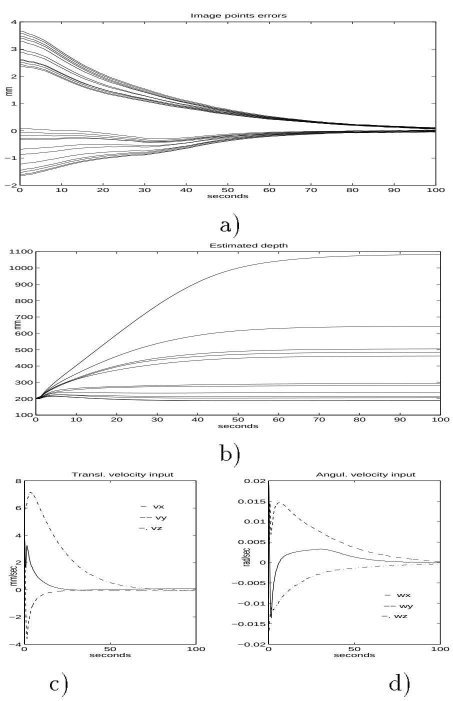

Figure 2: 3-D generic displacement ($n = 12$ control points of spline used): a) image points errors; b) estimated depth parameters; c) camera velocity components; d) camera angular velocity components.

5 Conclusions

The problem of controlling the relative pose between a robotic camera and a generic object of interest has been addressed by a nonlinear adaptive visual control system design in the discrete-time domain. The dynamic visual system is expressed in terms of image features, i.e. track-able 2-D points. Since the visual sampling period is not negligible at the current state of technology, the stability analysis is performed on a discrete-time representation of the camera-object visual model. In case of unknown depth parameters, the proposed 3-D nonlinear adaptation procedure, based on Lyapunov-design and prediction errors, ensures the ultimate boundedness of the whole state vector. An upper-bound of the residual set extension is also found. Robotic experiments show the effectiveness of the approach, and the satisfactory accuracy in the positioning of the camera. As future work, we are going to extend the proposed framework to the field of mobile robotics, and discrete-time nonlinear observers.

References

[1] B. Allotta and C. Colombo. "On the use of linear camera-object interaction models in visual servoing". *IEEE Transactions on Robotics and Automation.*, 15(2):346–352, 1999.

[2] F. Conticelli and B. Allotta. "Nonlinear controllability and stability analysis of image-based adaptive systems". In *IEEE Conference on Decision and Control, Phoenix, Arizona, USA*, 1999.

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

[4] J.T. Feddema and C.S.G. Lee. "Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera". *IEEE Trans. on Systems, Man and Cybernetics*, 20(5):1172–1183, 1990.

[5] J. Hill and W. T. Park. "Real time control of a robot with a mobile camera". *Proc. 9th ISIR, Washington, D.C.*, pages 233–246, Mar. 1979.

[6] S. Hutchinson, G.D. Hager, and P.I Corke. "Tutorial on visual servo control". *IEEE Trans. Rob. Autom.*, 12(5):651–670, 1996.

[7] A. Isidori. *Nonlinear Control Systems: An Introduction.* Springer Verlag, 1989.

[8] K. Kanatani. *Geometrical Computation for Machine Vision.* Oxford Science Publications, 1993.

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

[10] N.P. Papanikolopoulos, B.J. Nelson, and P.K. Khosla. "Six degree-of-freedom hand/eye visual tracking with uncertain parameters". *IEEE Trans. on Robot. Automat*, 11(5):725–732, 1995.

[11] D. Taylor, P.V Kokotović, R. Marino, and I. Kanellakopoulos. "Adaptive regulation of nonlinear systems with unmodeled dynamics". *IEEE Trans. on Automatic Control*, 34:405–412, 1989.