Proceedings of 2004 IEEE/RSJ International Conference on
Intelligent Robots and Systems
September 28 - October 2, 2004, Sendai, Japan

Robustness of Central Catadioptric Image-Based Visual Servoing to Uncertainties on 3D Parameters

Youcef Mezouar
LASMEA, Université B. Pascal
63177 Aubière Cedex, France
Youcef.Mezouar@lasmea.univ-bpclermont.fr

Ezio Malis
INRIA, Sophia Antipolis
2004, route des Lucioles, France
Ezio.Malis@sophia.inria.fr

Abstract— **This paper concerns the stability analysis of image-based visual servoing methods with respect to uncertainties on the 3D parameters introduced in the central catadioptric interaction matrix. Motivated by the growing interest for omnidirectional sensors on robotic applications and particularly on vision-based control, we extend recent results obtained for conventional cameras to the entire class of central catadioptric systems. In this paper, it is shown that with such sensors extreme care must be taken when approximating 3D parameters to ensure stability of the image-based control law.**

I. Introduction

In the last years, the use of visual observations to control the motions of robots has been extensively studied (approach referred in the literature as visual servoing). This is known to be a very efficient and flexible method for applications such as the positioning of a robot and the tracking of objects using the visual informations provided by an in-hand camera.

Three groups of vision-based control laws have been proposed in the literature namely position-based, image-based and hybrid-based control [9], [11], [12]. Contrarily to model-based visual servoing methods, image based visual servoing does not need the knowledge of the full model of the target. On the other hand, it is necessary to provide some information about depths of the object in the camera frame. Until recently, it was generally believed that a rough approximation of the depth distribution was sufficient to ensure the stability of the control law and most of the effort has been concentrated on the solution of convergence problems. Indeed image-based visual servoing is a local control solution which, even under perfect modeling can fail if the initial camera displacement is too big. Some methods have been proposed to address this issue based on the choice of features sharing good properties [18] or based on path planning [15]. Nevertheless, if the environment is completely unknown and the robot is uncalibrated the stability of the visual servoing in presence of depth estimation errors, can become a serious issue. In [13], the stability of image-based visual servoing with respect to depth distributions errors has been studied. It has been shown that the robustness domain is not so wide when conventional camera is used. In this paper, we extend the stability analysis to the entire class of central catadioptric cameras which includes the conventional ones. This work is motivated by the growing interest for omnidirectional sensors in robotics application. Indeed, conventional cameras suffer from restricted field of view. Many applications in vision-based robotics, such as mobile robot localisation [4] and navigation [20], can benefit from panoramic field of view provided by omnidirectional cameras. Clearly, visual servoing applications can also benefit from cameras with a wide field of view since such methods make assumptions on the link between the initial, current and desired images. They require correspondences between the visual features extracted from the initial image with those obtained from the desired one. These features are then tracked during the camera (and/or the object) motion. If these steps fail the visually based robotic task can not be achieved [6]. Some methods have been proposed to resolve this deficiency based on path planning [14], switching control [7], zoom adjustment [17], geometrical and topological considerations [8]. However, such strategies are sometimes delicate to adapt to generic setup. Omnidirectional cameras naturally overcome this problem. As a consequence, such sensors has been successfully integrated as part of a closed loop feedback control system [5], [16], [19], [2].

In the literature, there have been several methods proposed for increasing the field of view of cameras systems [3]. One effective way is to combine mirrors with conventional imaging system. The obtained sensors are referred as catadioptric imaging systems. The resulting imaging systems have been termed central catadioptric when a single projection center describes the world-image mapping. From a practical view point, a single center of projection is a desirable property for an imaging system [1]. Baker and Nayar in [1] derive the entire class of catadioptric systems with a single viewpoint.

In this paper, we study the robustness of image-based methods with respect to errors on the 3D parameters introduced in the interaction matrix when a central catadioptric camera is used as sensor.

II. Theoretical background

In this section, we describe the projection model for central catadioptric cameras and then we focus on eye-in-hand image-based visual servoing methods.

0-7803-8463-6/04/$20.00 ©2004 IEEE

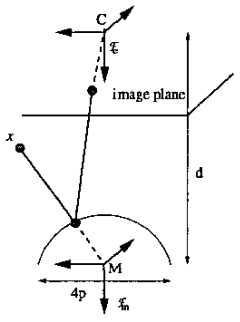


Fig. 1. Generic camera model

A. General camera model

As already mentioned in the introduction, a single center of projection is a desirable property for an imaging system. A single center implies that a point in 3D space projects into a single point in the image plane. Conventional perspective cameras are single view point sensors. As shown in [1], a central catadioptric system can be built by combining an hyperbolic, elliptical or planar mirror with a perspective camera and a parabolic mirror with an orthographic camera. To simplify notations conventional perspective cameras will be embedded in the set of central catadioptric cameras. In [10], a unifying theory for central panoramic systems is presented. According to this generic model, all central panoramic cameras can be modeled by a central projection onto a sphere followed by a central projection onto the image plane (see Fig. 1). This generic model can be parametrized by the couple (ξ, φ) (see Tab.I and refer to [2]) describing the type of sensor and the shape of the mirror. Setting $\xi = 0$, the general projection model becomes the well known perspective projection model.

camera	Mirror surface	ξ	φ
Parabolic	$z = \frac{x^2+y^2}{2a_p} - \frac{a_p}{2}$	1	$1+2p$
Hyperbolic	$\frac{(z+\frac{d}{2})^2}{a_h^2} - \frac{x^2+y^2}{b_h^2} = 1$	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d+2p}{\sqrt{d^2+4p^2}}$
Elliptical	$\frac{(z+\frac{d}{2})^2}{a_e^2} + \frac{x^2+y^2}{b_e^2} = 1$	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d-2p}{\sqrt{d^2+4p^2}}$
Planar	$z = \frac{d}{2}$	0	1
conventional	none	0	1

TABLE I

CENTRAL CATADIOPTRIC CAMERAS DESCRIPTION:
a_p, a_h, b_h, a_e, b_e depend only of the mirror intrinsic parameters d and p

Let $\mathcal{F}_c$ and $\mathcal{F}_m$ be the frames attached to the conventional camera and to the mirror respectively. Suppose that $\mathcal{F}_c$ and $\mathcal{F}_m$ are related by a translation along the Z-axis. The centers C and M of $\mathcal{F}_c$ and $\mathcal{F}_m$ will be called optical center and principal projection center respectively. Let $\mathcal{X}$ be a 3D point with coordinates $\mathbf{X} = [X\ Y\ Z]^T$ with respect to $\mathcal{F}_m$. After setting $\rho = \sqrt{X^2+Y^2+Z^2}$, let $\mathbf{m} = (x, y, 1)$ be the point (in normalized homogeneous coordinated) projected into a virtual plane according to the generic projection model [10]:

$$\mathbf{m} = \left(\frac{X}{Z+\xi\rho}, \frac{Y}{Z+\xi\rho}, 1\right) \tag{1}$$

Taking into account the intrinsic camera and mirror parameters, the 3D point $\mathcal{X}$ is projected in the real image plane to a point with homogeneous pixel coordinates $\mathbf{p} = (u, v, 1)$:

$$\mathbf{p} = \mathbf{KMm} \tag{2}$$

where the upper triangular matrix $\mathbf{K}$ contains the camera intrinsic parameters, and the diagonal matrix $\mathbf{M}$ the mirror intrinsic parameters:

$$\mathbf{M} = \begin{bmatrix} \varphi-\xi & 0 & 0 \\ 0 & \varphi-\xi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Notice that a catadioptric sensor with a planar mirror is equivalent to a conventional perspective camera. Using approximations $\widehat{\mathbf{K}}$ and $\widehat{\mathbf{M}}$ of the camera and mirror intrinsic parameters $\mathbf{K}$ and $\mathbf{M}$ and a measured image point $\mathbf{p}$ it is possible to compute the corresponding normalized point from equation (2): $\widehat{\mathbf{m}} = \widehat{\mathbf{M}}^{-1}\widehat{\mathbf{K}}^{-1}\mathbf{p}$. Obviously if the camera and mirror intrinsic parameters are perfectly known $\widehat{\mathbf{K}} = \mathbf{K}$ and $\widehat{\mathbf{M}} = \mathbf{M}$ the normalized coordinates are perfectly estimated $\widehat{\mathbf{m}} = \mathbf{m}$.

B. Interaction matrix of central catadioptric cameras

Consider a 3D point $\mathcal{X}_i$ with coordinates $\mathbf{X}_i = (X_i, Y_i, Z_i)$ with respect to $\mathcal{F}_m$ and its normalized image coordinates extracted from $\mathbf{m}$: $\mathbf{s}_i = (x_i, y_i)$. The derivative of $\mathbf{s}_i$ with respect to time is:

$$\dot{\mathbf{s}}_i = \mathbf{L}_i\mathbf{v}$$

where $\mathbf{v}$ is the velocity of the camera and $\mathbf{L}_i$ is the interaction matrix. The interaction matrix can be derived by differentiating the function $\mathbf{f}(\mathbf{X})$ with respect to the camera pose evaluated at the origin or by following [2]. $\mathbf{L}_i$ can be decomposed into two sub-matrices $\mathbf{L}_i = [\mathbf{A}_i\ \mathbf{B}_i]$ with :

$$\mathbf{A}_i = \begin{bmatrix} -\frac{1+x_i^2(1-\xi(\sigma_i+\xi))+y_i^2}{\rho_i(\sigma_i+\xi)} & \frac{\xi x_i y_i}{\rho_i} & \frac{x_i\sigma_i}{\rho_i} \\ \frac{\xi x_i y_i}{\rho_i} & -\frac{1+x_i^2+y_i^2(1-\xi(\sigma_i+\xi))}{\rho_i(\sigma_i+\xi)} & \frac{y_i\sigma_i}{\rho_i} \end{bmatrix} \tag{3}$$

$$\mathbf{B}_i = \begin{bmatrix} x_i y_i & -\frac{(1+x_i^2)\sigma_i-\xi y_i^2}{\sigma_i+\xi} & y_i \\ \frac{(1+y_i^2)\sigma_i-\xi x_i^2}{\sigma_i+\xi} & -x_i y_i & -x_i \end{bmatrix} \tag{4}$$

where $\sigma_i = \sqrt{1+(1-\xi^2)(x_i^2+y_i^2)}$. Equations (3) and (4) present the general central catadioptric interaction matrix as a function of image coordinates $\mathbf{s}_i$, the distance $\rho_i = \sqrt{X_i^2+Y_i^2+Z_i^2}$ and sensor parameter ξ. Notice that if $\xi = 0$, then the interaction matrix $\mathbf{L}_i$ is the well known interaction matrix for conventional perspective cameras.

C. Image-based visual servoing

The goal of imaged-based visual servoing is to position a robot by controlling the current position of the robot such that the current measured image features $\mathbf{s}$ reach their reference $\mathbf{s}^*$. Consider the following task function [9]:

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

where $\widehat{\mathbf{L}}^+$ is the pseudo-inverse of an approximation of the true (2n×6) interaction matrix. In [9], 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:

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

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:

$$\mathbf{v} = -\lambda\,\mathbf{e} \quad (6)$$

In order to compute the control law it is necessary to provide the approximated interaction matrix $\widehat{\mathbf{L}}$. Plugging equation (6) into equation (5), we obtain the following closed-loop equation:

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

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:

$$\dot{\mathbf{e}} = \lambda\mathbf{Q}\mathbf{e} \quad (8)$$

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:

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

The matrix $\mathbf{Q} = \mathbf{Q}(\widehat{\mathbf{K}}, \mathbf{K}, \widehat{\mathbf{M}}, \mathbf{m}, \widehat{\mathbf{z}}, \mathbf{z})$ depends on two set of unknown parameters. Obviously, if $\mathbf{K} = \widehat{\mathbf{K}}$, $\widehat{\mathbf{M}} = \mathbf{m}$ 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.

III. Stability Analysis

In this section we show that we can perform the stability analysis for the generic model after rewriting the interaction matrix in order to have a structure similar to the structure of the standard perspective interaction matrix.

A. The interaction matrix revisited

Let us denote $\eta_i = \rho_i/Z_i = \sqrt{1 + X_i^2/Z_i^2 + Y_i^2/Z_i^2}$, the coordinates of the image point can be rewritten as:

$$x_i = \frac{X_i/Z_i}{1+\xi\eta_i}$$

$$y_i = \frac{Y_i/Z_i}{1+\xi\eta_i}$$

By combining the two previous equations, it is easy to show that η_i is the solution of the following second order equation:

$$-\eta_i^2 + (x_i+y_i)^2(1+\xi\eta_i) + 1 = 0$$

with the following positive solution:

$$\eta_i = \frac{\sigma_i - \xi(x_i^2+y_i^2)}{1-\xi(x_i^2+y_i^2)} \quad (9)$$

Equation (9) shows that η_i can be computed as a function of image coordinates $\mathbf{s_i}$ and sensor parameter ξ. The matrix $\mathbf{A}_i$ can thus be rewritten as:

$$\mathbf{A}_i = \frac{1}{Z_i}\begin{bmatrix} -\frac{1+x_i^2(1-\xi(\sigma_i+\xi))+y_i^2}{\eta_i(\sigma_i+\xi)} & \frac{\xi x_i y_i}{\eta_i} & \frac{x_i\sigma_i}{\eta_i} \\ \frac{\xi x_i y_i}{\eta_i} & -\frac{1+x_i^2+y_i^2(1-\xi(\sigma_i+\xi))}{\eta_i(\sigma_i+\xi)} & \frac{y_i\sigma_i}{\eta_i} \end{bmatrix}$$

Note that, only the depth Z_i is unknown in the matrix $\mathbf{A}_i$. If we consider the $(2n\times 1)$ vector $\mathbf{s} = (\mathbf{s_1}, \mathbf{s_2}, \cdots, \mathbf{s_n})$, the corresponding $(2n\times 6)$ interaction matrix is $\mathbf{L}(\mathbf{z},\mathbf{s}) = (\mathbf{L}_1, \mathbf{L}_2, \cdots, \mathbf{L}_n)$ and the time derivative of $\mathbf{s}$ is:

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

Matrix $\mathbf{L}(\mathbf{z},\mathbf{s})$ can be decomposed into two $(2n\times 3)$ sub-matrices:

$$\mathbf{L}(\mathbf{z},\mathbf{s}) = [\mathbf{A}(\mathbf{z},\mathbf{s})\ \mathbf{B}(\mathbf{s})]$$

where $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ and $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_n)$. 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:

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

is a (2n×2n) diagonal matrix containing the depth distribution $\mathbf{z}$.

B. stability analysis

Let us suppose that the camera and mirror parameters are perfectly known (i.e. $\widehat{\mathbf{K}} = \mathbf{K}$, $\widehat{\mathbf{M}} = \mathbf{m}$). 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 $\widehat{\mathbf{z}}$:

$$\widehat{\mathbf{L}}(\widehat{\mathbf{z}},\mathbf{s}) = [\ \mathbf{A}(\widehat{\mathbf{z}},\mathbf{s})\quad \mathbf{B}(\mathbf{s})\ ]$$

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 $\mathbf{A}(\mathbf{z},\mathbf{s})$:

$$\mathbf{A}(\widehat{\mathbf{z}},\mathbf{s}) = \mathbf{D}(\widehat{\mathbf{z}})\ \mathbf{C}(\mathbf{s}) = \mathbf{\Gamma}^{-1}(\widehat{\mathbf{z}},\mathbf{z})\mathbf{A}(\mathbf{z},\mathbf{s}) \quad (10)$$

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

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

From equation (10) one can deduce that:

$$\mathbf{L}(\mathbf{z},\mathbf{s}) = [\ \mathbf{\Gamma}\mathbf{A}(\widehat{\mathbf{z}},\mathbf{s})\quad \mathbf{B}(\mathbf{s})\ ]$$

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

$$\mathbf{L} = \widehat{\mathbf{L}} + \mathbf{\Delta}[\ \mathbf{A}(\widehat{\mathbf{z}},\mathbf{s})\quad \mathbf{0}\ ]$$

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

$$\mathbf{Q} = -\widehat{\mathbf{L}}^+\mathbf{L} = -\mathbf{I} - \left[\ \widehat{\mathbf{L}}^+\mathbf{\Delta}\widehat{\mathbf{A}}\quad \mathbf{0}\ \right]$$

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:

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

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

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

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:

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

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, $\psi_j = 1$. 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 $\mathbf{F}$.

C. Necessary and sufficient conditions

The eigenvalues of $\mathbf{F}$ are the roots of the characteristic polynomial:

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

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:

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

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). 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. Similarly to [13] we can also find simple sufficient conditions in order to obtain an approximation of the robustness domain.

IV. Simulation Results

The stability results obtained in the previous section have been tested with simulations for an eye-in-hand configuration. Two 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 show the influence of the errors of the depth distribution on the visual servoing.

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:

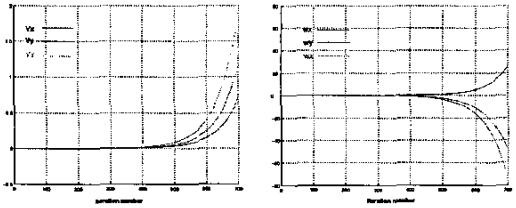
$$Z_i = d/\mathbf{n}^{\top}\mathbf{m}$$

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 $\widehat{d}$:

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

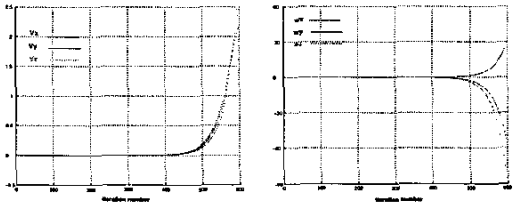
then: $\gamma_i = \frac{\widehat{Z}_i}{Z_i} = \frac{\widehat{d}}{d}\frac{\mathbf{n}^{\top}\mathbf{m}_i}{\widehat{\mathbf{n}}^{\top}\mathbf{m}_i}$ and $\psi_i = \frac{\gamma_i}{\gamma_j} = \frac{\mathbf{n}^{\top}\mathbf{m}_i}{\widehat{\mathbf{n}}^{\top}\mathbf{m}_i}\frac{\widehat{\mathbf{n}}^{\top}\mathbf{m}_i}{\mathbf{n}^{\top}\mathbf{m}_i}$
As expected, the stability of the visual servoing does not depends on $\widehat{d}$ but only on $\widehat{\mathbf{n}}$. Figure 2 to Figure 5 show the stability regions for conventional, parabolic and hyperbolic cameras 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 are cases not considered here.

First, note that all central cameras have similar stability region. Thus one can not expect to increase significantly the stable region by changing the type of central camera. When considering 3 image points (see Figure 2), the corresponding stable region is not so wide. Note that, adding a point inside the triangle defined by the others points (see Figures 3) only slightly modifies the stability region. 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 4, the stability region in green is even reduced. 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 5 the stability



(a) Translation velocity (b) Rotation velocity

Fig. 6. Unstable image-based visual servoing with conventional camera



(a) Translation velocity (b) Rotation velocity

Fig. 7. Unstable image-based visual servoing with parabolic camera

region considerably increase. However, even in this very favorable case, there exist an important red instability region.

In the second set of experiments (refer to Figures 6 and 7), an image-based positioning task of a six degrees of freedom manipulator has been simulated. The initial camera displacement is very small $t = -[0,001\ 0,001\ 0,001]$ meters and $r = -[0,48\ 0,96\ 1,44]$ degrees. The true depths are $\mathbf{z}^* = (0.9443\ 0.8470\ 0.9417\ 0.8498\ 0.9540\ 0.9412)\,m$ while the estimated depths distribution is $\widehat{z} = (0.9500\ 0.8500\ 0.9500\ 0.8600\ 0.9460\ 0.9300)\,m$. Despite the maximal error on the estimated depths is only $1,2\%$ of the true depth, one eigenvalue is positive for all central cameras (due of lack of space only figures for conventional and parabolic cameras are present). Thus, even starting very close to the reference position, after iteration 350 for the conventional camera (resp. 300 for the parabolic camera) the translation and the rotation errors start to grow. In the beginning, the control law seems to be stable since the other dominant eigenvalues have negative real part.

V. Conclusions

In this paper, we have shown that extreme care must be taken when approximating 3D parameters of a target for image-based visual servoing with central catadioptric cameras. Indeed, for these sensors the stability region in presence of errors on 3D parameters is not very large. It has been noticed that for all these sensors the stability region is similar and thus one can not expect to enlarge it by simply changing the type of central camera used.

References

[1] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *Int. Journal of Computer Vision*, 35(2):1–22, November 1999.

[2] J. P. Barreto, F. Martin, and R. Horaud. Visual servoing/tracking using central catadioptric images. In *8th Int. Symp. on Experimental Robotics, ISER'02*, pages 863–869, Bombay, India, July 2002.

[3] R. Benosman and S. Kang. *Panoramic Vision*. Springer Verlag, 2000.

[4] P. Blaer and P.K. Allen. Topological mobile robot localization using fast vision techniques. In *IEEE Int. Conf. on Robotics and Automation, ICRA'02*, pages 1031–1036, Washington, USA, May 2002.

[5] D. Burshka, J. Geiman, and G. Hager. Optimal landmark confi guration for vision based control of mobile robot. In *IEEE Int. Conf. on Robotics and Automation, ICRA'03*, pages 3917–3922, Tapei, Taiwan, September 2003.

[6] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. *The Confluence of Vision and Control, LNCIS Series, Springer Verlag*, 237:66–78, 1998.

[7] G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino. A switching control law for keeping features in the fi eld of view in eye-in-hand visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'03*, pages 3929–3934, Taipei, Taiwan, September 2003.

[8] N. J. Cowan, J. D. Weingarten, and D. E. Koditschek. Visual servoing via navigation functions. *IEEE Trans. on Robotics and Automation*, 18(4):521–533, August 2002.

[9] 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.

[10] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical implications. In *European Conf. on Computer Vision, ECCV'00*, volume 29, pages 159–179, Dublin, Ireland, May 2000.

[11] 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.

[12] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):238–250, April 1999.

[13] E. Malis and P. Rives. Robustness of image-based visual servoing with respect to depth distribution errors. In *IEEE Int. Conference on Robotics and Automation, ICRA'03*, Taipei, Taiwan, September 2003.

[14] Y. Mezouar and F. Chaumette. Path planning for robust image-based control. *IEEE Trans. on Robotics and Automation*, 18(4):534–549, August 2002.

[15] Y. Mezouar and F. Chaumette. Optimal camera trajectory with image-based control. *Int. Journal of Robotics Research*, 22(10):781–804, October 2003.

[16] A. Paulino and H. Araujo. Multiple robots in geometric formation: Control structure and sensing. In *Int. Symp. on Intelligent Robotic Systems*, pages 103–112, Reading, UK, July 2000.

[17] E. Malis S. Benhimane. Vision-based control with respect to planar and non-planar objects using a zooming camera. In *IEEE Int. Conf. on Advanced Robotics, ICAR'04*, pages 863–869, July 2003.

[18] O. Tahri and F. Chaumette. Application of moment invariants to visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'03*, pages 4276–4281, Taipeh, Taiwan, May 2003.

[19] R. Vidal, O. Shakernia, and S. Sastry. Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation. In *IEEE Int. Conf. on Robotics and Automation, ICRA'03*, Taipei, Taiwan, September 2003.

[20] N. Winter, J. Gaspar, G. Lacey, and J. Santos-Victor. Omnidirectional vision for robot navigation. In *IEEE Workshop on Omnidirectional Vision, OMNIVIS'00*, pages 21–28, South Carolina, USA, June 2000.

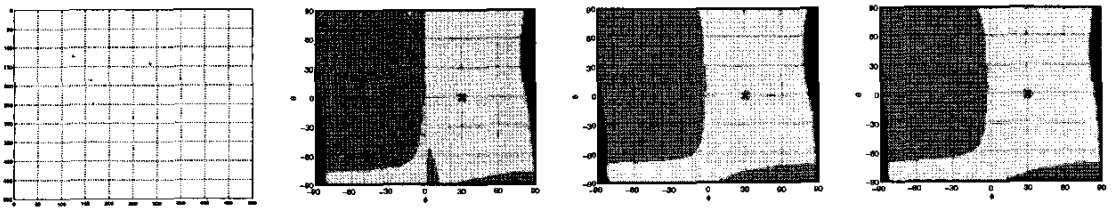


Fig. 2. (a) Image points with conventional camera; Stability regions: (b) conventional, (c) parabolic, (d) hyperbolic cameras

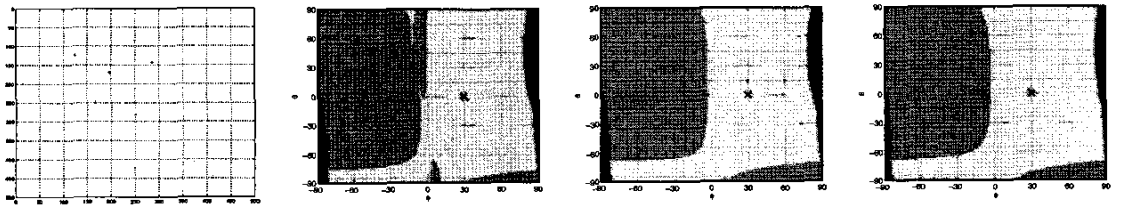


Fig. 3. (a) Image points with conventional camera; Stability regions: (b) conventional, (c) parabolic, (d) hyperbolic cameras

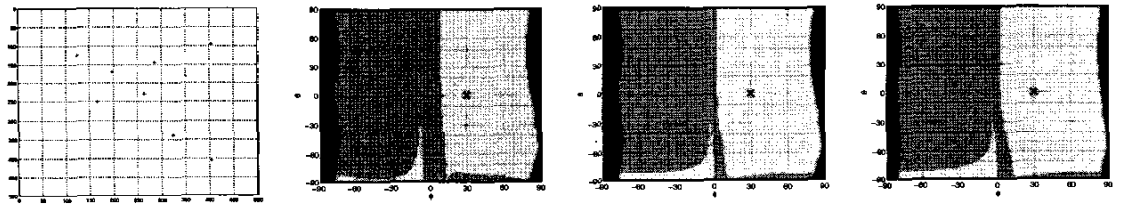


Fig. 4. (a) Image points with conventional camera; Stability regions: (b) conventional, (c) parabolic, (d) hyperbolic cameras

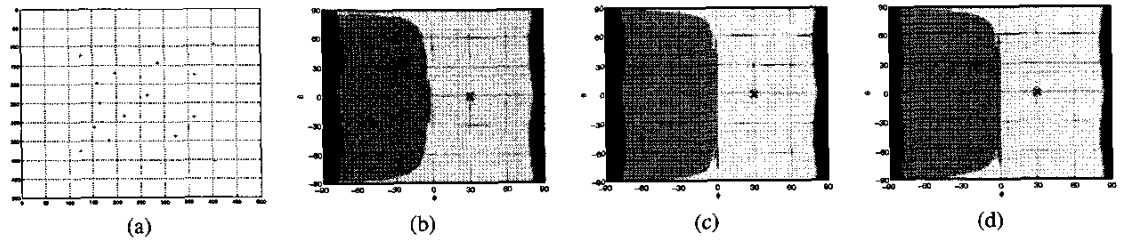


Fig. 5. (a) Image points with conventional camera; Stability regions: (b) conventional, (c) parabolic, (d) hyperbolic cameras