Robustness of Image-Based Visual Servoing with Respect to Depth Distribution Errors

Ezio Malis
I.N.R.I.A. Sophia Antipolis
2004, route des Lucioles, France

Patrick Rives
I.N.R.I.A. Sophia Antipolis
2004, route des Lucioles, France

Abstract— This paper concerns the stability analysis of image-based visual servoing methods with respect to uncertainties on the depths of the observed object. In the recent past, research on image-based visual servoing has been concentrated on potential problems of stability and on robustness with respect to camera calibration errors. Only little attention, if any, has been devoted to the robustness of image-based visual servoing to depth estimation errors. It is generally believed that a rough approximation of the depth distribution is sufficient to ensure the stability of the control law. In this paper, we prove that the robustness domain is not so wide and that an extreme care must be taken when approximating the depth distribution.

I. 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. Several vision-based control laws have been proposed in the literature [1]. Contrarily to model-based visual servoing methods, image-based visual servoing does not need the knowledge of the full model of the target [2]. On the other hand, it is necessary to provide some information about the depths of the object in the camera frame. It is generally believed that a rough approximation of the depth distribution is sufficient to ensure the stability of the control law. However, if the environment is completely unknown and the robot is uncalibrated the stability of the visual servoing in the presence of depth estimation errors, can become a serious issue. In the recent past, research on the stability of image-based visual servoing has been concentrated on the solution of convergence problems [3]. Indeed, image based is a local method which, even in the absence of calibration errors, can fail if the initial camera displacement is too big [3]. In order to avoid these potential convergence problems several possible approaches have been proposed. In hybrid approaches, some global information is introduced by estimating the camera displacement between the current and reference views [4] [5] [6]. The rotation of the camera is thus controlled directly in the Cartesian space while some image-based information is used to control the translation. The image-based information used in [4] consists of only one image point. Thus, the method does not need the estimation of the depths of all other points and it is stable for any positive approximation of the depth of point [7]. On the contrary, other hybrid approaches [6] and [5] use all available information in the image and thus they need the estimation of the depth distribution. More recently, a partitioned approach [8] has been proposed in order to avoid the camera displacement reconstruction but the approach is also strongly dependent on the depth distribution estimation. For this reason, the stability analysis of hybrid (except for [7]) and partitioned approaches is as difficult as the stability analysis of the standard image-based visual servoing. 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 [9] [10]. Thus, the error in the image is maintained small at each iteration of the control law. Even if interpolation approaches are an elegant solution to potential convergence problems of the standard image-based visual servoing, a fundamental theoretical question stands: how is the image-based visual servoing robust with respect to calibration errors ? 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] [12] [13], often considering a simplified model for the camera intrinsic parameters but always supposing that the depth distribution was perfectly estimated. In this paper, we investigate the robustness of the image-based visual servoing method with respect to errors on the depth distribution. The analysis proposed in this paper is not limited to purely image-based visual servoing methods but it could also be extended to recent hybrid and partitioned methods which use several image-based features in their control law [6], [5], [8].

II. Theoretical Background

A. Perspective projection

Let $\mathscr{F}_0$ be a frame attached to an object represented by the homogeneous coordinates of a discrete set of n 3D points $\mathscr{X}_i = (X_i, Y_i, Z_i, 1)$ ($i = \{1, 2, ..., n\}$). Let $\mathscr{F}$ 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 $\mathscr{X}_i \in \mathbb{P}^3$ is projected to the point $\mathbf{m}_i \in \mathbb{P}^2$ with normalized homogeneous coordinates:

$$\mathbf{m}_i = \frac{1}{Z_i} \left[\begin{array}{cc} \mathbf{R}_0 & \mathbf{t}_0 \end{array} \right] \mathscr{X}_i = (x_i, y_i, 1) \tag{1}$$

where $\mathbf{R}_0$ and $\mathbf{t}_0$ are respectively the rotation and the translation between frame $\mathscr{F}_0$ and $\mathscr{F}$. Pinhole cameras perform a perspective projection of a 3D point. The information measured by the camera is an image point:

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

where the upper triangular matrix $\mathbf{K}$ contains the camera intrinsic parameters. Using an approximation $\widehat{\mathbf{K}}$ of the camera intrinsic parameters $\mathbf{K}$ and a measured image point $\mathbf{p}_i$ it is possible to compute the corresponding normalized point from equation (2): $\widehat{\mathbf{m}}_i = \widehat{\mathbf{K}}^{-1}\mathbf{p}_i$. Obviously, if the camera intrinsic parameters are perfectly known $\widehat{\mathbf{K}} = \mathbf{K}$ then the normal coordinates are perfectly estimated $\widehat{\mathbf{m}}_i = \mathbf{m}_i$.

B. Image-based visual servoing

Consider the $(2n \times 1)$ vector $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n)$, where $\mathbf{s}_i = (x_i, y_i)$ is the (2×1) vector containing the normalized coordinates extracted from $\mathbf{m}_i$. The derivative of $\mathbf{s}_i$ with respect to time is:

$$\dot{\mathbf{s}}_i = \mathbf{L}_i(Z_i, \mathbf{s}_i)\,\mathbf{v}$$

where $\mathbf{v}$ is the velocity of the camera and $\mathbf{L}_i$ is the (2×6) interaction matrix [2] which can be decomposed into two (2×3) sub-matrices $\mathbf{L}_i = \left[\begin{array}{cc} \mathbf{A}_i & \mathbf{B}_i \end{array} \right]$:

$$\mathbf{A}_i(Z_i, \mathbf{s}_i) = \frac{1}{Z_i} \left[\begin{array}{ccc} -1 & 0 & x_i \\ 0 & -1 & y_i \end{array} \right] \tag{3}$$

$$\mathbf{B}_i(\mathbf{s}_i) = \left[\begin{array}{ccc} x_i y_i & -(1+x_i^2) & y_i \\ (1+y_i^2) & -x_i y_i & -x_i \end{array} \right] \tag{4}$$

If we consider the derivative of vector $\mathbf{s}$ we have:

$$\dot{\mathbf{s}} = \mathbf{L}(\mathbf{z}, \mathbf{s})\,\mathbf{v}$$

where $\mathbf{L} = (\mathbf{L}_1, \mathbf{L}_2, ..., \mathbf{L}_n)$ is the $(2n \times 6)$ interaction matrix. Again, $\mathbf{L}$ can be decomposed into two $(2n \times 3)$ sub-matrices:

$$\mathbf{L}(\mathbf{z}, \mathbf{s}) = \left[\begin{array}{cc} \mathbf{A}(\mathbf{z}, \mathbf{s}) & \mathbf{B}(\mathbf{s}) \end{array} \right]$$

where $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n)$ and $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_n)$. Due to the form of matrix $\mathbf{A}(\mathbf{z}, \mathbf{s})$ we also have:

$$\mathbf{A}(\mathbf{z}, \mathbf{s}) = \mathbf{D}(\mathbf{z})\,\mathbf{C}(\mathbf{s})$$

where:

$$\mathbf{D}(\mathbf{z}) = \mathrm{diag}\left(\frac{1}{Z_1}, \frac{1}{Z_1}, \frac{1}{Z_2}, \frac{1}{Z_2}, ..., \frac{1}{Z_n}, \frac{1}{Z_n} \right)$$

is a $(2n \times 2n)$ diagonal matrix containing the depth distribution $\mathbf{z}$. Consider the following task function [2]:

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

where $\widehat{\mathbf{L}}^+$ is the pseudo-inverse of an approximation of the true $(2n \times 6)$ interaction matrix. In [2], the matrix $\widehat{\mathbf{L}}^+$ is supposed to be constant while in this paper we consider the most general case when the matrix is not constant. In that case, the derivative of the task function is:

$$\dot{\mathbf{e}} = \frac{d\widehat{\mathbf{L}}^+}{dt}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^+\dot{\mathbf{s}} = (\mathbf{O}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^+\mathbf{L})\,\mathbf{v} \tag{5}$$

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 control law:

$$\mathbf{v} = -\lambda\,\mathbf{e} \tag{6}$$

In order to compute the control law it is necessary to provide the approximated interaction matrix $\widehat{\mathbf{L}}$.

III. STABILITY ANALYSIS

Plugging equation (6) into equation (5), we obtain the following closed-loop equation:

$$\dot{\mathbf{e}} = -\lambda\,(\mathbf{O}(\mathbf{s} - \mathbf{s}^*) + \widehat{\mathbf{L}}^+\mathbf{L})\mathbf{e} \tag{7}$$

It is well know from control theory that the non-linear system (7) is locally asymptotically stable in a neighborhood of $\mathbf{s} = \mathbf{s}^*$ if and only if the linearized system is stable:

$$\dot{\mathbf{e}} = \lambda\mathbf{Q}\mathbf{e} \tag{8}$$

where $\mathbf{Q} = -\widehat{\mathbf{L}}^+\mathbf{L}|_{\mathbf{s}=\mathbf{s}^*}$. The linear system (8) is asymptotically stable *if and only if* $\mathbf{Q}$ has eigenvalues with negative real part:

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

The matrix depends $\mathbf{Q} = \mathbf{Q}(\widehat{\mathbf{K}}, \mathbf{K}, \widehat{\mathbf{z}}, \mathbf{z})$ on two set of unknown parameters. Obviously, if $\mathbf{K} = \widehat{\mathbf{K}}$ and $\widehat{\mathbf{z}} = \mathbf{z}$ then $\mathbf{Q} = \mathbf{I}$ and the system is stable. The objective of the robustness analysis is to known if the system is stable in the presence of unavoidable calibration errors. Note that, an adaptive estimation of the depth distribution is possible only supposing that the robot is calibrated and in the presence of enough disparity in the image.

A. Known camera intrinsic parameters

Let us suppose that the camera parameters are perfectly known (i.e. $\widehat{\mathbf{K}} = \mathbf{K}$). Thus, the normalized points are perfectly estimated $\widehat{\mathbf{s}} = \mathbf{s}$ and the uncertainties on the estimated interaction matrix only depends on the depth distribution $\widehat{\mathbf{z}}$:

$$\widehat{\mathbf{L}}(\widehat{\mathbf{z}}, \mathbf{s}) = \left[\begin{array}{cc} \mathbf{A}(\widehat{\mathbf{z}}, \mathbf{s}) & \mathbf{B}(\mathbf{s}) \end{array} \right]$$

It is easy to verify that the estimated sub-matrix $\mathbf{A}(\widehat{\mathbf{z}}, \mathbf{s})$ can be written as a function of the true sub-matrix $\mathbf{A}(\mathbf{z}, \mathbf{s})$:

$$\mathbf{A}(\widehat{\mathbf{z}}, \mathbf{s}) = \mathbf{D}(\widehat{\mathbf{z}})\,\mathbf{C}(\mathbf{s}) = \Gamma^{-1}(\widehat{\mathbf{z}}, \mathbf{z})\mathbf{A}(\mathbf{z}, \mathbf{s}) \tag{9}$$

where, setting $\gamma_i = \widehat{Z}_i/Z_i$ the ratio between the estimated and true depths, the diagonal matrix Γ is:

$$\Gamma = \mathbf{D}(\mathbf{z})\mathbf{D}^{-1}(\widehat{\mathbf{z}}) = \mathrm{diag}\left(\gamma_1, \gamma_1, \gamma_2, \gamma_2, ..., \gamma_n, \gamma_n \right)$$

From equation (9) one can deduce that:

$$\mathbf{L}(\mathbf{z},\mathbf{s}) = \begin{bmatrix} \Gamma \mathbf{A}(\widehat{\mathbf{z}},\mathbf{s}) & \mathbf{B}(\mathbf{s}) \end{bmatrix}$$

Setting $\Delta = \Gamma - \mathbf{I}$ one can deduce that:

$$\mathbf{L} = \widehat{\mathbf{L}} + \Delta \begin{bmatrix} \mathbf{A}(\widehat{\mathbf{z}},\mathbf{s}) & \mathbf{0} \end{bmatrix}$$

Setting $\widehat{\mathbf{A}} = \mathbf{A}(\widehat{\mathbf{z}},\mathbf{s})$, the matrix $\mathbf{Q}$ is:

$$\mathbf{Q} = -\widehat{\mathbf{L}}^{+}\mathbf{L} = -\mathbf{I} - \begin{bmatrix} \widehat{\mathbf{L}}^{+}\Delta\widehat{\mathbf{A}} & \mathbf{0} \end{bmatrix}$$

If $\widehat{\mathbf{L}}$ is full rank, the pseudo-inverse of the matrix can be written as:

$$\widehat{\mathbf{L}}^{+} = \begin{bmatrix} \widehat{\mathbf{A}}^{\natural} \\ \widehat{\mathbf{B}}^{\natural} \end{bmatrix}$$

where $\widehat{\mathbf{A}}^{\natural}$ is a generalized inverse of $\widehat{\mathbf{A}}$ (i.e. $\widehat{\mathbf{A}}^{\natural}\widehat{\mathbf{A}} = \mathbf{I}$) and $\widehat{\mathbf{B}}^{\natural}$ is a generalized inverse of $\widehat{\mathbf{B}}$ (i.e. $\widehat{\mathbf{B}}^{\natural}\widehat{\mathbf{B}} = \mathbf{I}$). Note also that $\widehat{\mathbf{A}}^{\natural}\widehat{\mathbf{B}} = 0$ and $\widehat{\mathbf{B}}^{\natural}\widehat{\mathbf{A}} = 0$. Matrix $\mathbf{Q}$ can be rewritten as:

$$\mathbf{Q} = -\mathbf{I} - \begin{bmatrix} \widehat{\mathbf{A}}^{\natural} \\ \widehat{\mathbf{B}}^{\natural} \end{bmatrix} \begin{bmatrix} \Delta\widehat{\mathbf{A}} & \mathbf{0} \end{bmatrix} = - \begin{bmatrix} \mathbf{I} + \widehat{\mathbf{A}}^{\natural}\Delta\widehat{\mathbf{A}} & \mathbf{0} \\ \widehat{\mathbf{B}}^{\natural}\Delta\widehat{\mathbf{A}} & \mathbf{I} \end{bmatrix}$$

Setting again $\Delta = \Gamma - \mathbf{I}$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} -\widehat{\mathbf{A}}^{\natural}\Gamma\widehat{\mathbf{A}} & \mathbf{0} \\ -\widehat{\mathbf{B}}^{\natural}\Gamma\widehat{\mathbf{A}} & -\mathbf{I} \end{bmatrix}$$

Thus, the closed-loop matrix is block lower triangular. In this case, it is well known that the eigenvalues of $\mathbf{Q}$ are the eigenvalues of the two (3×3) matrices $\mathbf{Q}_{11}$ and $\mathbf{Q}_{22}$. Since $\mathbf{Q}_{22} = -\mathbf{I}$ its eigenvalues are negative for any choice of the depth distribution. The analysis is limited to the eigenvalues of the following matrix:

$$\mathbf{Q}_{11} = -\widehat{\mathbf{A}}^{\natural}\Gamma\widehat{\mathbf{A}} = \sum_{i=1}^{n} \gamma_i \widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i$$

where $\widehat{\mathbf{A}}_i^{\natural}$ are sub-matrices of matrix $\widehat{\mathbf{A}}$. Note that:

$$\widehat{\mathbf{A}}^{\natural}\widehat{\mathbf{A}} = \sum_{i=1}^{n} \widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i = \mathbf{I}$$

The first important results of the analysis is that *the depth distribution can be estimated up to a positive scalar factor.* The scalar factor only influence the performance of the servoing but not its stability since it does not change the sign of the eigenvalues. Thus, without loss of generality we can factor $\gamma_j > 0$ from the sum:

$$\mathbf{Q}_{11} = -\gamma_j \sum_{i=1}^{n} \psi_i \widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i = \gamma_j \mathbf{F}$$

where $\psi_i = \gamma_i/\gamma_j$ and, obviously, $\psi_j = 1$. We will see in section III-A.2 how it is possible to select the best j. Since $\gamma_j > 0$, $\mathbf{Q}_{11}$ is stable if and only if $\mathbf{F}$ is stable. Therefore, we can focus on the stability of $\mathbf{F}$.

1) Necessary and sufficient conditions: The eigenvalues of $\mathbf{F}$ are the roots of the characteristic polynomial:

$$\lambda^3 - \text{tr}(\mathbf{F})\lambda^2 + \frac{1}{2}(\text{tr}(\mathbf{F})^2 - \text{tr}(\mathbf{F}^2))\lambda - \det(\mathbf{F}) = 0$$

where tr and det are respectively the trace and the determinant of a matrix. The necessary and sufficient conditions for the roots of the polynomial to have negative real part are obtained from the Routh-Hurwitz Theorem:

$$\begin{aligned} \text{tr}(\mathbf{F}) &< 0 \\ \text{tr}(\mathbf{F}^2) - \text{tr}(\mathbf{F})^2 &< 0 \\ \det(\mathbf{F}) &< 0 \\ \text{tr}(\mathbf{F})(\text{tr}(\mathbf{F})^2 - \text{tr}(\mathbf{F}^2)) - 2\det(\mathbf{F}) &< 0 \end{aligned}$$

The necessary and sufficient conditions can be used to test the stability of the servoing and to obtain the robustness domain (see for example the simulations in section IV-A). However, for a large number of parameters the computation time can be high. In some cases, it is preferable to have a simple test in order to know, given a bound on the precision of depths estimates $|\psi_i| \leq \overline{\psi}_i$, if the eigenvalues are negative. In the next section, we present simple sufficient conditions in order to obtain an approximation of the robustness domain.

2) Sufficient conditions: Since $\psi_j = 1$, we can rewrite the (3×3) matrix $\mathbf{F}$ as:

$$\mathbf{F} = -\widehat{\mathbf{A}}_j^{\natural}\widehat{\mathbf{A}}_j - \sum_{i=1, i\neq j}^{n} \psi_i \widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i$$

from equation $\widehat{\mathbf{A}}_j^{\natural}\widehat{\mathbf{A}}_j = \mathbf{I} - \sum_{i=1, i\neq j}^{n}\widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i$, thus:

$$\mathbf{F} = -\mathbf{I} - \sum_{i=1, i\neq j}^{n} \delta_i \widehat{\mathbf{A}}_i^{\natural}\widehat{\mathbf{A}}_i = -\mathbf{I} + \mathbf{E}(\delta)$$

where $\delta = (\delta_1, \delta_2, ..., \delta_m)$ and $\delta_i = \psi_i - 1$. Matrix $\mathbf{F}$ can be regarded as a perturbation of the matrix $-\mathbf{I}$, where $\mathbf{E}(\delta)$ is the perturbation matrix. Let us define the SPECTRAL VARIATION of a matrix $\tilde{\mathbf{M}}$ with respect to a matrix $\mathbf{M}$ as [14]:

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

The Bauer-Fike theorem [14] states that:

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

In our case, applying the Bauer-Fike theorem to the spectral variation $\mathbf{F}$ with respect to $-\mathbf{I}$ we obtain:

$$\text{sv}(\mathbf{F}) = \max_i |\tilde{\lambda}_i + 1| \leq \|\mathbf{E}(\delta)\|$$

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

$$\max_i |\tilde{\lambda}_i + 1| < 1 \tag{10}$$

which implies $\tilde{\lambda}_i < 0$. From the definition of spectral variation, all others eigenvalues λ_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 $\mathbf{E}(\delta) = -\sum_{i=1,i\neq j}^{n} \delta_i \widehat{\mathbf{A}}_i^\natural \widehat{\mathbf{A}}_i$:

$$\|\mathbf{E}\| \leq \sum_{i=1,i\neq j}^{n} |\delta_i| \|\widehat{\mathbf{A}}_i^\natural \widehat{\mathbf{A}}_i\|$$

setting $\mu_i = \|\widehat{\mathbf{A}}_i^\natural \widehat{\mathbf{A}}_i\| > 0$, the condition (10) can be imposed by bounding the previous inequality:

$$\sum_{i=1,i\neq j}^{n} \mu_i |\delta_i| < 1 \qquad (11)$$

In the inequality, each error $|\delta_i|$ is weighted by the scalars μ_i. The smaller is μ_i the larger can be $|\delta_i|$. Thus, the best choice for the point γ_j is $\mu_j = \max_k \mu_k$. Inequality (11) define a polygonal region whose axis are weighted by the scalars μ_i. The volume of the region $V = \prod_{i=1,i\neq j}^{n} \mu_i$ gives a measure of the robustness domain. If we suppose that the precision of measurement is the same for all points $|\delta_i| \leq \delta$ then:

$$\delta < 1/\sum_{i=1,i\neq j}^{n} \mu_i \qquad (12)$$

This is a very simple test for the local stability.

B. Unknown camera intrinsic parameters

If the camera intrinsics parameters are unknown the analysis is not simplified as in the previous case. The matrix $\mathbf{Q}$ is not upper triangular any more and depends on the estimated camera intrinsics parameters $\widehat{\mathbf{K}}$ [11]. Thus, to test the stability of the system one must consider all six eigenvalues of the matrix. Consequently, the stability regions are reduced and a bigger precision in the depth estimation is required.

IV. SIMULATION RESULTS

The stability results obtained in the previous section have been tested with simulations. Three set of tests have been carried out. In the first set, the target is planar. In this case, whatever is the number of points on the plane, the stability analysis only depends on the estimated normal to the plane. Thus, from the necessary and sufficient conditions we obtain the *exact* robustness domain. In the second set, we consider a 3D object. In that case, the sufficient conditions provide an *approximate* robustness domain. Finally, in the last simulations we show the influence of the errors of the depth distribution on the visual servoing.

A. Planar objects

When the object is planar, the depths are related to the normal $\mathbf{n}$ to the plane and proportional to the distance d of the plane from the center of projection:

$$Z_i = d/\mathbf{n}^\top \mathbf{m}$$

where $\mathbf{n}$ is a unit vector which is a function of two parameters $\mathbf{n}(\theta,\phi) = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$. The estimated depth $\widehat{Z}_i$ can be obtained using an approximation of $\widehat{\mathbf{n}}(\widehat{\theta}, \widehat{\phi})$ and $\widehat{d}$:

$$\widehat{Z}_i = d/\widehat{\mathbf{n}}^\top \mathbf{m}$$

then:

$$\gamma_i = \frac{\widehat{Z}_i}{Z_i} = \frac{\widehat{d}}{d}\frac{\mathbf{n}^\top \mathbf{m}_i}{\widehat{\mathbf{n}}^\top \mathbf{m}_i} \quad \text{and} \quad \psi_i = \frac{\gamma_i}{\gamma_j} = \frac{\mathbf{n}^\top \mathbf{m}_i}{\widehat{\mathbf{n}}^\top \mathbf{m}_i}\frac{\widehat{\mathbf{n}}^\top \mathbf{m}_i}{\mathbf{n}^\top \mathbf{m}_i}$$

As expected, the stability of the visual servoing does not depends on $\widehat{d}$ but only on $\widehat{\mathbf{n}}$. Figure 1 show the stability regions as a function of $(\widehat{\theta}, \widehat{\phi})$ for an increasing number of points on the same plane. The true normal is $\mathbf{n} = (0.5, 0, 0.866)$ (i.e. $\theta = 0$ and $\phi = \pi/6$). In the green region all the eigenvalues are negatives, the system is locally asymptotically stable. In the red region at least one eigenvalue is positive, and the system is locally unstable. Finally, the normals obtained in the blue region are discarded since we obtain at least a negative depth, which is impossible. When considering 3 image points (see Figure 1(a)), the corresponding stable region is not so wide. Note that, adding a point inside the triangle defined by the others points (see Figure 1(c)) only slightly modifies the stability region (compare Figure 1(b) and (d)). Thus, when learning the reference image, one can think that it is probably better to chose points spread in the image. Unfortunately, the stability analysis shows that it is not always true. Indeed, if we add 4 more points as in Figure 1(e), the stability region in green is even reduced (see Figure 1(f)). Note that, if we have absolutely no idea on the 3D position of the plane, a simple guess $\widehat{\mathbf{n}} = (0, 0, 1)$ makes the visual servoing unstable. On the other hand, if many points are well distributed in all the image as in Figure 1(g) the stability region considerably increase (see Figure 1(h)). However, even in this very favorable case, there exist an important red instability region.

B. Non-planar objects

When the target is non-planar, it is easier to use the sufficient condition. In the simulation, we show the stability regions for 3 and 4 points since they can be represented in a plot. In the first case, we find $\mu_1 = 0.1861$ and $\mu_2 = 0.2188$ see Figure 2(a)). After adding a point not on the plane, we find $\mu_1 = 0.1861$, $\mu_2 = 0.2188$ and $\mu_3 = 0.2235$ see Figure 2(b)). For higher dimensional problems, the volume of the convex polyhedron gives an

idea of the precision required in the measurement of the depth distribution.
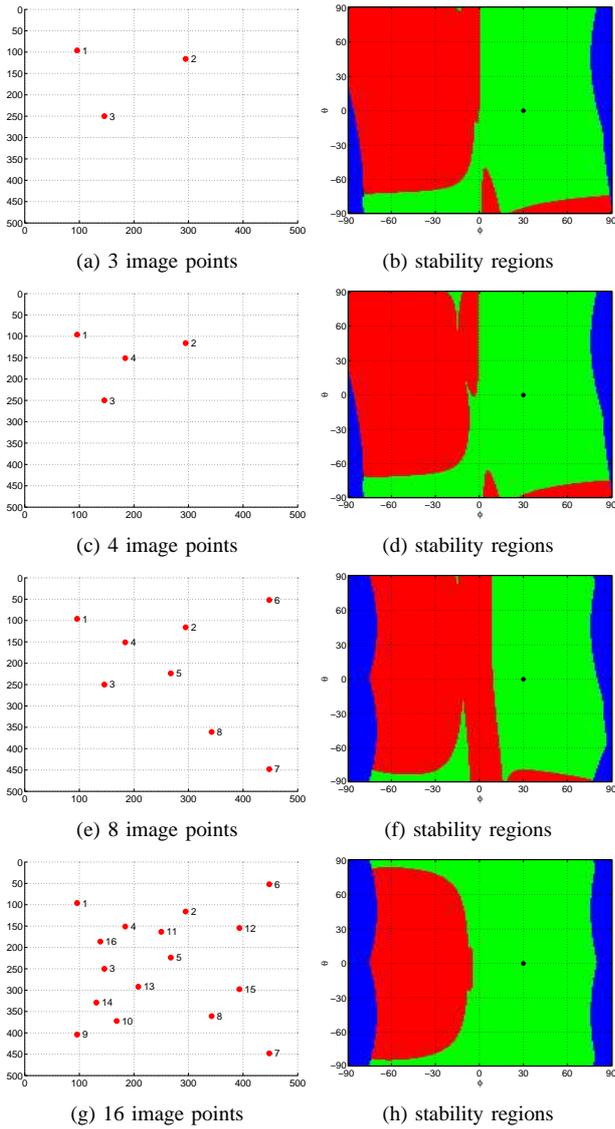

(a) 3 image points   (b) stability regions

(c) 4 image points   (d) stability regions

(e) 8 image points   (f) stability regions

(g) 16 image points   (h) stability regions

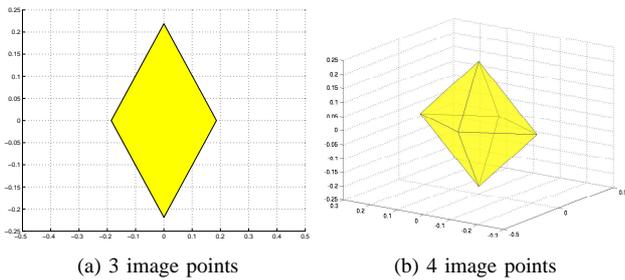Fig. 1. Stability regions for a planar object.

(a) 3 image points   (b) 4 image points

Fig. 2. Stability regions for a non planar object.

C. Visual servoing

We simulated several examples to show the usefulness of the theoretical analysis proposed in the paper. For all examples, the initial camera displacement is very small: $t = -\begin{bmatrix} 0,001 & 0,001 & 0,001 \end{bmatrix}$ meters and $r = \begin{bmatrix} 0,48 & 0,96 & 1,44 \end{bmatrix}$ degrees. Firstly, we verified that the visual servoing is locally stable when the sufficient conditions are satisfied. From equation (12) we find that if $|\delta_i| < \delta = 0.06$ then the visual servoing is stable. The true depths are $\mathbf{z}^* = (0.91, 0.83, 0.87, 0.86, 0.91, 0.89)$ m while the estimated depth is an average of the true depths $\widehat{\mathbf{z}}^* = (0.88, 0.88, 0.88, 0.88, 0.88, 0.88)$ m. Note that the maximum error is 6 % of the true depth. From the data given above, we find $\delta_1 = 0.047$, $\delta_2 = 0.052$, $\delta_4 = 0.017$, $\delta_5 = 0.039$ and $\delta_6 = 0.021$. Thus, the sufficient condition is verified and the visual servoing is stable. However, it must be noticed that if an error of 10 % on the intrinsic parameters is added, the visual servoing is not stable any more. This prove that, in the presence of camera calibration errors the stability region is reduced. In the first simulation, we show that if the necessary and sufficient conditions are not verified the visual servoing is unstable. The true depth distribution is $\mathbf{z} = (0.94, 0.85, 0.94, 0.85, 0.95, 0.94)$ m, while the estimated depth distribution is $\widehat{\mathbf{z}} = (0.91, 0.91, 0.91, 0.91, 0.91, 0.91)$ m. Despite the maximal error on the estimated depths is only 8% of the true depth, one eigenvalue is positive. Thus, even starting very close to the reference position, after iteration 200 the translation and rotation errors start to grow (see Figures 3(c) and 3(d)). In the beginning, the control law seems to be stable since the others dominant eigenvalues have negative real part. Note that, in this case the sufficient condition is also not satisfied and we know in advance that the servoing will be unstable. Thus, it is useful to know if the precision on the depth distribution is good enough before starting the servoing.

(a) Translation velocity   (b) Rotation velocity
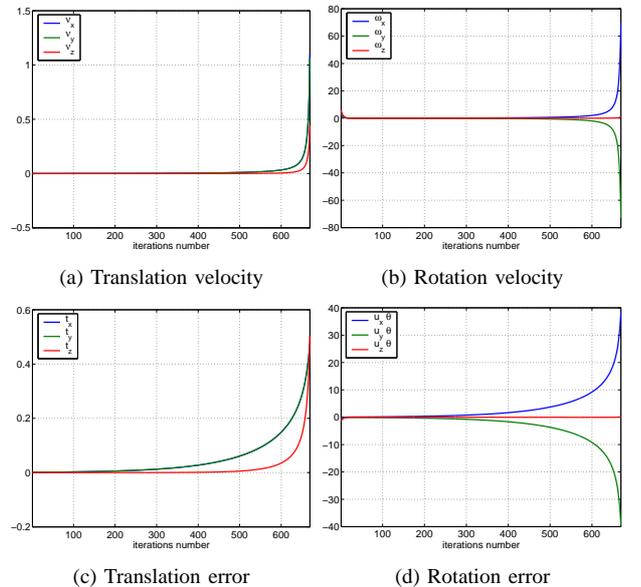
(c) Translation error   (d) Rotation error

Fig. 3. Unstable image-based visual servoing.

The second simulation is similar to the first one but with the addition of a Gaussian noise ($\sigma = 1$ pixel) in the image. The noise in the image accelerate the divergence of the servoing (see Figure 4).
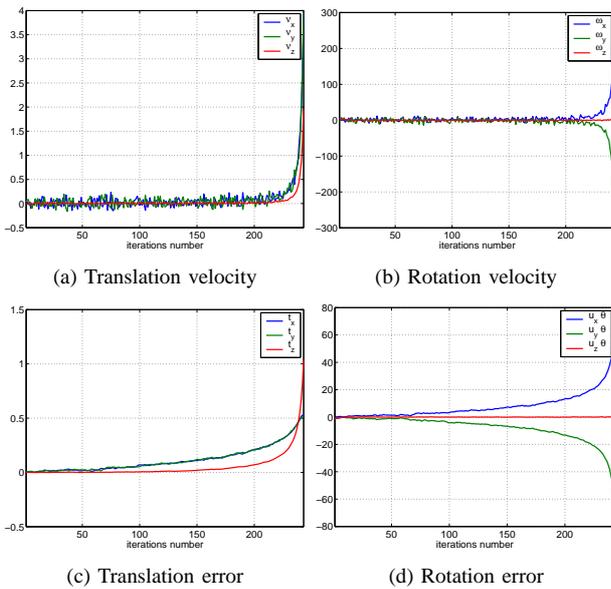

(a) Translation velocity   (b) Rotation velocity

(c) Translation error   (d) Rotation error

Fig. 4. Unstable image-based visual servoing with noise.

V. CONCLUSION

In this paper, we have shown that extreme care must be taken when approximating the depth distribution of a target for image-based visual servoing. Indeed, the stability region in the presence of errors on the depth distribution is not very large. As a consequence, if the target geometry is completely unknown it is necessary to accurately estimate the depth. Future work will be devoted to the off-line estimation of the depth distribution with a precision bounded by those obtained from the sufficient condition proposed in this paper.

ACKNOWLEDGMENTS

VI. REFERENCES

[1] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Transaction on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, October 1996.

[2] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Transaction on Robotics and Automation*, vol. 8, no. 3, pp. 313–326, June 1992.

[3] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The confluence of vision and control*, D. Kriegman, G. Hager and A. Morse Editors, vol. 237 of *LNCIS Series*, pp. 66–78. Springer Verlag, 1998.

[4] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 d visual servoing," *IEEE Transaction on Robotics and Automation*, vol. 15, no. 2, pp. 234–246, April 1999.

[5] 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*, Sidney, Australia, April 1999, pp. 99–108.

[6] K. Deguchi, "Optimal motion control for image-based visual servoing by decoupling translation and rotation," in *IEEE International Conference on Intelligent Robots and Systems*, October 1998, vol. 2, pp. 705–711.

[7] 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*, vol. 18, no. 2, pp. 176–186, April 2002.

[8] P. Corke and S. Hutchinson, "A new partitioned approach to image-based visual servo control," *IEEE Transaction on Robotics and Automation*, vol. 14, no. 4, pp. 507–515, Aug. 2001.

[9] K. Hashimoto and T. Noritsugu, "Enlargement of stable region in visual servo," in *IEEE Conference on Decision and Control*, Sydney, Australia, December 2000, pp. 3927–3932.

[10] Y. Mezouar and F. Chaumette, "Design and tracking of desirable trajectories in the image space by integrating mechanical and visibility constraints," in *IEEE International Conference on Robotics and Automation*, Seoul, South Korea, May 2001.

[11] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," in *3rd International Symposium on Experimental Robotics*, Kyoto, Japan, October 1993.

[12] C. C. Cheah, S. Kawamura, and S. Arimoto, "Feedback control for robotic manipulator with uncertain kinematics and dynamics," in *IEEE International Conference on Robotics and Automation*, Belgium, May 1998, vol. 4, pp. 3607–3612.

[13] L. Deng, F. Janabi-Sharifi, and W. J. Wilson, "Stability and robustness of visual servoing methods," in *IEEE International Conference on Robotics and Automation*, Washington D.C., May 2002, vol. 2, pp. 1605–1609.

[14] G. W. Stewart and Ji-guang Sun, *Matrix perturbation theory*, Computer Science and Science Computing. Harcourt Brace Jovanovich, 1990.