Stability Analysis of Invariant Visual Servoing and Robustness to Parametric Uncertainties

Ezio Malis

Stability Analysis of Invariant Visual Servoing and Robustness to Parametric Uncertainties

Ezio Malis

I.N.R.I.A. - ICARE Research Group, Sophia Antipolis, France.

Abstract. This paper concerns the stability analysis of a new visual servoing approach which is invariant on camera intrinsic parameters. Contrarily to standard methods, the invariant visual servoing approach can be used with a zooming camera or when the reference image is learned with a camera different from that used for servoing. Even if the error computed in an invariant space does not depend on the camera intrinsic parameters, they are needed to estimate the interaction matrix which links the camera velocity to the displacements of the features in the invariant space. Thus, calibration errors can affect the stability of the control law. For this reason, it is important to study the robustness of the proposed vision-based control with respect to uncertainties on the parameters of the system.

1 Introduction

Visual servoing is a very flexible method for the control of uncalibrated dynamic systems evolving in an unknown environment. Typical applications of visual servoing are the positioning of a robot and the tracking of objects using the information provided by an in-hand camera. The visual servoing approaches proposed in the literature [16,18] can be classified depending on the a priori knowledge available on the parameters of the system and on the observed object. If a 3D model of the object is available we can use a "*model-based*" approach [35,28], while if the 3D model of the object is unknown we must use a "*model-free*" approach [12,24]. Model-free methods, needs a preliminary learning step during which a reference image of the object is stored (teaching-by-showing). After the camera and/or the object have been moved, several vision-based control methods [2,12,25] have been proposed in order to drive the robot back to the reference position. When the current image observed by the camera is identical to the reference image the robot is back to the desired position. The model-free approach has the advantage of avoiding the knowledge of the model but it cannot be used with a zooming camera. If the camera *intrinsic* parameters (e.g. the focal length) change during the servoing, then the reference image must be learned again. Both model-based and model-free approaches are useful but, depending on the "a priori" knowledge we have of the scene, we must switch between them. In order to solve this problem, I propose in this paper a unified approach to vision-based control which can be used whether the model of the object is known or not [22]. The key idea of the unified approach, which is an extension of the work presented

in [21], is to build a reference in a projective space which can be computed if the model is known or if an image of the object is available. Thus, only one low level visual servoing technique must be implemented at once. The new unified approach is called *invariant* visual servoing since we work in a projective space which is invariant to camera intrinsic parameters [21] and at the same time invariant to the knowledge of the 3D model of the object [22]. Contrarily to standard model-free approaches, this allows us to use the invariant visual servoing approach with a zooming camera or to learn the reference image with a camera different from that used for servoing [20]. There are various ways in which invariance to camera parameters can be obtained [14,34,21]. In [21] invariance to all the camera intrinsic parameters has been obtained by selecting three interest points to build a projective transformation. Consequently, the selection of the three points raised the problem of the best choice. The problem has been solved in [22] by building the projective transformation from all points available in the image. The control in the invariant space can be carried-out within the task-function framework [29] and its structure is very similar to standard image-based approaches [12,16,18]. Even if the task function of the invariant visual servoing does not depend on the camera intrinsic parameters, they are needed to estimate the interaction matrix which links the camera velocity to the displacements of the features in the invariant space [23]. Thus, calibration errors can affect the stability of the control law. In the recent past, research on the stability of image-based visual servoing has been concentrated on the solution of convergence problems [5]. Indeed, the image-based approach is a local method which, even in the absence of calibration errors, can fail if the initial camera displacement is too large [5]. In order to avoid these potential convergence problems several possible solutions have been proposed: hybrid, partitioned and interpolation approaches. In hybrid approaches, some global information is introduced by estimating the camera displacement between the current and reference views [25,27,8]. The rotation of the camera is thus controlled directly in the Cartesian space while some image-based information is used to control the translation. More recently, a partitioned approach [7] has been proposed in order to avoid the camera displacement reconstruction. Another solution to potential stability problem of the image-based approach is provided by interpolation approaches. These methods define a path in the image by interpolating initial and reference image features [17,26,23]. Thus, the error in the image is maintained small at each iteration of the control law. Interpolation approaches are an elegant solution to potential convergence problems of local approaches. In the case of the invariant visual servoing it is even possible to define a path in the projective space such that the robot follows a straight line in the Cartesian space [23]. Even using interpolation approaches, the problem of finding the local robustness domain of the vision-based control law has not been yet solved. Due to the complexity of the problem, only few theoretical results have been obtained concerning the stability analysis of image-based visual

servoing in the presence of calibration errors. The theoretical analysis has been carried out only in very simple cases [11], often considering a simplified model for the camera intrinsic parameters [6,19,9] but always supposing that the depth distribution was perfectly estimated [11,6,19,9,23]. The objective of this paper is to to study the robustness of visual servoing with respect to both intrinsic and extrinsic camera parameters. The application of robust control analysis tools [1,4,31] allows us to find an approximation of the robustness domain of the invariant visual servoing.

2 Modeling

2.1 Perspective projection

Let $\mathcal{F}_0$ be a frame attached to an object represented by the homogeneous coordinates of a discrete set of n 3D points $\boldsymbol{\mathcal{X}}_i = (X_i, Y_i, Z_i, 1)$ $(i = \{1, 2, ..., n\})$. Let $\mathcal{F}$ be the current camera frame and let the origin of the frame coincide with the center of projection. Let the plane of projection be parallel to the plane $(\overrightarrow{x}, \overrightarrow{y})$. Without loss of generality we can suppose that the distance between the two planes is 1. A 3D point $\boldsymbol{\mathcal{X}}_i \in \mathbb{P}^3$ is projected to the point $\mathbf{m}_i \in \mathbb{P}^2$ with normalized homogeneous coordinates:

$$\mathbf{m}_i = \frac{1}{Z_i} \left[\mathbf{R}_0 \ \mathbf{t}_0 \right] \boldsymbol{\mathcal{X}}_i = (x_i, y_i, 1) \tag{1}$$

where $\mathbf{R}_0$ and $\mathbf{t}_0$ are respectively the rotation and the translation between frame $\mathcal{F}_0$ and $\mathcal{F}$.

2.2 Camera model

Pinhole cameras perform a perspective projection of a 3D point. The information measured by the camera is an image point $\mathbf{p}_i$:

$$\mathbf{p}_i = \mathbf{K}\, \mathbf{m}_i = (u_i, v_i, 1) \tag{2}$$

The triangular matrix $\mathbf{K}(t)$ contains the camera intrinsic parameters:

$$\mathbf{K}(t) = \begin{bmatrix} f & sf & u_0 \\ 0 & rf & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where f is the focal length (measured in pixels), u_0 and v_0 are the coordinates of the principal point (in pixels), s is the skew and r is the aspect ratio. In most of the papers dealing with the stability analysis of visual servoing several parameters are often supposed to be known. While s and r can be accurately estimated this is not true for the principal point. It has been shown not only that its calibration is very sensitive to noise [13] but also that its inaccurate

location has a significant effect on the calibration of others parameters or on the reconstruction accuracy [15]. The non-singular (3×3) matrix $\mathbf{K}(t)$ defines a projective transformation from the normalized coordinate system $\mathcal{M}$ to the image coordinate system $\mathcal{P}$. Thus, an approximation $\widehat{\mathbf{K}}$ of the matrix is needed in order to estimate a normalized point from a measured image point: $\widehat{\mathbf{m}}_i = \widehat{\mathbf{K}}^{-1}\,\mathbf{p}_i = (\widehat{x}_i, \widehat{y}_i, 1)$.

2.3 Invariance to camera intrinsic parameters

Suppose that n points are available. Using all the image points, with projective coordinates $\mathbf{p}_i = (u_i, v_i, 1)$, and all the the normalized points, with projective coordinates $\mathbf{m}_i = (x_i, y_i, 1)$, we can compute the following symmetric (3×3) matrices:

$$\mathbf{S}_p = \frac{1}{n}\sum_{i=1}^{n} \mathbf{p}_i\mathbf{p}_i^\top \qquad \text{and} \qquad \mathbf{S}_m = \frac{1}{n}\sum_{i=1}^{n} \mathbf{m}_i\mathbf{m}_i^\top \tag{4}$$

Since $\mathbf{p}_i = \mathbf{K}\mathbf{m}_i$, the matrix $\mathbf{S}_p$ can be written as a function of $\mathbf{S}_m$ and of the camera intrinsic parameters $\mathbf{K}$:

$$\mathbf{S}_p = \frac{1}{n}\sum_{i=1}^{n} \mathbf{p}_i\mathbf{p}_i^\top = \mathbf{K}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{m}_i\mathbf{m}_i^\top\right)\mathbf{K}^\top = \mathbf{K}\,\mathbf{S}_m\,\mathbf{K}^\top \tag{5}$$

If the points are not collinear and $n > 3$ then $\mathbf{S}_p$ and $\mathbf{S}_m$ are positive definite matrices and they can be written, using a Cholesky decomposition, as:

$$\mathbf{S}_p = \mathbf{T}_p\,\mathbf{T}_p^\top \qquad \text{and} \qquad \mathbf{S}_m = \mathbf{T}_m\,\mathbf{T}_m^\top \tag{6}$$

where both $\mathbf{T}_p$ and $\mathbf{T}_m$ are (3×3) non-singular upper triangular matrices. Thus, from equations (5) and (6) we obtain:

$$\mathbf{T}_p = \mathbf{K}\,\mathbf{T}_m \tag{7}$$

The matrix $\mathbf{T}_p$ defines a projective transformation from the projective image space $\mathcal{P} \in \mathbb{P}^2$ to a new projective space $\mathcal{Q} \in \mathbb{P}^2$. Similarly, the matrix $\mathbf{T}_m$ defines a projective transformation from the projective space $\mathcal{M} \in \mathbb{P}^2$ to the same space $\mathcal{Q} \in \mathbb{P}^2$. We can compute the same vectors $\mathbf{q}_i \in \mathcal{Q}$ from image points:

$$\mathbf{q}_i = \mathbf{T}_p^{-1}\mathbf{p}_i \tag{8}$$

or from the knowledge of model of the object and the desired position:

$$\mathbf{q}_i = \mathbf{T}_m^{-1}\mathbf{m}_i \tag{9}$$

Equations (8) and (9) define the same point in $\mathcal{Q}$ since:

$$\mathbf{q}_i = \mathbf{T}_p^{-1}\mathbf{p}_i = \mathbf{T}_m^{-1}\mathbf{K}^{-1}\mathbf{p}_i = \mathbf{T}_m^{-1}\mathbf{K}^{-1}\mathbf{K}\mathbf{m}_i = \mathbf{T}_m^{-1}\mathbf{m}_i$$

As a consequence, the new projective space $\mathcal{Q}$ is independent on camera intrinsic parameters ($\mathbf{T}_m^{-1}$ and $\mathbf{m}_i$ do not depends on camera intrinsic parameters). Thus, the position of the camera can be controlled by driving the set of invariants $\mathbf{q}_i \ \forall i \in \{1, 2, ..., n\}$ to some reference values $\mathbf{q}_i^* \ \forall i \in \{1, 2, ..., n\}$ while the zoom of the camera can be controlled separately.

3 Vision-based control

The control of the pose of the camera rigidly attached to the robot end-effector is achieved by minimizing an error computed in the space $\mathcal{Q}$ which only depends on the camera pose. Let $\mathbf{t}$ and $\mathbf{R}$ be respectively the translation and the rotation between the reference camera frame $\mathcal{F}^*$ and the current camera frame $\mathcal{F}$. Let $\mathbf{r} = \theta\mathbf{u}$ be the (3×1) vector containing the axis of rotation $\mathbf{u}$ and the angle of rotation θ. Then, $\boldsymbol{\xi} = (\mathbf{t}, \mathbf{r})$ is a (6×1) vector containing global coordinates of an open subset $\mathcal{S} \subset \mathbb{R}^3 \times SO(3)$. The key idea of the proposed visual servoing approach is to always work in a projective space $\mathcal{Q} \in \mathbb{P}^2$ which can be computed from points belonging to the image space $\mathcal{P} \in \mathbb{P}^2$ (if the model is unknown) or points belonging to the projective space $\mathcal{M} \in \mathbb{P}^2$ (if the model is known). The approach does not need the explicit calibration of the camera and can be used even if the camera is zooming as shown in Figure 1.

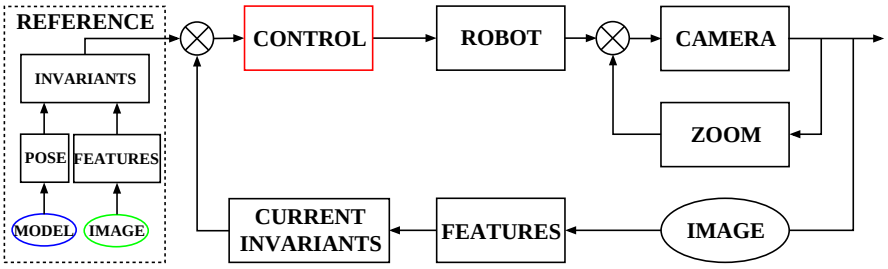


Fig. 1. Block diagram of the invariant control approach.

Similarly to the standard image-based approach, the control of the camera is achieved by stacking all the reference points of space $\mathcal{Q}$ in a (3n×1) vector $\mathbf{s}^*(\boldsymbol{\xi}^*) = (\mathbf{q}_1^*, \mathbf{q}_2^*, \cdots, \mathbf{q}_n^*)$. Similarly, the points measured in the current camera frame are stacked in the (3n×1) vector $\mathbf{s}(\boldsymbol{\xi}) = (\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n)$. If $\mathbf{s}(\boldsymbol{\xi}) = \mathbf{s}^*(\boldsymbol{\xi}^*)$ then $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ and the camera is back to the reference position whatever the camera intrinsic parameters. The derivative of vector $\mathbf{s}$ is:

$$\dot{\mathbf{s}} = \mathbf{L}\,\mathbf{v} \tag{10}$$

where the (3n×6) matrix $\mathbf{L}$ is called the interaction matrix and $\mathbf{v}$ is the velocity of the camera. The interaction matrix depends on current normalized

points $\mathbf{m}_i(\boldsymbol{\xi}) \in \mathcal{M}$ ($\mathbf{m}_i$ can be computed from image points $\mathbf{m}_i = \mathbf{K}^{-1}\,\mathbf{p}_i$), on the invariant points $\mathbf{q}_i(\boldsymbol{\xi}) \in \mathcal{Q}$ and on the current depth distribution $\mathbf{z}(\boldsymbol{\xi}) = (Z_1, Z_2, ..., Z_n)$. Consider the following (6×1) task function:

$$\mathbf{e} = \widehat{\mathbf{L}}^{+}(\mathbf{s} - \mathbf{s}^*)$$

where $\widehat{\mathbf{L}}^{+}$ is the pseudo-inverse of an approximation of the true (3n×6) interaction matrix. In [12], the interaction matrix is supposed to be constant. In this paper, we consider the most general case when $\mathbf{C}$ depends on $\mathbf{s}$ as needed in [22]. In that case :

$$\dot{\mathbf{e}} = \frac{d\widehat{\mathbf{L}}^{+}}{dt}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^{+}\dot{\mathbf{s}} = (\mathbf{O}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^{+}\mathbf{L})\,\mathbf{v} \tag{11}$$

where $\mathbf{O}(\mathbf{s} - \mathbf{s}^*)$ is a 6×6 matrix such that $\mathbf{O}(\mathbf{s} - \mathbf{s}^*)|_{\mathbf{s}=\mathbf{s}^*} = 0$. Consider the following proportional control law:

$$\mathbf{v} = -\lambda\,\mathbf{e} \tag{12}$$

where λ is a positive scalar factor which tunes the speed of convergence. Using this control law the robot can be driven back to the reference position.

4 Stability Analysis

In this section, the local stability of the control law (12) is analyzed. The local stability is valid only in a neighborhood of the equilibrium point. However, when the initial error is large, it is possible to sample the trajectory and consider only small errors at each iteration of the control law [23]. Plugging equation (12) into (11), we obtain the following closed-loop equation:

$$\dot{\mathbf{e}} = -\lambda(\mathbf{O}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^{+}\mathbf{L})\mathbf{e} \tag{13}$$

It is well know from control theory that the non-linear system (13) is locally asymptotically stable in a neighborhood of $\mathbf{s} = \mathbf{s}^*$ if and only if the linearized system is stable:

$$\dot{\mathbf{e}} = -\lambda\mathbf{Q}\mathbf{e} \tag{14}$$

where $\mathbf{Q} = \widehat{\mathbf{L}}^{+}\mathbf{L}|_{\mathbf{s}=\mathbf{s}^*}$. The linear system (14) is asymptotically stable *if and only if* $\mathbf{Q}$ has eigenvalues with positive real part:

$$\mathrm{real}(\mathrm{eig}(\mathbf{Q})) = \mathrm{real}(\mathrm{eig}(\widehat{\mathbf{L}}^{+}\mathbf{L})) > 0$$

The matrix depends on two set of unknown parameters: $\mathbf{Q} = \mathbf{Q}(\widehat{\mathbf{K}}, \widehat{\mathbf{z}})$. Obviously if $\mathbf{K} = \widehat{\mathbf{K}}$ and $\widehat{\mathbf{z}} = \mathbf{z}$ then $\mathbf{Q} = \mathbf{I}$ and the system is stable. If $\mathbf{Q}$ is full rank then $\mathbf{e} = 0$ is the only equilibrium point of the system (i.e. if $\|\mathbf{e}\|$

decreases then it decreases towards $\mathbf{e} = 0$). It is well known from control theory that if $\mathbf{Q} > 0$ then the norm of the task function $\|\mathbf{e}\|$ decreases to zero. However, we need to prove that so does the error $\mathbf{s} - \mathbf{s}^*$ in the absence of local minima and singularities. The problem to know if, and in which case, a local minimum can be found is beyond the aim of this paper and it will be addressed in future work. In this paper, $\mathbf{Q}$ is supposed to be full rank. As already mentioned, the local asymptotic stability of the system can be proved considering the system linearized around $\mathbf{e} = 0$ (i.e. $\boldsymbol{\xi} = \boldsymbol{\xi}^*$):

$$\dot{\mathbf{e}} = -\lambda \mathbf{Q}(\boldsymbol{\xi})|_{\mathbf{e}=0}\mathbf{e} = -\lambda \mathbf{Q}(\boldsymbol{\xi}^*)\mathbf{e}$$

where $\mathbf{Q}(\boldsymbol{\xi}^*) = \widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)\mathbf{L}(\boldsymbol{\xi}^*)$. The system is locally stable if $\mathbf{Q}(\boldsymbol{\xi}^*) > 0$ since in that case $\mathbf{Q}(\boldsymbol{\xi}^*)$ has eigenvalues with positive real part. However, to prove the local asymptotic convergence of $\mathbf{e}$ to zero, we need also to show that $\mathbf{s} - \mathbf{s}^*$ never belongs to $\mathrm{Ker}(\widehat{\mathbf{J}}^+)$. This means that it exists a neighborhood $\mathcal{U}$ of $\boldsymbol{\xi}^*$ such that $\mathbf{e} = \widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) \neq 0$, $\forall \boldsymbol{\xi} \in \mathcal{U}$ (i.e. $\mathbf{e} = 0$ only if $\mathbf{s}(\boldsymbol{\xi}) = \mathbf{s}^*$). Let us suppose that $\mathbf{s}(\boldsymbol{\xi}) \neq \mathbf{s}^*$ and therefore $\boldsymbol{\xi} \neq \boldsymbol{\xi}^* = 0$. The Taylor development of $\mathbf{s}(\boldsymbol{\xi})$ in a neighborhood of $\boldsymbol{\xi}^* = 0$ is:

$$\mathbf{s} - \mathbf{s}^* = \mathbf{L}(\boldsymbol{\xi}^*)\ \boldsymbol{\xi} + O^2(\boldsymbol{\xi}) \tag{15}$$

Multiplying by $\boldsymbol{\xi}^T\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)$ (where $\boldsymbol{\xi}^T\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*) \neq 0$ since $\boldsymbol{\xi} \neq 0$ and $\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)$ is full rank) both sides of equation (15) we obtain:

$$\boldsymbol{\xi}^T\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) = \boldsymbol{\xi}^T\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)\mathbf{L}(\boldsymbol{\xi}^*)\boldsymbol{\xi} + O^3(\boldsymbol{\xi})$$

remember that if $\mathbf{Q} = \widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)\mathbf{L}(\boldsymbol{\xi}^*) > 0$ then $\boldsymbol{\xi}^T\mathbf{Q}\boldsymbol{\xi} \geq 2\sigma\|\boldsymbol{\xi}\|^2$, where $\sigma > 0$ is the minimum singular value of the positive definite matrix $\mathbf{Q} + \mathbf{Q}^T$. If $\widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) = 0$ then:

$$0 \geq 2\sigma\|\boldsymbol{\xi}\|^2 + O^3(\boldsymbol{\xi})$$

that means:

$$\|\boldsymbol{\xi}\|^2 \leq |O^3(\boldsymbol{\xi})|$$

which is impossible since, by definition of $O^3(\boldsymbol{\xi})$, it exists a neighborhood of $\boldsymbol{\xi}^*$ in which:

$$\|\boldsymbol{\xi}\|^2 > |O^3(\boldsymbol{\xi})|$$

Therefore, $\mathbf{e} = \widehat{\mathbf{L}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) \neq 0$ if $\mathbf{s} \neq \mathbf{s}^*$ in a neighborhood of $\boldsymbol{\xi}^*$ and the system is locally asymptotically stable.

5 Robustness to Parametric Uncertainties

The parameters used to compute the control law are only roughly known and some of the can vary with time when the camera is zooming. Let $\mathbf{g} =$

$(g_1, g_2, ..., g_p)$ a vector containing p parameters (in our case $p = n + 5$ where n is the total number of depth parameters and 5 is the total number of camera intrinsic parameters). The closed loop matrix $\mathbf{Q}$ can be expressed as a function of the true and estimated parameters $\mathbf{Q} = \mathbf{Q}(\mathbf{g}, \widehat{\mathbf{g}})$. Obviously, when the parameters are perfectly known $\mathbf{Q}(\mathbf{g}, \widehat{\mathbf{g}})|_{\widehat{\mathbf{g}}=\mathbf{g}} = \mathbf{I}$. Matrix $\mathbf{Q}$ can be viewed as a set of multivariate polynomials $\mathbf{Q}(\mathbf{g})$ whose variables are the parameters $\mathbf{g} = (g_1, g_2, ..., g_p)$. As already mentioned, the system is locally stable if and only if the eigenvalues of $\mathbf{Q}$ have positive real part. The eigenvalues of the (6×6) matrix $\mathbf{Q}$ are the roots of the characteristic polynomial:

$$p(\lambda, \mathbf{g}) = \sum_{k=0}^{6} c_k(\mathbf{g})\lambda^k$$

where the coefficients $c_k(\mathbf{g})$ are polynomial functions of the uncertain parameters $\mathbf{g}$. Given the measurement precision on the parameters $\mathbf{g}_i \in \left[\underline{g}_i, \overline{g}_i\right]$, it is possible to test the stability of the system. The necessary and sufficient conditions for the roots of the polynomial to be positive are obtained from the Routh-Hurwitz stability criterion without explicitly computing them. To check if the polynomial is stable we need to transform the bounds on the uncertainty into bounds on the coefficients of the polynomial. Indeed, if $\mathbf{g} \in \left[\underline{g}, \overline{g}\right]$ then $c_k(\mathbf{g}) \in [\underline{c}_k, \overline{c}_k]$. Thanks to the Kharitonov theorem [1], the stability of the uncertain polynomial can thus quickly checked using the Routh-Hurwitz stability criterion on the Kharitonov polynomials. However, the bounds on the coefficients of the polynomials are often conservative. Thus, if the Kharitonov polynomials are not stable we cannot conclude on the stability of the original system.

An approximation of the robustness domain can be obtained by bounding directly the eigenvalues of the closed-loop matrix. After setting $\tilde{g}_i = g_i - \widehat{g}_i$, matrix $\mathbf{Q}$ can be approximated as:

$$\mathbf{Q}(\mathbf{g}) \approx \mathbf{Q}_0 + \sum_{i=1}^{p} \tilde{g}_i \mathbf{Q}_i \tag{16}$$

where $\mathbf{Q}_0 = \mathbf{I}$, $\mathbf{Q}_i = \frac{\partial \mathbf{Q}(\mathbf{g})}{\partial g_i}|_{\mathbf{g}=\widehat{\mathbf{g}}}$ and $\tilde{\mathbf{g}}_i = (\tilde{g}_1, \tilde{g}_2, ..., \tilde{g}_p)$. Matrix $\mathbf{Q}$ can be regarded as a perturbation of the identity matrix $\mathbf{I}$. Let us define the SPECTRAL VARIATION of $\tilde{\mathbf{M}}$ with respect to $\mathbf{M}$ [31]:

$$\mathrm{sv}_{\mathbf{M}}(\tilde{\mathbf{M}}) = \max_i \min_j |\tilde{\lambda}_i - \lambda_j|$$

The Bauer-Fike theorem states that [31]:

$$\mathrm{sv}_{\mathbf{M}}(\tilde{\mathbf{M}}) \leq \|\tilde{\mathbf{M}} - \mathbf{M}\|$$

In our case, the spectral variation $\mathbf{S}$ with respect to $\mathbf{I}$ is:

$$\mathrm{sv}_{\mathbf{I}}(\mathbf{Q}) = \max_i |\tilde{\lambda}_i - 1| \leq \|\mathbf{E}\|$$

Thus, a simple sufficient condition for the stability of $\mathbf{Q}$ is $\|\mathbf{E}\| < 1$. Indeed, if $\|\mathbf{E}\| < 1$ then:

$$\max_i |\tilde{\lambda}_i - 1| < 1$$

which implies $\tilde{\lambda}_i > 0$. From the definition of spectral variation, all others eigenvalues $\lambda_k \ \forall k$ are such that $|\tilde{\lambda}_k - 1| \leq |\tilde{\lambda}_i - 1|$. Thus, $|\tilde{\lambda}_k - 1| < 1$ which means $\tilde{\lambda}_k > 0 \ \forall k$. Now, since $\mathbf{E} = \sum_{i=1}^{n} \tilde{g}_i \mathbf{Q}_i$:

$$\|\mathbf{E}\| \leq \sum_{i=1}^{m} |\tilde{g}_i| \|\mathbf{Q}_i\|$$

setting $\mu_i = \|\mathbf{Q}_i\| > 0$ the condition can be imposed by bounding the previous inequality:

$$\sum_{i=1, i \neq j}^{n} \mu_i |\tilde{g}_i| < 1$$

in this equation, the error $|\tilde{g}_i|$ is weighted by the scalars μ_i. The smaller is μ_i the less is the influence of the error $|\tilde{g}_i|$ on the stability of the system. Numerical examples show that the control is particularly robust to uncertainties on camera intrinsic parameters. As it will be shown in the experimental results, more than 50 % error of the focal length estimation can be tolerated. On the other hand, the system is less robust to uncertainties on the depth distribution. As an example, 10 % error on the depth distribution can be enough to make the system unstable.

6 Open problems

The standard procedure for visual servoing approaches is to design a control law for the nominal system [16,18]. Only few authors [3,32,33] have applied robust control techniques in order to directly take into account uncertainties at the design level. Despite these technique may be very conservatives their application to the invariant visual servoing scheme seems a possible solution to deal with the problem of large uncertainties on both camera intrinsic and extrinsic parameters. Indeed, if the environment is completely unknown and the system is uncalibrated the stability of the visual servoing in the presence of large errors on the parameters, can become a serious issue since the robustness domain is unknown. In addition to the problem of providing some information about the unknown depths of the object in the camera frame (i.e. the camera *extrinsic* parameters), the proposed invariant approach introduces new control problems since several intrinsic parameters can vary with time. Since it is generally impossible to estimate accurately on-line all the time-varying parameters, it is necessary to take into account the influence

of such uncertainties on the stability of the visual servoing directly at the design level. Another unresolved problem is the control of the zoom of the camera. Indeed, the zoom can be used to enlarge the field of view (zoom out) if the object is getting out of the image and to reduce the field of view (zoom in) to improve the extraction of the visual features. Unfortunately, these two objectives cannot be achieved at the same time and a compromise must be found.

7 Experimental Results

The vision-based control approach proposed in the paper has been validated on the 3 d.o.f. system Argés (see Figure 2) at INRIA Sophia-Antipolis. The Argés monocular system is an experimental platform used to develop active vision algorithms. The hardware is made of on-the-shelf components:

- a Computer controlled CCD Camera Acom1 PAL with a f=5.9 to 47.2 mm zoom-lens, automatic AGC 18dB and motor iris, a numerical auto-focus on 10bits, white balance, plus rs232C and video interface;
- a Pan-tilt turret, from RobotSoft, with a resolution of 3.086 minutes of arcs, a 4 lbs capacity and a speed up to 300 deg/sec, using constant current bipolar motor drives, via a rs232C interface;
- a linear degree of freedom, from CharlyRobot, with a resolution of 0.1 mm, using a slow screw driven control;

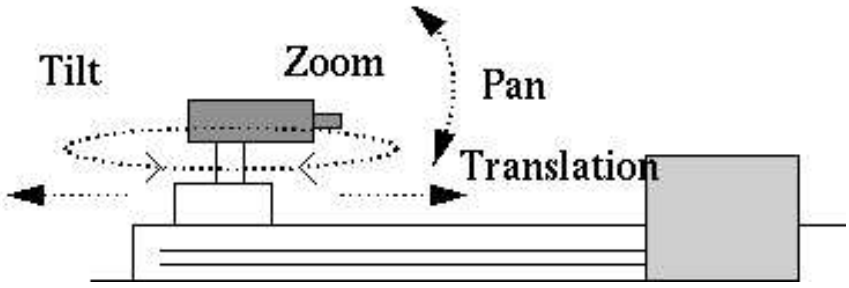
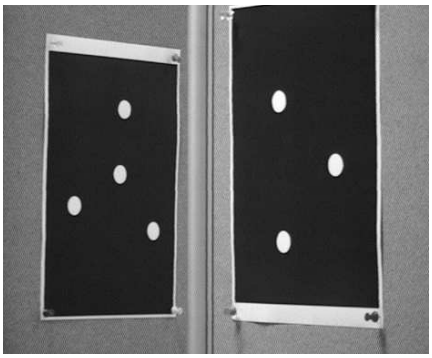


Fig. 2. The 3 d.o.f. robot Argés with a zooming camera.

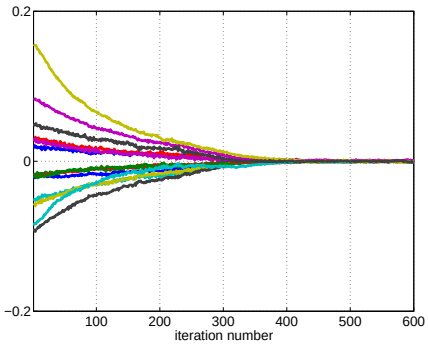
The objective of the experiment is to position the camera with respect to an object represented by 7 points (see Figure 3(a)). Suppose that the model is unknown (i.e. model-based approaches cannot be used) and that the camera is zooming during the servoing (i.e. standard model-free approaches cannot be used). On the other hand, we can use the approach proposed in the paper. The camera, with a focal length $f^* = 2500$ pixels, is driven to the reference position and the corresponding image is stored (see Figure 3(a)). Then, the

camera is displaced to its initial position. The initial displacement is $t_x = 350$ mm for the linear degree of freedom, $r_x = 4$ degrees and $r_y = 17$ degrees (i.e. tilt and pan respectively). At the initial position, the camera zooms out and the focal length is changed to $f_0 = 1670$ pixels. The corresponding initial image is given in Figure 3(b). The problem of matching/tracking features, common to all visual servoing techniques, is beyond the aim of this paper and it has been already investigated in the literature [10]. In the experiments, I focus on the general properties of the vision-based control approaches, therefore I consider that the matching/tracking problem has already been solved.

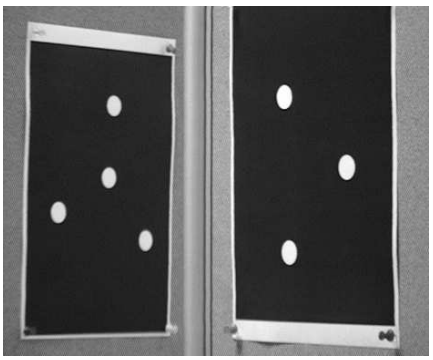


(a) Reference image

(b) Initial image



(c) Error in the invariant space

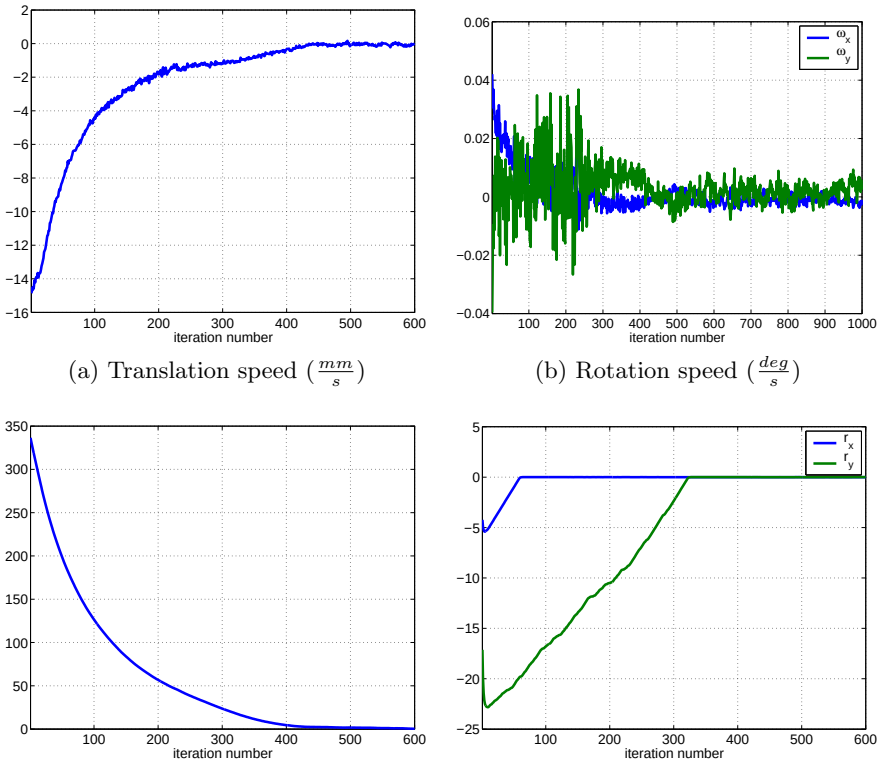


(d) Final image

Fig. 3. Camera positioning a with respect to 7 non coplanar points.

From the initial and reference images we can compute the invariants $\mathbf{q}_i$ and $\mathbf{q}_i^*$ $\forall i \in \{1, 2, ..., n\}$ in the projective space $\mathcal{Q}$. Using the control law plotted in Figures 4(c) and (d), the error $\mathbf{q}_i - \mathbf{q}_i^*$ is zeroed (except for noise) (Figure 3(c)). Consequently, the camera is back to the reference position (i.e. the translational and rotational errors in Figures 4(e) and (f) converge to zero

with an accuracy of 1 mm and 0.1 degrees). Since the camera parameters are unknown we use in the control law an extremely bad approximation of the focal length $\widehat{f} = 600$ pixels and we suppose that the principal point is in the center of the image. Despite the camera internal parameters $\widehat{\mathbf{K}}$ and the depth distribution $\widehat{\mathbf{z}}$ are only approximated, the control law is stable and converges. Obviously, if the calibration errors are big the performance of the visual servoing decreases (long time of convergence, unpredictable behavior). On the other hand, we can use the zoom of the camera to improve the performance of the servoing. During the servoing, the zoom is used to enlarge the field of view of the camera if the object is getting out of the image and to bound the size of the object to improve the robustness of features extraction. At the convergence, the camera focal length is $f \approx 2720$ (see Figure 4(b)). Consequently, the camera is back to the reference position, the images at the convergence (see Figure 4(d)) is not identical to the reference image because the camera has different focal lengths.



(a) Translation speed ($\frac{mm}{s}$)

(b) Rotation speed ($\frac{deg}{s}$)

(c) Translation error (mm)

(d) Rotation error (deg)

Fig. 4. Experimental results of the invariant visual servoing approach.

8 Conclusion

In this paper, it has been shown that the invariant visual servoing approach is robust to uncertainties on the parameters of the system. However, only a rough approximation of the robustness domain has been obtained. Since several intrinsic parameters can vary with time, and it is generally impossible to estimate accurately on-line all the time-varying parameters, it is necessary to take into account the influence of such uncertainties on the stability of the visual servoing. Thus, the approach could be considerably improved by directly applying robust control techniques at the design level.

Appendix A

The interaction matrix in the invariant space is obtained by stacking togheter all the interaction matrices $\mathbf{L}_{qi}$ relative to the points $\mathbf{q}_i$

$$\mathbf{L}_q = (\mathbf{L}_{q1}, \mathbf{L}_{q2}, ..., \mathbf{L}_{qn})$$

From equation (9), and knowing that $\dot{\mathbf{T}}_m^{-1} = -\mathbf{T}_m^{-1}\dot{\mathbf{T}}_m\mathbf{T}_m^{-1}$, we obtain the derivative of $\mathbf{q}_i$:

$$\dot{\mathbf{q}}_i = \dot{\mathbf{T}}_m^{-1}\mathbf{m}_i + \mathbf{T}_m^{-1}\dot{\mathbf{m}}_i = \mathbf{T}_m^{-1}\left(\dot{\mathbf{m}}_i - \mathbf{A}\mathbf{m}_i\right)$$

where $\mathbf{A}$ is the following triangular matrix:

$$\mathbf{A} = \dot{\mathbf{T}}_m\mathbf{T}_m^{-1} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The matrix $\mathbf{A}$ can be obtained by solving the following equation:

$$\dot{\mathbf{S}}_m = \dot{\mathbf{T}}_m\mathbf{T}_m^{\top} + \mathbf{T}_m\dot{\mathbf{T}}_m^{\top} = \dot{\mathbf{T}}_m\mathbf{T}_m^{-1}\mathbf{T}_m\mathbf{T}_m^{\top} + \mathbf{T}_m\mathbf{T}_m^{\top}\mathbf{T}_m^{-\top}\dot{\mathbf{T}}_m^{\top} = \mathbf{A}\mathbf{S}_m + \mathbf{S}_m\mathbf{A}^{\top}$$

The entries of the matrix $\mathbf{A}$ are linear functions of the camera velocity:

$$\mathbf{A}\mathbf{m}_i = \mathbf{C}_i\ \mathbf{v}$$

where $\mathbf{C}_i$ is a (3×6) matrix. Similarly, $\dot{\mathbf{m}}_i$ is a linear function of the camera velocity:

$$\dot{\mathbf{m}}_i = \mathbf{L}_{mi}\ \mathbf{v}$$

where:

$$\mathbf{L}_{mi} = \begin{bmatrix} -\frac{1}{Z_i} & 0 & \frac{x_i}{Z_i} & x_i y_i & -(1+x_i^2) & y_i \\ 0 & -\frac{1}{Z_i} & \frac{y_i}{Z_i} & 1+y_i^2 & -x_i y_i & -x_i \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the interaction matrix relative to the point $\mathbf{q}_i$ can be written as:

$$\mathbf{L}_{qi} = \mathbf{T}_m^{-1}\left(\mathbf{L}_{mi} - \mathbf{C}_i\ \right)$$

References

1. B. Barmish. *New Tools for Robustness of Linear Systems*. Macmilian, New York, 1994.
2. R. Basri, E. Rivlin, and I. Shimshoni. Visual homing: surfing on the epipole. *International Journal of Computer Vision*, 33(2):22–39, 1999.
3. D. Bellot and P. Danes. Towards an lmi approach to multicriteria visual servoing. In *European Control Conference*, Porto, Potugal, September 2001.
4. S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.
5. F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. Kriegman, G. Hager, and A. Morse, editors, *The confluence of vision and control*, volume 237 of *LNCIS Series*, pages 66–78. Springer Verlag, 1998.
6. C. C. Cheah, S. Kawamura, and S. Arimoto. Feedback control for robotic manipulator with uncertain kinematics and dynamics. In *IEEE Int. Conf. on Robotics and Automation*, volume 4, pages 3607–3612, Leuven, Belgium, May 1998.
7. P. Corke and S. Hutchinson. A new partitioned approach to image-based visual servo control. *IEEE Transactions on Robotics and Automation*, 14(4):507–515, August 2001.
8. K. Deguchi. Optimal motion control for image-based visual servoing by decoupling translation and rotation. In *IEEE Int. Conf. on Intelligent Robots and Systems*, volume 2, pages 705–711, October 1998.
9. L. Deng, F. Janabi-Sharifi, and W. J. Wilson. Stability and robustness of visual servoing methods. In *IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 1605–1609, Washington D.C., May 2002.
10. Y. Dufournaud, C. Schmid, and R. Horaud. Matching images with different resolutions. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 618–618, Hilton Head Island, South Carolina, USA, June 2000.
11. B. Espiau. Effect of camera calibration errors on visual servoing in robotics. In *3rd Int. Symposium on Experimental Robotics*, Kyoto, Japan, October 1993.
12. B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3):313–326, June 1992.
13. O. Faugeras. *Three-dimensionnal computer vision: a geometric viewpoint*. MIT Press, Cambridge, MA, 1993.
14. G. D. Hager. Calibration-free visual control using projective invariance. In *IEEE Int. Conf. on Computer Vision*, pages 1009–1015, MIT, Cambridge (USA), June 1995.
15. R. Hartley and R. Kaucic. Sensitivity of calibration to principal point position. In *European Conference on Computer Vision*, volume 2, pages 433–446. Copenhagen, Denmark, May 2002.
16. K. Hashimoto. *Visual Servoing: Real Time Control of Robot manipulators based on visual sensory feedback*, volume 7 of *World Scientific Series in Robotics and Automated Systems*. World Scientific Press, Singapore, 1993.
17. K. Hashimoto and T. Noritsugu. Enlargement of stable region in visual servo. In *IEEE Conference on Decision and Control*, pages 3927–3932, Sydney, Australia, December 2000.
18. S. Hutchinson, G. D. Hager, and P. I. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5):651–670, October 1996.

19. R. Kelly, R. Carelli, O. Nasisi, B. Kuchen, and F. Reyes. Stable visual servoing of camera-in-hand robotic systems. *IEEE Trans. on Mechatronics*, 5(1):39–48, March 2000.
20. E. Malis. Vision-based control using different cameras for learning the reference image and for servoing. In *IEEE/RSJ International Conference on Intelligent Robots Systems*, volume 3, pages 1428–1433, Maui, Hawaii, November 2001.
21. E. Malis. Visual servoing invariant to changes in camera intrinsic parameters. In *International Conference on Computer Vision*, volume 1, pages 704–709, Vancouver, Canada, July 2001.
22. E. Malis. An unified approach to model-based and model-free visual servoing. In *European Conference on Computer Vision*, volume 4, pages 433–447, Copenhagen, Denmark, May 2002.
23. E. Malis. Vision-based control invariant to camera intrinsic parameters: stability analysis and path tracking. In *IEEE International Conference on Robotics and Automation*, volume 1, Washington, D.C., USA, May 2002.
24. E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *IEEE Transaction on Robotics and Automation*, 18(2):176–186, April 2002.
25. E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):234–246, April 1999.
26. Y. Mezouar and F. Chaumette. Design and tracking of desirable trajectories in the image space by integrating mechanical and visibility constraints. In *IEEE Int. Conference on Robotics and Automation*, Seoul, South Korea, May 2001.
27. G. Morel, J. Szewczyk, S. Boudet, and J. Pot. Explicit incorporation of 2d constraints in vision based control of robot manipulators. In *Proc. ISER'99 : Experimental Robotics IV*, pages 99–108, Sidney, Australia, April 1999.
28. J. G. P. Martinet. Position based visual servoing using a nonlinear approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 531–536, Kyongju, Korea, October 1999.
29. C. Samson, M. Le Borgne, and B. Espiau. *Robot Control: the Task Function Approach*, volume 22 of *Oxford Engineering Science Series*. Clarendon Press, Oxford, UK, 1990.
30. C. Samson, M. Le Borgne, and B. Espiau. *Robot Control: the Task Function Approach*, volume 22 of *Oxford Engineering Science Series*. Clarendon Press, Oxford, UK, 1991.
31. G. W. Stewart and J.-g. Sun. *Matrix perturbation theory*. Computer Science and Science Computing. Harcourt Brace Jovanovich, 1990.
32. M. Sznaier, B. Murphy, and O. Camps. An lpv approach to synthesizing robust active vision systems. In *IEEE Conference on Decision and Control*, pages 2545–2550, Sydney, Australia, December 2000.
33. S. Tarbouriech and P. Soueres. Advanced control strategies for the visual servoing scheme. In *IFAC Symposium on Robot Control, SYROCO*, volume 2, pages 457–462, September 2000.
34. M. Werman, S. Banerjee, S. Dutta Roy, and M. Qiu. Robot localization using uncalibrated camera invariants. In *IEEE Int. Conference on Computer Vision and Pattern Recognition*, volume II, pages 353–359, Fort Collins, CO, June 1999.
35. W. J. Wilson, C. C. W. Hulls, and G. S. Bell. Relative end-effector control using cartesian position-based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5):684–696, October 1996.